UNIVERSITY OF GENOA

Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics and Maternal and Child Sciences

PhD in Pediatric Sciences

Curriculum: Muscular, neurodegenerative and metabolic diseases of developmental age

Exome Sequencing as a Dynamic and Integrated Diagnostic Platform in Pediatric Neurological and Neurodevelopmental Disorders: A Real-World Single-Center Analysis

Relator:
Professor Pasquale Striano

Candidate:
Dr. Elisabetta Amadori

Supervisors:
Dr. Francesca Madia
Dr. Michele Iacomino

Academic year 2024-2025

Abstract

Background: Next-generation sequencing has reshaped pediatric neurogenetics, where neurological and neurodevelopmental disorders exhibit overlapping phenotypes and substantial genetic heterogeneity. Whole-exome sequencing is widely adopted as a first- or second-tier test, though real-world performance may differ from research-based estimates.

Rationale and Objectives: Reported diagnostic yields often derive from selected or multicenter cohorts and may not reflect routine tertiary-care practice. This single-center exploratory analysis aimed to estimate real-world exome sequencing (ES) diagnostic yield, compare panel-restricted and agnostic exome-wide interpretation, evaluate the impact of family configuration, and quantify the incremental contribution of systematic reanalysis and read-depth–based copy number variants (CNVs) detection.

Materials and Methods: We analyzed 817 pediatric patients who underwent ES at IRCCS Giannina Gaslini Institute (November 2021–October 2025). Testing was performed as singletons (43.9%), trios (51.3%), or duos (4.8%), and interpreted using either virtual phenotype-driven panels (37.6%) or an agnostic exome-wide approach (62.4%). Reanalysis was conducted in selected non-diagnostic cases. Read-depth–based CNV detection was implemented in 2024 in a defined subset with orthogonal validation.

Results: Baseline diagnostic yield was 23.5% (192/817), significantly higher with agnostic than panel-based interpretation (29.6% vs 13.4%, $p < 0.001$). In the overall cohort, trio testing outperformed singletons (29.1% vs 15.0%). Reanalysis (10.0% of cases) achieved a 27.2% incremental yield among initially negative individuals (22/81), corresponding to a +2.7% absolute increase. In the agnostic-only strict-exome subset, incremental yield was 24.3% (+1.8%), mainly driven by periodic bioinformatic reanalysis (55.6%) and updated clinical reassessment (33.3%). CNV calling ($n = 215$) provided 12 additional diagnoses (5.6%) and secondary findings were identified in 6.4%.

Discussion and Conclusions: Despite the inherent limitations of this single-center exploratory analysis, ES demonstrated superior performance with agnostic exome-wide interpretation and trio-based testing in a real-world tertiary pediatric cohort. Integration of CNV analysis and systematic reanalysis further increased diagnostic yield, supporting the role of ES as a dynamic and expandable platform for individualized, clinically integrated genomic diagnostics, particularly in neurological and neurodevelopmental disorders.

Summary

1 Introduction

1.1 Next-Generation Sequencing in Pediatric Neurological and Neurodevelopmental Disorders

Next-generation sequencing (NGS) has profoundly reshaped medical genetics, enabling parallel interrogation of thousands of genes and transforming the diagnostic approach to genetically heterogeneous disorders. This paradigm shift is particularly impactful in pediatric neurology, where neurodevelopmental disorders (NDDs) frequently present with complex, overlapping, and etiologically diverse phenotypes. Whole-exome sequencing (WES) has emerged as a central diagnostic modality in rare diseases, given that coding regions—although comprising only 1–2% of the genome—harbor a substantial proportion of pathogenic variants responsible for Mendelian conditions [1].

NDDs encompass a broad spectrum of conditions characterized by developmental impairments affecting personal, social, academic, or occupational functioning. These include intellectual disability (ID), global developmental delay (GDD), autism spectrum disorder (ASD), and early-onset epileptic encephalopathies (EE), and may overlap with epilepsy and cerebral palsy. When milder phenotypes are considered, NDDs affect up to 1 in 6 children, with a significant proportion presumed to have a genetic etiology [2]. The marked phenotypic variability and locus heterogeneity underlying these disorders strongly support the adoption of genome-wide diagnostic strategies.

Accordingly, international guidelines increasingly recommend Exome Sequencing (ES)/Genome Sequencing (GS) as a first- or second-tier test in individuals with ID, GDD, and/or congenital anomalies, demonstrating superior diagnostic yield compared with traditional sequential testing and a measurable impact on clinical management [3]. The genetic architecture of NDDs is complex, encompassing de novo variants, autosomal recessive (AR) inheritance (homozygous or compound heterozygous), mosaicism, and structural alterations [3-4]. In pediatric neurology, implementation of ES has substantially reduced the duration and fragmentation of the so-called “diagnostic odyssey”, particularly in patients with non-specific or multisystemic phenotypes [5].

Clinical evidence confirms its utility. In a cohort of 868 children undergoing clinical ES, a molecular diagnosis was achieved in 27% of cases, increasing to 34% in ID and 32% in GDD subgroups [6]. Meta-analyses report overall diagnostic yields ranging from approximately 36% to 43% across heterogeneous rare-disease cohorts, although variability reflects differences in phenotyping depth, case selection, and analytical

pipelines [7-8]. These data underscore both the clinical value of ES and the methodological determinants of its performance.

1.2 Technical framework and interpretative governance

WES comprises three core phases: enrichment of coding regions via hybrid capture, massively parallel sequencing, and bioinformatic processing—including alignment, variant calling, annotation, and classification [9]. Analytical pipelines detect single nucleotide variants (SNVs), small insertions/deletions (indels), and, depending on design, copy number variants (CNVs) inferred from read-depth data [9]. Variant interpretation is performed according to standardized American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) criteria—pathogenic (P), likely pathogenic (LP), variant of uncertain significance (VUS), likely benign (LB), and benign (B)—thereby ensuring consistency and inter-laboratory reproducibility [10].

Clinical implementation of ES necessitates structured governance of ethical and communicative aspects, particularly regarding VUS and secondary findings. The ACMG defines secondary findings as P or LP variants unrelated to the primary indication but identified in genes associated with actionable conditions; recommendations for their reporting are periodically updated to support responsible genomic practice [11]. In pediatric settings, these considerations are amplified by patient age, familial implications, and the need for multidisciplinary counseling.

1.3 Analytic strategies: panel-based and agnostic approaches

Clinical interpretation of ES data generally follows two strategies: panel-based (virtual gene panel) and agnostic (exome-wide) approaches [12–15].

In the panel-based approach, ES data are generated but interpretation is restricted to a predefined gene set selected according to phenotype, Human Phenotype Ontology (HPO) terms [16], clinical suspicion, or curated disease panels [12,17]. This strategy reduces interpretative burden and limits the number of variants requiring ACMG/AMP classification [12,18], while potentially decreasing the likelihood of incidental findings, though not eliminating them [11,19–20].

Conversely, the agnostic approach interrogates the entire exome without predefined gene restrictions [12]. Variant prioritization integrates allele frequency thresholds in

population databases [21], predicted functional impact, segregation analysis (e.g., trio-based), phenotype-driven algorithms [22], and curated gene–disease validity frameworks [23–25]. This strategy is particularly advantageous in highly heterogeneous or multisystemic phenotypes and facilitates identification of emerging gene–disease associations [26]. Many centers adopt tiered workflows that combine both approaches, balancing diagnostic yield, interpretative sustainability, and clinical governance [18,20].

1.4 Family configuration and diagnostic performance

Diagnostic yield is influenced not only by analytical strategy but also by family configuration [7]. ES may be performed as singleton (proband only), duo (proband plus one parent), or trio (proband plus both parents). Trio analysis improves detection of de novo variants and clarifies inheritance patterns, enhancing interpretative precision. However, it increases costs, involves testing of unaffected individuals with potential incidental findings, and may fail to identify inherited dominant variants in the context of reduced penetrance or variable expressivity [27–28].

1.5 Limitations and complementary strategies

Despite its established clinical utility, ES retains intrinsic limitations. By targeting primarily coding regions, it does not systematically interrogate regulatory and deep intronic sequences [12]. Additionally, heterogeneous coverage and capture-related biases—particularly in GC-rich, repetitive, or low-mappability regions—may reduce analytical sensitivity and increase false-negative rates [12,29].

Although not originally optimized for structural variant detection, advances in read-depth–based algorithms (e.g., GATK-gCNV, ExomeDepth, XHMM, CANOES, CODEX2, CLAMMS, CNVkit) now enable clinically meaningful identification of multi-exonic Copy Number Variants (CNVs) in adequately covered regions [30–31]. Integration of CNV-calling pipelines into routine workflows has increased diagnostic yield and resolved a proportion of previously unsolved cases without immediate escalation to whole-GS [32].

Finally, periodic data reanalysis has emerged as a critical complementary strategy. The continuous expansion of gene–disease knowledge, refinement of annotation resources, and updates in classification frameworks justify systematic reassessment of negative or inconclusive cases. The ACMG recognizes reanalysis as an integral component of

responsible genomic practice [33], and empirical evidence demonstrates incremental diagnostic gains without additional sequencing [34].

2 Rationale and Objectives

ES is now established as a first-tier diagnostic tool in pediatric neurology and NDDs, however several aspects of its real-world performance remain insufficiently characterized. Most published diagnostic yield estimates derive from multicenter consortia or highly selected research cohorts and may not fully reflect the heterogeneity, operational constraints, and interpretative dynamics of routine tertiary-care practice. A context-specific evaluation within a single pediatric center over a defined recent time frame is therefore warranted to provide a realistic estimate of its clinical impact.

In this framework, the present exploratory analysis assesses the diagnostic yield achieved over the last 3–4 years in routine clinical practice and examines key methodological determinants of performance. Particular attention is given to the analytical strategies adopted for exome interpretation, comparing—within methodological constraints—phenotype-driven panel-based approaches with gene-agnostic exome-wide interrogation in order to clarify their relative and incremental diagnostic contributions in a real-world setting.

Recognizing the dynamic evolution of genomic knowledge, the analysis further quantifies the additional yield obtained through systematic reanalysis, including extension from initial panel-based interpretation to exome-wide assessment, thereby addressing the temporal dimension of genomic diagnostics. In parallel, it evaluates the contribution of CNV detection following implementation of a read-depth–based algorithm in a defined patient sub-cohort, estimating the added diagnostic value of expanded bioinformatic pipelines.

Finally, given the marked clinical heterogeneity of pediatric neurological and neurodevelopmental disorders—including EE, GDD, ID, movement disorders, neuromuscular diseases, and complex syndromic phenotypes—the analysis explores phenotype-specific diagnostic probabilities to inform patient selection, pre-test counseling, and prioritization strategies within tertiary-care genomic medicine.

3 Material and Methods

3.1 Patient cohort

The patient cohort comprised 817 individuals who underwent ES (either panel-based or agnostic approach) as part of their routine diagnostic work-up at the Medical Genetics Unit of IRCCS Giannina Gaslini Institute between November 2021 and October 2025.

Pediatric patients presenting with neurological signs or symptoms and/or neurodevelopmental disorders who were referred to IRCCS Giannina Gaslini Institute for comprehensive clinical and genetic diagnostic evaluation—primarily through the Units of Pediatric Neurology and Neuromuscular Diseases and Child and Adolescent Neuropsychiatry, or following admission to other hospital departments for concomitant acute conditions—were considered in this exploratory analysis.

Clinical presentations were heterogeneous, with patients exhibiting a broad spectrum of clinical features and variable combinations thereof. Patients who underwent fast-track targeted gene panel testing were excluded.

For genetic assessment, clinical data were collected by the referring specialist or clinical geneticist and recorded using a structured checklist. The available family configuration for genetic analysis (singleton, proband plus one parent, or trio) was specified and documented for each case.

Written informed consent was obtained prior to initiating genetic testing. Consent was signed for the proband and, when parental samples were included in the analysis, separately by each parent for their own genetic testing. The consent form also included specific information regarding the possibility of receiving secondary findings.

Subsequently, each reported clinical feature was translated into HPO terms to ensure standardized phenotypic characterization and was then encoded using PhenoTips terminology. Clinical and genetic data were systematically entered into an internal institutional database of IRCCS Giannina Gaslini Institute.

The data used for the current analysis were fully anonymized and are not in any way traceable to individual patients.

This work was conducted in accordance with institutional policies and ethical standards and was based on collected clinical data obtained during routine diagnostic practice.

3.2 Exome Sequencing and Analytical Workflow

ES was performed in all probands and available family members according to the family-based configuration, using genomic DNA extracted from peripheral blood or, in selected cases, alternative tissue sources. Sequencing procedures were technically uniform across the cohort; however, data interpretation followed two distinct analytical strategies according to the clinical request of the referring specialist, reflecting real-world clinical implementation.

In the first cohort, ES data were analyzed using an agnostic exome-wide approach without prior gene restriction. Variant prioritization was performed across the entire coding genome, enabling identification of pathogenic variants in both established and newly described disease genes. Structured reanalysis of exome-wide data was undertaken when clinically indicated or following significant advances in gene–disease knowledge, supporting a longitudinal interpretative model.

In the second cohort, ES data were initially analyzed using a virtual gene panel strategy restricted to predefined phenotype-driven gene sets *(Supplementary Table S1)*, including:

(i) epilepsy and EE genes,

(ii) cortical malformation (CM)-associated genes, and

(iii) genes related to the mTOR signaling pathway.

In cases with negative panel-based results, escalation to unrestricted exome-wide reinterpretation was performed upon request of the referring specialist, implementing a tiered diagnostic framework within the same sequencing dataset.

From 2024 onward, CNV detection from exome data was incorporated into the clinical workflow through a dedicated bioinformatic pipeline. Consequently, CNV analysis was available only for a later, numerically unbalanced subset of patients (cases sequentially numbered 603-817), irrespective of the primary analytical strategy.

3.3 Variant Interpretation and Reporting

Variants were filtered for rarity, predicted functional impact, segregation consistency, and clinical annotation, and classified according to ACMG/AMP criteria [10]. P and LP variants were considered diagnostic. In selected cases, VUS were considered potentially contributory when supported by strong genotype–phenotype concordance and convergent in silico evidence.

Putative CNVs identified through exome-based read-depth analysis were confirmed using orthogonal methods, including array comparative genomic hybridization (array-CGH) or, in selected cases, multiplex ligation-dependent probe amplification (MLPA).

All findings were interpreted within a multidisciplinary framework involving laboratory geneticists and clinical specialists and were communicated through structured clinical reports. Suggestive VUS, when identified, were included in the clinical report. Secondary findings were disclosed according to patient preferences documented in pre-test informed consent.

Informed consent for the publication of anonymized clinical and genetic data was obtained from all participants.

Detailed bioinformatic procedures are reported in the *Supplementary Methods* section.

3.4 Dataset Structure and Variable Definitions

The analytical dataset was structured with one row per tested individual. Key variables captured sex, clinical features, prior genetic investigations, requested analytic approach (panel-based vs agnostic), family configuration (singleton/duo/trio), CNV-calling availability, baseline diagnostic status, molecular findings (gene/CNV, zygosity, inheritance mode, parental origin), ACMG classification, reanalysis indicators (reanalysis performed, type of reanalysis, changes in approach, and results after reanalysis), and secondary findings.

Diagnostic yield was defined as the proportion of patients with a positive result at baseline. Incremental diagnostic yield from reanalysis was defined as the proportion of additional diagnoses identified among baseline non-diagnostic cases that underwent reanalysis.

Reanalysis, when performed, was categorized as: (i) periodic bioinformatic reinterpretation, (ii) clinically triggered reinterpretation, (iii) complementary functional analysis, (iv) family extension (singleton→duo/trio or duo→trio), and/or (v) escalation from panel-based to agnostic exome-wide interpretation. Outcomes were recorded as diagnostic or non-diagnostic.

CNV-related analyses were restricted to the subset of cases with CNV calling performed, and denominators were explicitly reported.

Time-trend analyses were not performed due to incomplete capture of report dates.

3.5 Outcomes

The primary outcome was baseline diagnostic yield.

Secondary outcomes included:

(i) diagnostic yield stratified by analytic approach, family configuration, and sex;

(ii) diagnostic yield in the agnostic-only strict-exome subset;

(iii) incremental diagnostic yield from reanalysis;

(iv) counts of diagnostic genes;

(v) CNV-related diagnostic yield;

(vi) descriptive cohort characteristics and prevalence of clinical signs and symptoms in the overall cohort and stratified by analytic approach and diagnostic yield;

(vii) VUS reporting rates at baseline and after reanalysis;

(viii) frequency of secondary findings;

(ix) selected illustrative case descriptions.

3.6 Statistical analysis

Analyses were primarily descriptive. Categorical variables were summarized as counts and percentages; continuous variables as mean (SD) or median (IQR), as appropriate.

Baseline diagnostic yield was reported with exact 95% confidence intervals (CIs), overall and stratified by analytic approach, family configuration, sex, and panel subgroup (when applicable).

Between-group comparisons used Fisher's exact or χ^2 tests as appropriate. McNemar's test was applied for paired categorical comparisons. Given the non-random allocation of analytic strategies, comparisons were interpreted as exploratory.

Incremental diagnostic yield from reanalysis was estimated among baseline non-diagnostic cases undergoing reanalysis, with exact 95% CIs. Reanalysis categories were summarized as non-mutually exclusive when applicable.

A dedicated sensitivity analysis restricted to the agnostic exome-only cohort was conducted to assess robustness and avoid workflow heterogeneity. This analysis excluded cases in which diagnostic changes were driven by a shift from panel-based to exome-wide interpretation or by new sample sequencing. We considered only reanalysis strictly related to the exome (e.g., periodic bioinformatics reanalysis, reanalysis triggered by

updated clinical information, functional study, or family extension from singleton to duo/trio or duo to trio).

CNV-related analyses were restricted to the CNV-called subset, reporting its denominator. Within this subset, results were further stratified by SNV/indel and CNV status to describe composite molecular diagnoses.

VUS reporting rates were summarized at baseline and post-reanalysis, and transitions were descriptively analyzed.

An exploratory multivariable logistic regression model was fitted with diagnostic outcome as the dependent variable and analytic approach, family configuration, sex, CNV-called status, and prior genetic testing as candidate predictors. Duos were excluded from the primary model to avoid unstable estimates due to small sample size, and a sensitivity model including duos was fitted for comparison.

Results were reported as odds ratios with 95% CIs. Statistical significance was set at two-sided $\alpha = 0.05$. Analyses were performed using R (version 4.5.2).

Additional methodological details are reported in the *Supplementary Methods* section.

3.7 Selection Bias and Methodological Considerations

Within this real-world, single-center and exploratory analysis, potential sources of selection bias and confounding were identified a priori. Allocation to analytical strategies (panel-based vs agnostic exome-wide interpretation) was not randomized but reflected routine clinical practice and temporal evolution of laboratory workflows, introducing potential indication bias and phenotype-related confounding. Specifically, the panel-based cohort was enriched for epilepsy/EE, whereas the agnostic approach encompassed a broader and more heterogeneous phenotypic spectrum, potentially influencing observed differences in diagnostic yield.

Family configuration (trio, duo, singleton) was similarly determined by clinical feasibility and parental availability, limiting comparability across groups. Furthermore, the later implementation of CNV calling, applied to a numerically unbalanced subset, introduces potential temporal bias and reduces internal uniformity of the cohort.

The multivariable logistic regression model assessing baseline diagnostic yield was therefore specified as exploratory and hypothesis-generating. In the absence of randomization and given limited subgroup sizes, effect estimates should be interpreted as associative rather than causal.

4 Results

4.1 Cohort characteristics

We analyzed 817 individuals. The cohort included 376 females (46.0%) and 441 males (54.0%).

Family configurations were singleton (359, 43.9%), trio (419, 51.3%), duo proband+mother (37, 4.5%), duo proband+father (2, 0.2%). Requested analytic approaches were agnostic exome-wide (510, 62.4%) and a virtual panel approach (307, 37.6%), which included EE panel (250, 30.6%), MTOR panel (32, 3.9%), and CM panel (25, 3.1%). The WES-only cohort (agnostic) with recorded exome results included 510 individuals, and was predominantly trio-based (413/510, 80.9%).

4.2 Baseline diagnostic yield

Overall baseline diagnostic yield was 192/817 (23.5%, 95% CI 20.3–26.2%). Yield was similar by sex (females 22.9% vs males 23.4%). By requested analytic approach, baseline yield was 29.6% for agnostic (151/510, 95% CI 25.1–33.2%), and 13.4% for virtual panel approach (41/307), specifically 12.8% for EE panel (32/250, 95% CI 8.9–17.6%), 28.1% for MTOR panel (9/32, 95% CI 13.7–46.7%), and 0/25 for CM panel (95% CI 0.0–13.7%) *(Figure 1)*. By family configuration, baseline yield was 15% in singletons (54/359), 29.1% in trios (122/419), 32.4% in proband+mother duos (12/37), and 1/2 in proband+father duos.

In the WES-only cohort, the overall diagnostic yield was 151/510 (29.6%). Stratified by family configuration, the yield was 27.1% in singletons (16/59), 29.5% in trios (122/413), and 34.2% in duos (13/38), including 12 proband-mother duos and 1 proband-father duo. A direct comparison of baseline positivity between agnostic and panel-based approaches showed higher yield for agnostic testing (29.6% vs 13.4%; Fisher's exact test $p < 0.001$). The composition of baseline positives by family configuration (singleton, duo, trio) within each approach is shown in *Figure 2*.

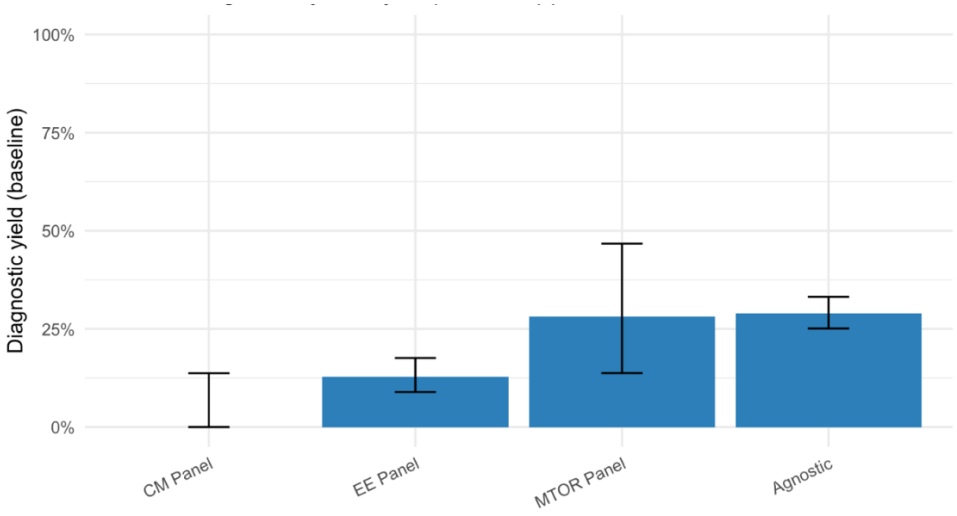


Figure 1. Baseline diagnostic yield by analytic approach

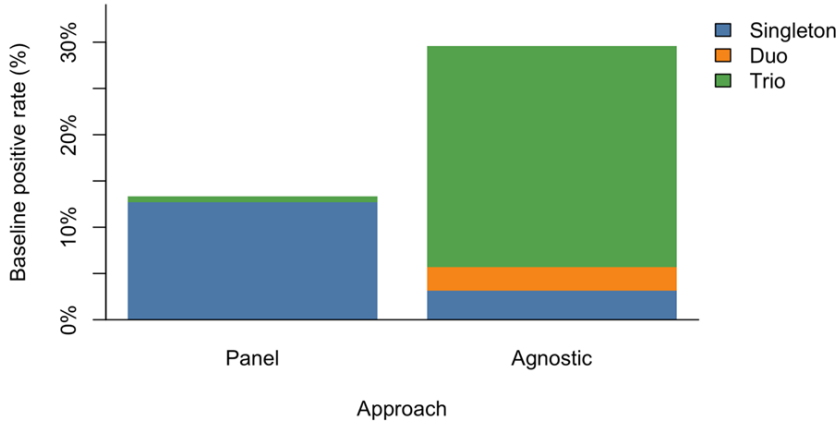


Figure 2. Baseline positive rate by approach with within-column distribution of family configuration (singleton, duo, trio). Agnostic: 151/510 (singleton n=16, duo n=13, trio n=122); Panel: 41/307 (singleton n=39, duo n=0, trio n=2)

4.3 Reanalysis and incremental yield

Eighty-two cases (10.0%) underwent reanalysis or new analyses; 81 were non-diagnostic at baseline, whereas one previously diagnosed case underwent reanalysis prompted by updated clinical information. Reanalysis identified 22 additional diagnoses. Incremental diagnostic yield among baseline non-diagnostic cases was 22/81 (27.2%, 95% CI 17.9–38.2%), corresponding to an absolute increase of 2.7% in the overall cohort (22/817). Reanalysis types included panel-to-exome escalation (41), other panels (4), periodic bioinformatics reanalysis (27), updated clinical reassessment (3), exome extension (5), complementary functional study (1), and WES sequencing performed on a different

biological specimen (1). In the agnostic-only, strict-exome reanalysis subset, incremental yield was 9/37 (24.3%, 95% CI 12.5–43.3%), corresponding to an absolute increase of 1.8% in the agnostic cohort *(Figures 3-5).*

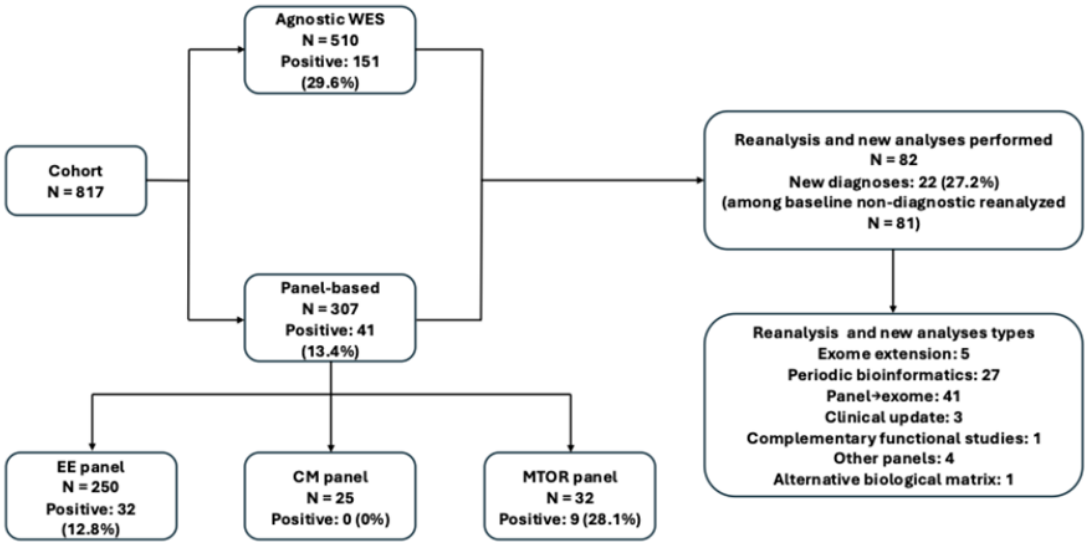


Figure 3. Flow of diagnostic yield by analytic approach and reanalysis in the overall cohort.

In *Figures 4-5*, variant classes according to ACMG criteria are also reported. All cases represented correspond to positive diagnostic outcomes; however, for graphical clarity, variants are displayed per patient. In AR conditions, biallelic variants assigned to the same ACMG class were grouped together, whereas combinations such as LP+VUS or P+LP reflect genotypes in which the two variants were classified into different ACMG categories. For the purposes of this analysis, these composite genotypes (biallelic variants) were treated as single diagnostic categories, while variant-level pathogenicity assessment will be presented separately.

Among baseline positives individuals with variant-class information available (n = 191; one case lacked classification data), variant-class composition was predominantly P-only (52.9%) and LP-only (35.1%), with smaller fractions of causative VUS-only (c-VUS-only, 8.4%), combined LP+P findings (1.0%), and LP+VUS findings (2.6%).

The distribution after reanalysis was similar. Among overall positives with available variant-class data (n = 213), class composition was P-only 52.6%, LP-only 34.7%, c-VUS-only 8.5%, LP+P 0.9%, and LP+VUS 3.3%.

Within the agnostic-only cohort, among baseline positive cases with variant-class data (n=150), class composition was P-only 52%, LP-only 34%, c-VUS-only 9.3%, LP+P 1.3%, and LP+VUS 3.3%. After reanalysis (positives n = 159), the distribution remained

comparable (P-only 51.6%, LP-only 33.3%, c-VUS-only 9.4%, LP+P 1.3%, LP+VUS 4.4%).

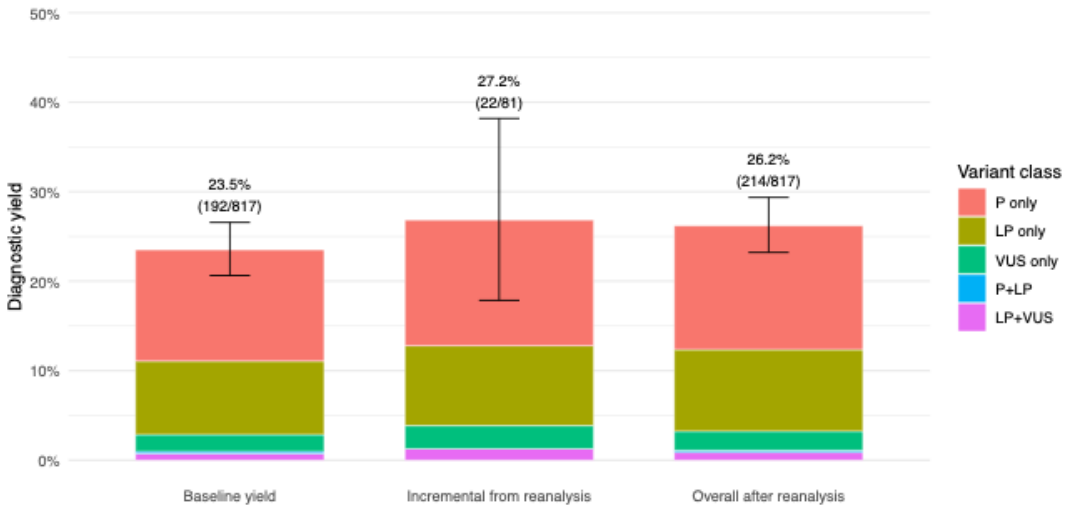


Figure 4. Diagnostic yield with variant-class composition in the overall cohort.

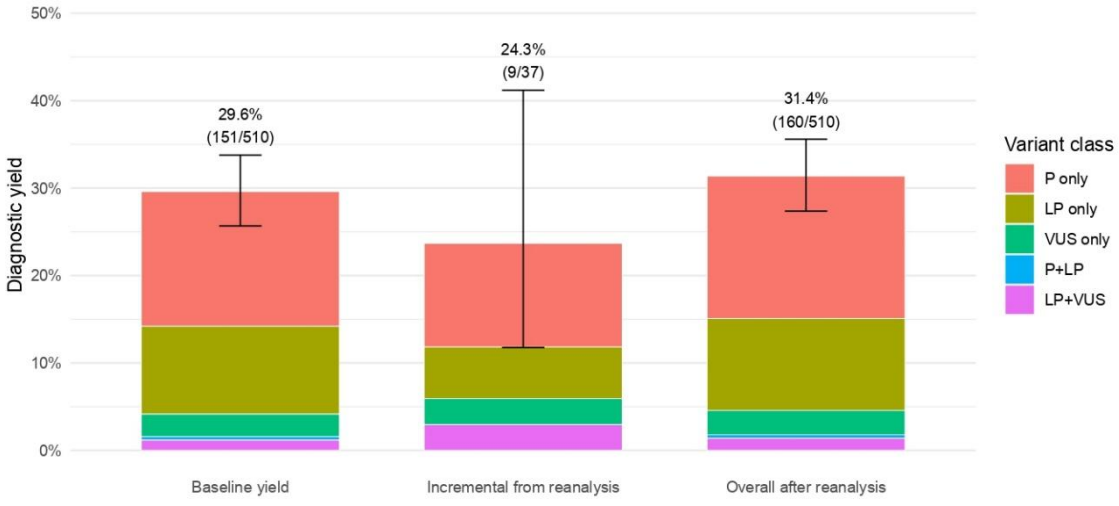


Figure 5. Diagnostic yield with variant-class composition in the agnostic exome-only subset.

Across reanalyzed cases with at least one positive outcome (n=22), diagnostic yield was primarily attributable to panel-to-exome escalation (59.1% of positives; 13/22), followed by periodic bioinformatics reanalysis (22.7%; 5/22), updated clinical reassessment (13.6%; 3/22), and functional study (4.5%; 1/22). The corresponding P/LP/VUS (VUS interpreted as causative) counts by cause were: panel-to- exome escalation (P=8, LP=6,

VUS=1), periodic bioinformatics reanalysis (P=2, LP=4, VUS=1), updated clinical reassessment (P=2, LP=1, VUS=2), and functional study (P=2). In the agnostic-only subset (n = 9 reanalysis positives), most diagnoses resulted from periodic bioinformatics reanalysis (55.6%; 5/9), followed by updated clinical reassessment (33.3%; 3/9), and functional study (11.1%; 1/9), with P/LP/VUS counts of 2/4/1, 2/1/2, and 2/0/0, respectively *(Figures 6-7)*.

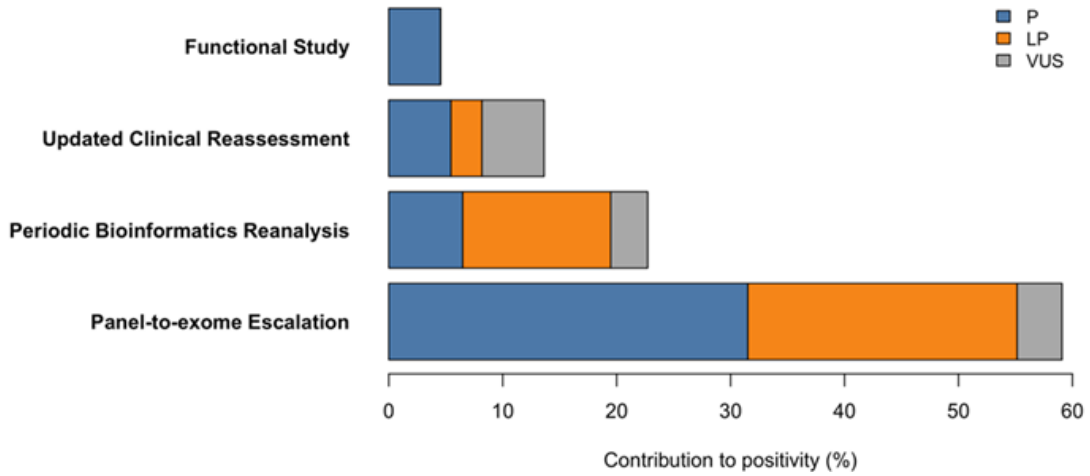


Figure 6. Overall cohort: contribution to positivity by reanalysis cause with stacked P/LP/VUS composition.

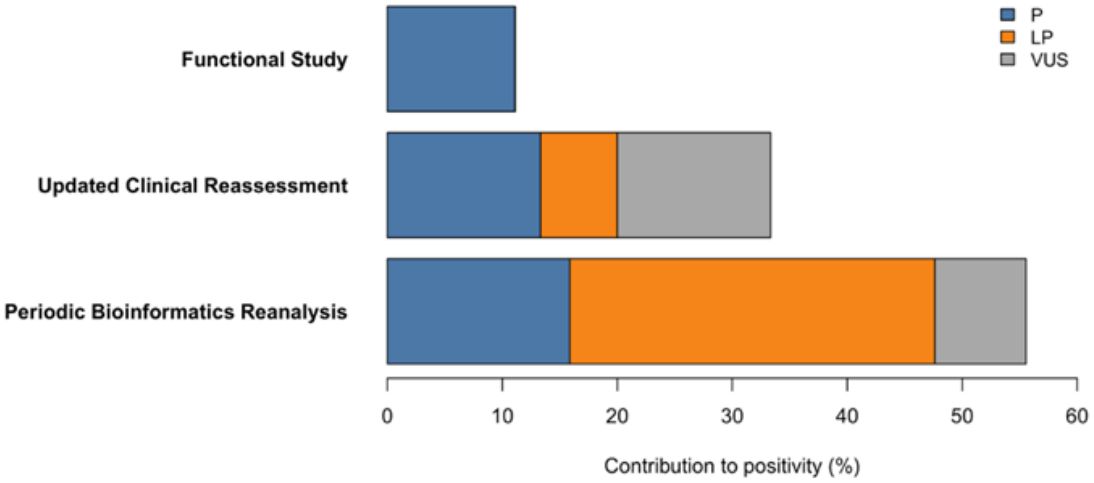


Figure 7. Agnostic-only cohort: contribution to positivity by reanalysis cause with stacked P/LP/VUS composition.

4.4 Genes and inheritance patterns

The most frequently identified genes at baseline were *PRRT2* (n= 10), *SCN2A* (n= 4), *GRIN2A* (n = 4), *CACNA1A* (n = 4) and *DEPDC5* (n = 3). Genes identified by each approach are summarized in *Table 1* and *Figure 8*. Some genes appear more than once, as they were identified in sibling pairs or within sibling groups. Specifically, *CHMP1A* was identified in dizygotic twins (one female and one male). *PRRT2* in two sisters and in a sibling pair (EE panel). *GDI1* in two brothers, *SORD* in three brothers, *IKBKG* in identical female twins, and *ASXL3* in two brothers.

New diagnoses after reanalysis most commonly involved *ANKRD11* (n = 2), with other genes observed once each *(Figure 9)*. The complete list of genes identified at baseline, together with those identified in the reanalysis, is reported in the supplementary materials *(Supplementary Table S2)*.

Table 1. Gene counts by approach at baseline

Approach	Gene	n (each)
Agnostic	*SATB2; SETD5; SORD*	3
	ADAR; ANKRD11; ARID2; ASXL3; ATL; CACNA1A; CHMP1A; CTNNB1; DEAF1; DEPDC5; GDI1; GRIN2B; IKBKG; ITPR1; KCNQ2; KIF1A; NSD1; PRRT2; SCN1A; SCN2A; SLC1A4; SLC5A6; ZEB2.	2
	ACAN; ADAMTS17; ADCY5; ADGRG1; ALG1; ARID1B; ARSA; ARX; ATAD3A; ATP6V0C; BTD; CDC45; CEP57; CHD2; CLCN2; CLCN4; CNOT3; COL4A1; CSF1R; CSNK2B; CUX2; DHDDS; DNM1; DNMT3A; DYNC1H1; ECHS1; EIF2AK2; EMC10; EP300; FBXW11; FGF8; FGFR3; FOXG1; GH1; GNB1; GPSM2; GRIN2A; HCN4; HIST1H1E; HPRT1; IGHMBP2; KARS1; KAT6A; KMT2A; KMT2B; MACF1; MAOA; MAPK1; MAST1; MBD5; MECP2; MED13L; MORC2; MTOR; NANS; NARS1; NF1; NFIA; NOTCH3; NPR2; NR2F1; OPHN1; PHIP; PIGN; POGZ; POLR3A; PPARG; PQBP1; PTCH1; PTEN; RAI1; RANBP2; SCN8A; SEMA6B; SLC2A1; SMARCE1; SMPD4; SOX2; SOX4; SPART; SYNGAP1; SZT2; TANC2; TCF4; TMX2; TOR1AIP1; VPS13B; WDR62; ZC4H2; ZDHHC9; ZNF462; ZNF699.	1
EE Panel	*PRRT2*	8
	GRIN2A	3
	CACNA1A; SCN2A	2
	ATP6V0A1; ATP6V0C; CHD2; CSNK2B; DEPDC5; DYRK1A; GABRA1; GABRG2; HCN1; KCNB1; KCNQ2; MECP2; PCDH19; SCN1A; SLC6A1; STXBP1; SYNGAP1.	1
MTOR Panel	*TSC1*	3
	TSC2	2
	NPRL2; PIK3CA; PTEN	1

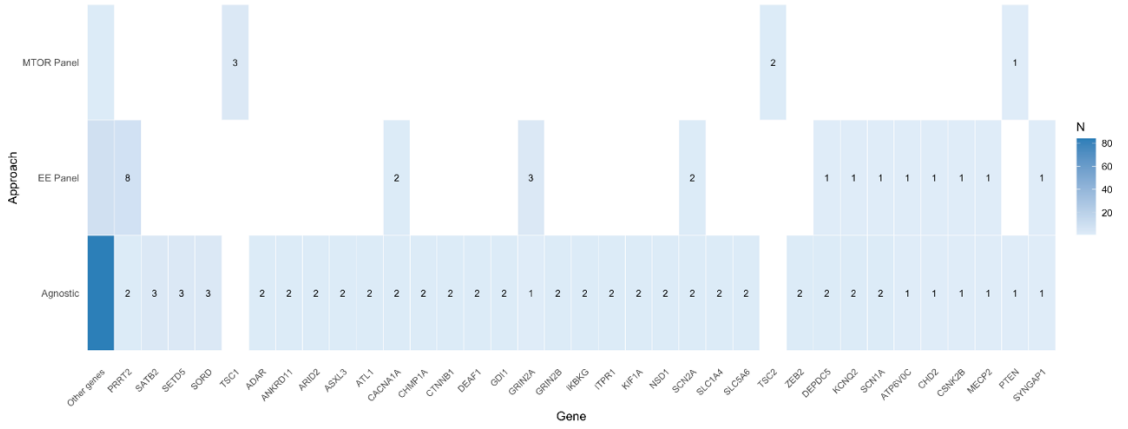


Figure 8. Genes identified by approach at baseline.

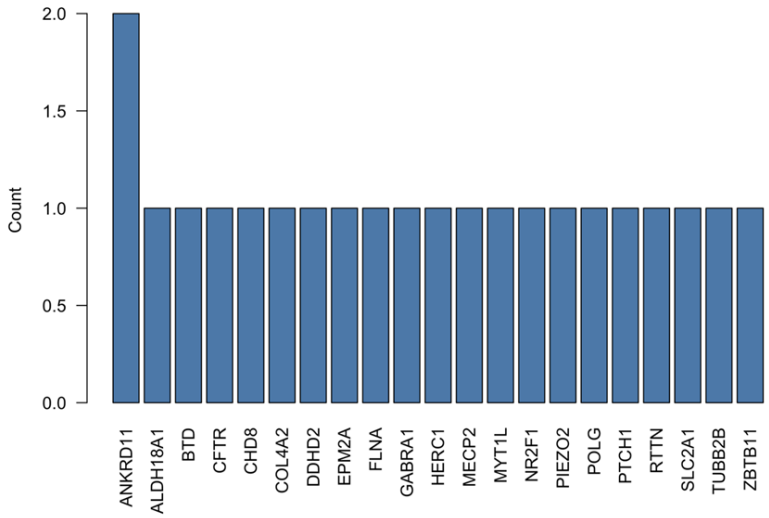


Figure 9. Genes identified after reanalysis.

Considering inheritance information (available for 193 baseline genetic observations and 23 reanalysis observations), baseline findings were predominantly autosomal dominant (AD, 73.1%), followed by autosomal recessive (AR, 19.2%) and X-linked (XL, 7.8%). Reanalysis showed a shift towards AR (47.8%), with AD at 43.5% and XL at 8.7%; overall (n = 216 genetic observations) remained skewed towards AD (69.9% AD, 22.2% AR, 7.9% XL). When examining subcategories of inheritance, the most frequent patterns were AD de novo (baseline 78/193 [40.4%], reanalysis 6/23 [26.1%], overall 84/216 [38.9%]), AD unspecified (baseline 44/193 [22.8%], reanalysis 1/23 [4.3%], overall 45/216 [20.8%]), and AR compound heterozygous (baseline 19/193 [9.8%], reanalysis 6/23 [26.1%], overall 25/216 [11.6%]) *(Figures 10-11)*. In the agnostic-only cohort, baseline inheritance (n = 152) was AD 67.1%, AR 24.3%, and XL 8.6%; reanalysis (n =

10) was AR 60.0% and AD 40.0% (no XL); overall (n = 162) was AD 65.4%, AR 26.5%, and XL 8.0%. The leading detailed categories in the agnostic cohort were AD de novo (baseline 73/152 [48.0%], overall 75/162 [46.3] %) and AR compound heterozygous (baseline 19/152 [12.5%], overall 25/162 [15.4%]), with AR homozygous also prominent (baseline 18/152 [11.8%], overall 18/162 [11.1%]), *(Figures 12-13)*.

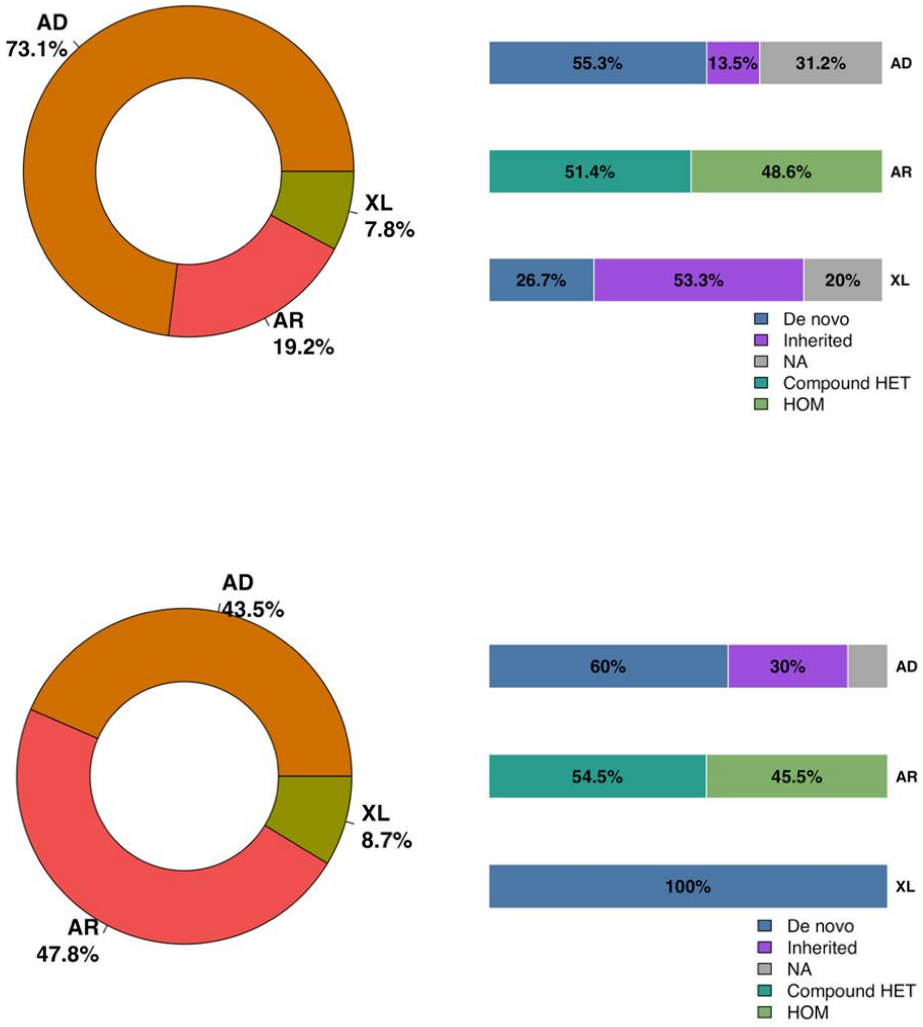


Figure 10. Mode of inheritance distributions: respectively, from top to bottom, baseline cohort, reanalysis cohort. Left: AD/AR/XL proportions; right: subtype inheritance breakdowns (100%).

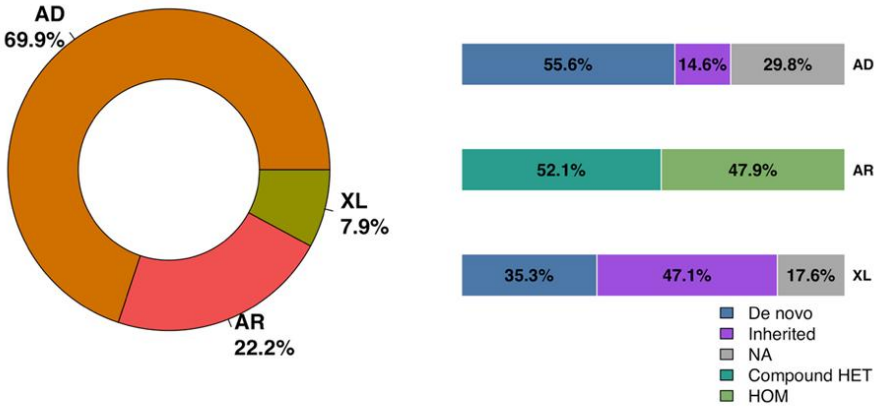


Figure 11. Mode of inheritance distributions: overall cohort (baseline + reanalysis). Left: AD/AR/XL proportions; right: subtype inheritance breakdowns (100%).

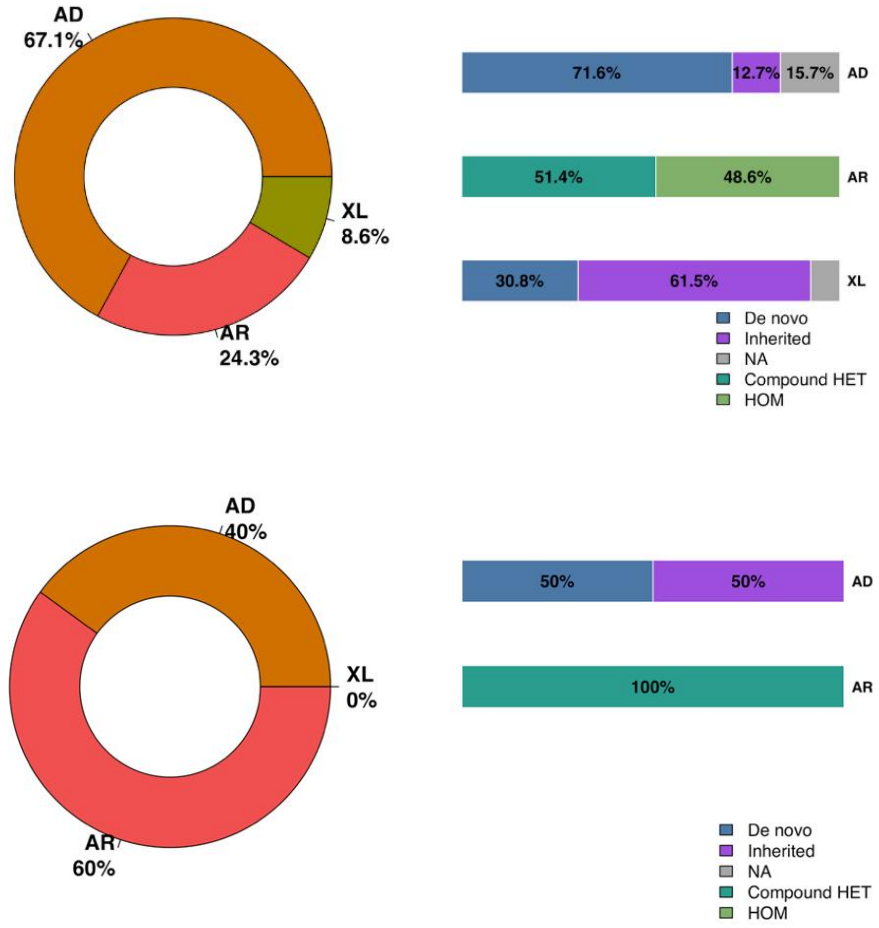


Figure 12. Mode of inheritance distributions: respectively, from top to bottom, baseline agnostic-only cohort and reanalysis agnostic-only cohort. Left: AD/AR/XL proportions; right: subtype breakdowns (100%).

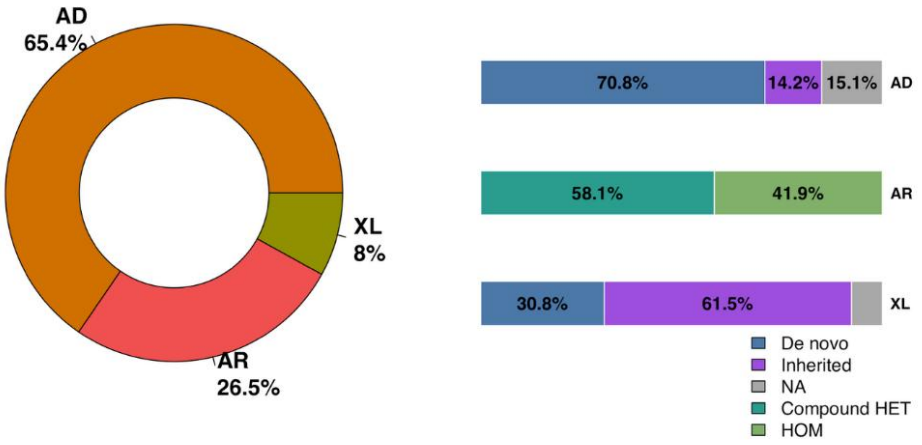


Figure 13. Mode of inheritance distributions: overall agnostic-only cohort (baseline + reanalysis). Left: AD/AR/XL proportions; right: subtype breakdowns (100%).

4.5 CNV findings

CNV calling was performed in 215 cases (panel + agnostic) with one case was excluded for technical reasons (panel cohort); 12 cases were CNV-positive (5.6%, 95% CI 2.9–9.5%), representing the incremental CNV diagnostic contribution within the CNV-called subset. Among 142 cases (agnostic-only) with both CNV calling and WES SNV/indel status available, SNV+/CNV+ occurred in 6 (4.2*%) (table 2)*, SNV-/CNV+ in 5 (3.5%) *(table 3)*, SNV+/CNV- in 36 (25.4%), and SNV-/CNV- in 95 (66.9%) *(Figures 14-15)*.

Table 2. CNV+/SNV+ combination

	CNV +	**SNV +**
Sex/Family-based approach	**Single-locus**	
M/trio	*EMC10*: deletion exone 1, paternal, pathogenic	*EMC10*, maternal
M/trio	*VPS13B*: deletion exones 4-7, maternal, pathogenic	*VPS13B*, paternal
	Double-hit	
F/trio	Deletion chr.2q14.1, de novo, pathogenic	*IKBKG*, maternal
F/trio	Deletion chr.2q14.1, de novo, pathogenic	*IKBKG*, maternal
M/trio	Duplication chr.22q11.2, maternal, pathogenic	*FGFR3*, maternal
M/trio	Deletion chr.17p12, maternal, pathogenic	*SATB2*, de novo

Single-locus: CNV and SNV/indel affecting the same gene. Double-hit: CNV and SNV/indel affecting different genes or chromosomal regions, consistent with dual molecular findings.

Table 3. SNV-/CNV+ combination

Sex/Family-based approach	Only CNV+
F/trio	Deletion chr.1p36.33-p36.23, de novo, pathogenic
F/trio	Duplication chr16p13.2, de novo, VUS
M/singleton	Duplication exones 1-5 in FGF12 (chr.3), de novo, pathogenic
F/trio	Deletion chr.18q12.1-12.2, de novo, not available (NA)
M/singleton	Duplication chr.15q11.2-13.3, NA, pathogenic

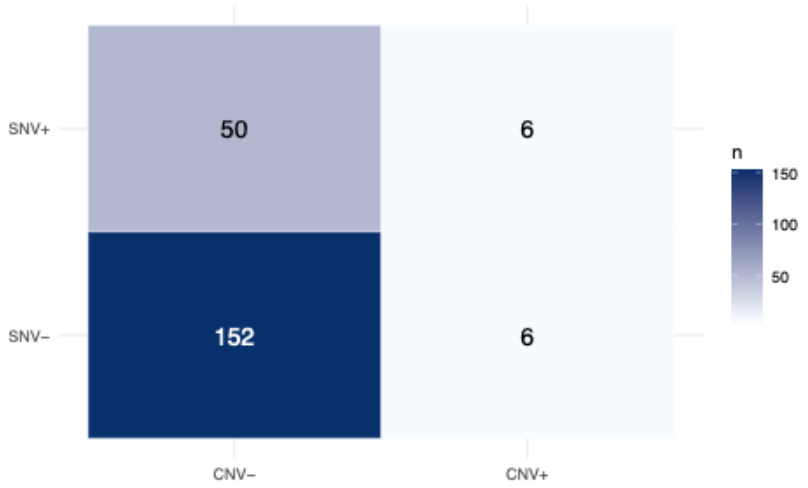


Figure 14. SNV/CNV combination matrix in the overall cohort, CNV-called subset.

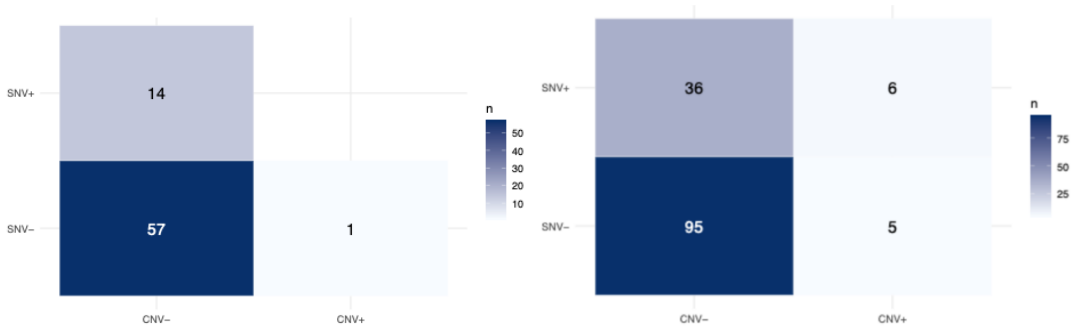


Figure 15. SNV/CNV combination matrix in the panel cohort (left) and in the agnostic-only cohort (right), CNV-called subset.

In the panel-based subset with CNV algorithm applied (n = 73), 1 case was CNV+ (1/73, 1,4%), 71 were CNV– (71/73, 97.3%), and 1 was excluded for technical reasons; SNV+/CNV+ was not observed, as the single CNV-positive case did not harbor a concomitant SNV. The single CNV-positive case consisted of a heterozygous four-exon

deletion in *TSC1*, confirmed by MLPA, identified in a male patient with cortical tubers on brain MRI.

4.6 Phenotype prevalence

In the overall cohort, the most prevalent clinical features were epilepsy (51.0%), brain MRI abnormalities (30.4%), developmental delay (23.1%), dysmorphism (20.3%), ID (15.5%), and EE (14.7%) *(Supplementary Figure S1)*. When stratified by baseline diagnostic yield *(Table 4)*, brain MRI abnormalities, developmental delay, dysmorphism, and hypotonia were more frequent among positive cases (e.g., brain MRI abnormalities 39.2% vs 27.7%; developmental delay 37.6% vs 18.9%; dysmorphism 33.3% vs 16.5%; hypotonia 18.0% vs 6.4%), whereas epilepsy and ASD were more frequent among non-diagnostic cases (54.9% vs 38.1% and 9.3% vs 3.7%, respectively; all $p<0.05$).

Phenotype prevalence differed by analytic approach: epilepsy, EE, and brain malformations were more frequent in the panel subset, whereas dysmorphism, developmental delay, brain MRI abnormalities, hypotonia, and ID were more frequent in the agnostic subset *(Table 5)*.

Table 4. Phenotype prevalence by baseline diagnostic status. p-values from Fisher's exact test.

Phenotype	Positive (n=189)	WT (n=625)	p
Epilepsy	72 (38.1%)	343 (54.9%)	6.3×10^{-5}
Brain MRI abnormalities	74 (39.2%)	173 (27.7%)	0.0038
Developmental delay	71 (37.6%)	118 (18.9%)	3.8×10^{-7}
Dysmorphism	63 (33.3%)	103 (16.5%)	1.6×10^{-6}
Intellectual disability	31 (16.4%)	95 (15.2%)	0.731
Epileptic encephalopathy	30 (15.9%)	89 (14.2%)	0.559
Brain malformations	22 (11.6%)	75 (12.0%)	1.000
Hypotonia	34 (18.0%)	40 (6.4%)	8.2×10^{-6}
Autism/ASD	7 (3.7%)	58 (9.3%)	0.013
Spasticity/spastic paraparesis	9 (4.8%)	16 (2.6%)	0.147
Visual impairment	10 (5.3%)	15 (2.4%)	0.054
Hearing loss	2 (1.1%)	7 (1.1%)	1.000
Neurodegenerative	3 (1.6%)	6 (1.0%)	0.441

Table 5. Phenotype prevalence by analytic approach (agnostic vs panel). p-values from Fisher's exact test.

Phenotype	Agnostic (n=510)	Panel (n=307)	p
Epilepsy	204 (40.0%)	199 (64.8%)	5.6×10^{-12}
Dysmorphism	139 (27.3%)	25 (8.1%)	5.9×10^{-12}
Developmental delay	152 (29.8%)	38 (12.4%)	4.6×10^{-9}
Epileptic encephalopathy	29 (5.7%)	47 (15.3%)	9.3×10^{-6}
Brain MRI abnormalities	168 (32.9%)	62 (20.2%)	8.1×10^{-5}
Hypotonia	60 (11.8%)	13 (4.2%)	2.0×10^{-4}
Intellectual disability	94 (18.4%)	33 (10.7%)	0.0037
Visual impairment	3 (0.6%)	0 (0.0%)	0.295
Neurodegenerative	2 (0.4%)	0 (0.0%)	0.530
Brain malformations	1 (0.2%)	1 (0.3%)	1.000
Spasticity/spastic paraparesis	0 (0.0%)	0 (0.0%)	1.000
Autism/ASD	0 (0.0%)	0 (0.0%)	1.000
Hearing loss	0 (0.0%)	0 (0.0%)	1.000

Differences for rarer phenotypes (e.g., visual impairment, neurodegenerative features, hearing loss, ASD, spasticity) were not statistically significant and often reflected very small counts.

Baseline and post-reanalysis positive rates by clinical signs and symptoms are shown for the overall cohort in *Supplementary Figure S2* and separately for agnostic and panel subsets *(Figure 16-17)*. In the panel subset, absolute positive rates were lower overall but increased after reanalysis for several clinical features, most notably epilepsy and brain MRI abnormalities (note: extended panel-to-exome analyses were excluded from the exome reanalysis cohort, as they were included in these figures/tables under the reanalyzed panel cohort). McNemar's tests indicated that only a subset of clinical features showed statistically significant within-symptom changes, with most comparisons remaining non-significant. In the agnostic subset, positive rates were highest for hypotonia, developmental delay, dysmorphism and EE, with only modest absolute changes after reanalysis *(Supplementary Tables S3-S6)*.

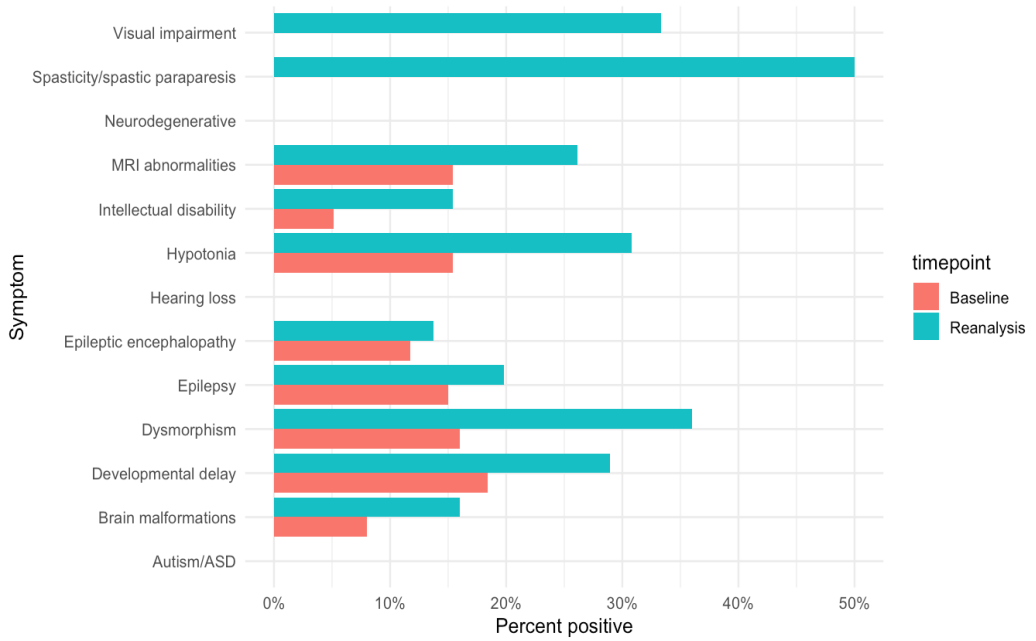


Figure 16. Panel subset: percentage positive by clinical signs and symptoms at baseline and after reanalysis.

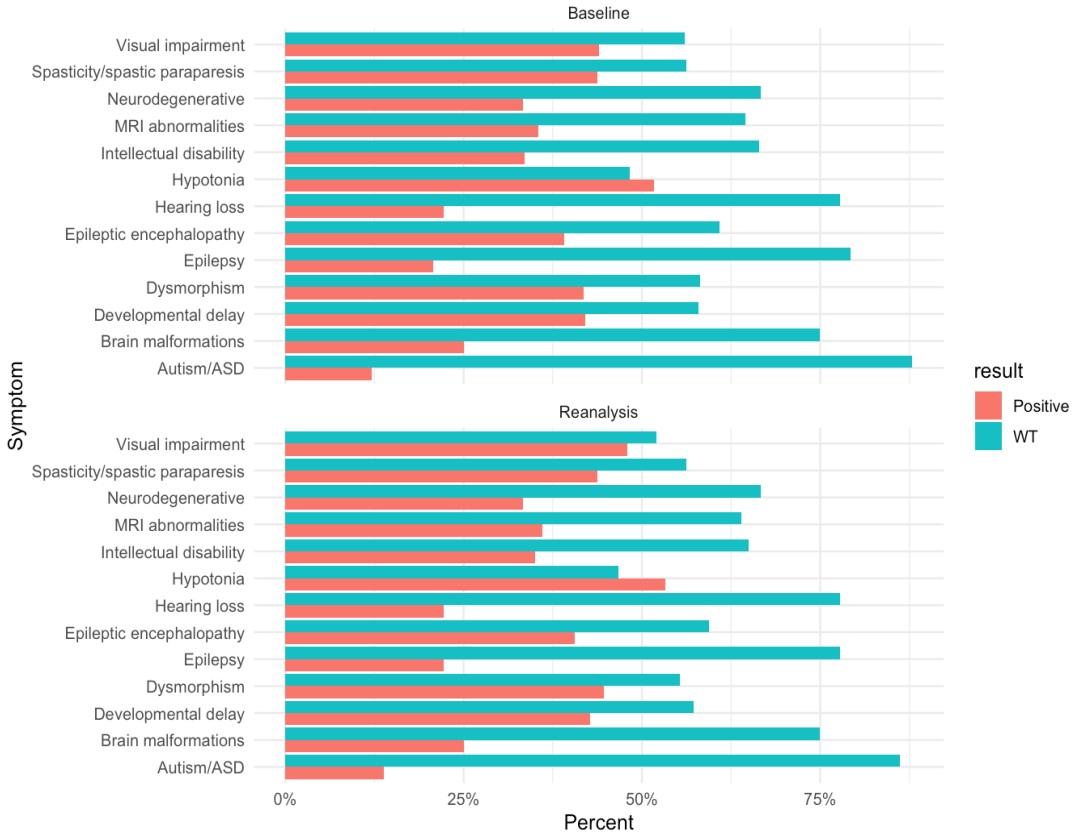


Figure 17. Agnostic subset: percentage positive by clinical signs and symptoms at baseline and after reanalysis.

4.7 Variant classes, VUS, and secondary findings

In this specific analysis, we considered individual variants rather than molecular diagnoses as the unit of measure. Accordingly, the denominator was the total number of variants identified, not the number of molecular diagnoses. In this context, biallelic variants were counted separately at the variant level. Baseline identified variants included 113 P, 80 LP, and 27 c-VUS. Reanalysis added 14 P, 11 LP, and 4 c-VUS, resulting in overall totals of 127 P, 91 LP, and 31 c-VUS.

When stratified by analytic approach, panel-based cases showed 23 P, 16 LP, and 2 c-VUS at baseline, with reanalysis contributing 8 P, 6 LP, and 1 c-VUS (overall totals: 31 P, 22 LP, 3 c-VUS). In the agnostic cohort, baseline counts were 90 P, 64 LP, and 25 c-VUS; reanalysis added 6 P, 5 LP, and 3 c-VUS, for overall totals of 96 P, 69 LP, and 28 c-VUS *(Figures 18-19)*.

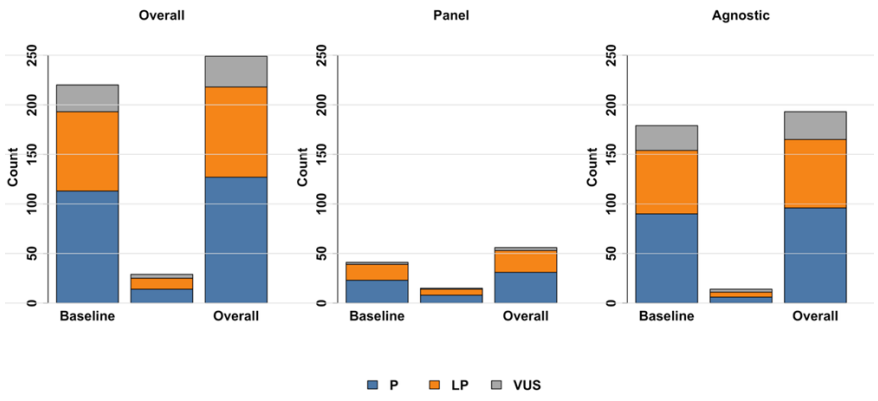


Figure 18. Variant-class counts (P/LP/causative VUS) at baseline, reanalysis, and overall, stratified by approach (overall, panel, agnostic).

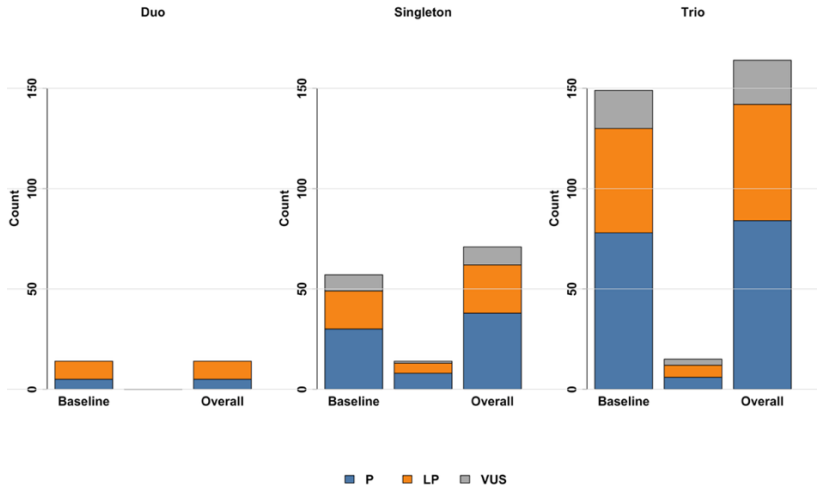


Figure 19. Variant-class counts (P/LP/causative VUS) at baseline, reanalysis, and overall, stratified by family configuration.

VUS reporting (excluding VUS interpreted as causative) occurred at baseline in 202/817 individuals (24.7%; 95% CI 21.8–27.8) and after reanalysis in 15/82 (18.3%; 95% CI 10.6–28.4). One individual had a VUS identified at baseline and an additional VUS upon reanalysis. Overall, 216/817 individuals (26.4%; 95% CI 23.4–29.6) had at least one VUS identified in either phase. By analytic approach, baseline VUS reporting was 22.8% (70/307; 95% CI 18.2–27.9) in panel-based testing and 25.9% (132/510; 95% CI 22.1–29.9) in the agnostic cohort. After reanalysis, rates were 27.3% (12/44; 95% CI 15.0–42.8) and 7.9% (3/38; 95% CI 1.7–21.4), respectively, resulting in overall proportions of 26.7% (82/307; 95% CI 21.8–32.0) and 26.3% (134/510; 95% CI 22.5–30.3), see *Figure 20* and *Supplementary Figure S3*.

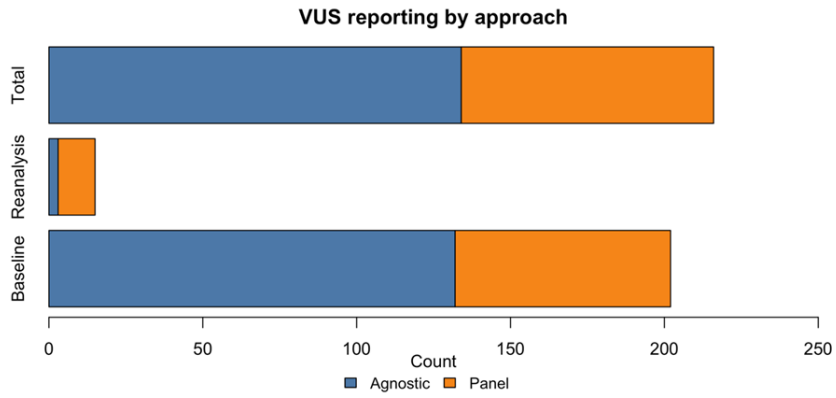


Figure 20. VUS reporting by approach (counts).

By family configuration, baseline VUS reporting was 26.7% in singletons (96/359; 95% CI 22.2–31.6), 21.5% in trios (90/419; 95% CI 17.6–25.7), and 41.0% in duos (16/39; 95% CI 25.6–57.9). Post-reanalysis rates were 27.1% (13/48; 95% CI 15.3–41.8), 3.2% (1/31; 95% CI 0.1–16.7), and 33.3% (1/3; 95% CI 0.8–90.6), yielding overall proportions of 30.4%, 21.7%, and 41.0%, respectively *(Figure 21 and Supplementary Figure S4)*. Longitudinal assessment showed persistent absence of VUS in 76.8% of individuals, with no VUS reported at either baseline or reanalysis. New VUS were identified at reanalysis in 18.3% of cases, whereas 4.9% had previously reported VUS that were confirmed without identification of additional variants.

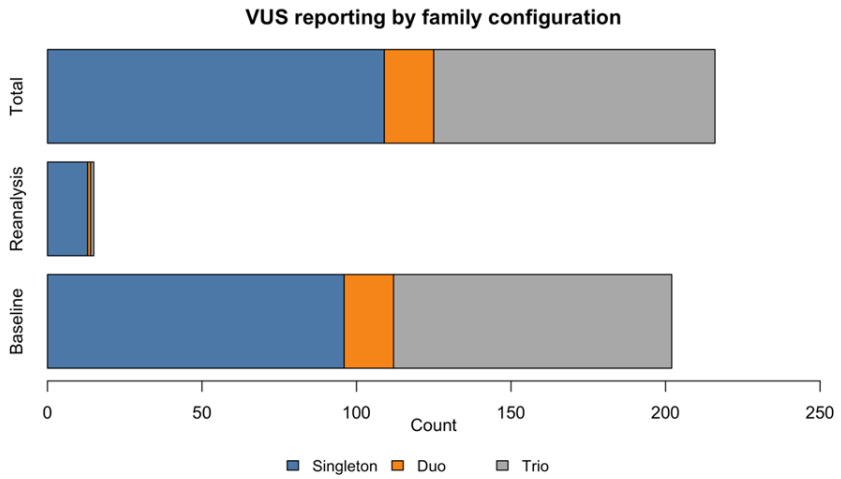


Figure 21. VUS reporting by family configuration (counts).

Secondary findings (SF) were reported in 52/817 individuals (6.4%), with the detailed list provided in *Supplementary Table S3 and Supplementary Figure S5*. Overall, 59 SF were identified. These were classified by clinical system as follows: hemoglobinopathies and erythrocyte enzymopathies (n = 25, 42.4%; *HBB*, *HBA1*, *HBA2*, *G6PD*); cardiovascular disorders (ACMG-recognized; n = 18, 30.5%; *MYH7*, *MYBPC3*, *TTN*, *FLNC*, *SCN5A*, *KCNQ1*, *RYR2*, *LDLR*, *APOB*); cancer predisposition syndromes (ACMG-recognized; n = 6, 10.2%; *BRCA1*, *MSH2*, *MSH6*, *MUTYH*, *SMAD4*, *BLM*); metabolic and lysosomal disorders (n = 3, 5.1%; *GLA*, *LIPE*, *HFE*); ophthalmologic diseases (n = 2, 3.4%; *ABCA4*, *PROM1*); renal or other rare inherited conditions (n = 2, 3.4%; *TBC1D8B*, *MPZL2*); and autoinflammatory (n = 1, 1.7%; *MEFV*), neuromuscular (n = 1, 1.7%; *RYR1*), and mitochondrial/genomic instability–related disorders (n = 1, 1.7%; *RRM2B*) *(Figure 22)*. In two cases (*MYBPC3* and *G6PD*), SF were identified exclusively in the tested mother and disclosed according to prior written informed consent, whereas in one case (*SCN5A*) the finding was withheld following parental opt-out of disclosure.

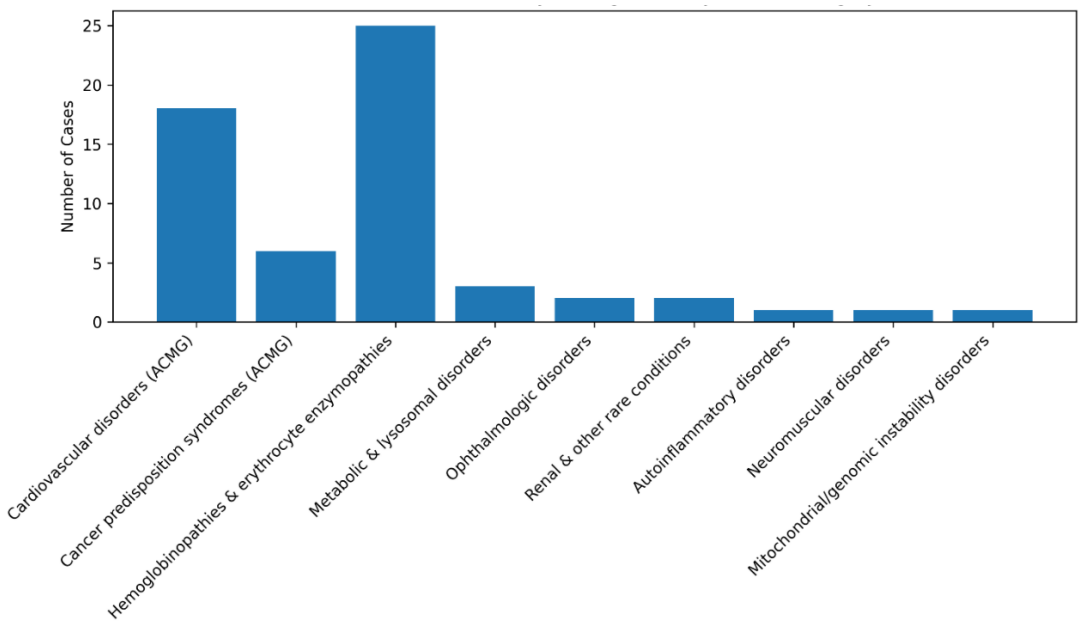


Figure 22. Distribution of secondary finding gene by clinical category.

4.8 Multivariable modeling

We present the exploratory logistic model for baseline diagnostic yield with trio as the reference family configuration (duos excluded).

Sex was not independently associated with outcome (male vs female OR 0.98, 95% CI 0.69–1.38; p = 0.901). Compared with the panel approach, the agnostic approach showed higher odds of a positive result, although the estimate was imprecise and did not reach conventional statistical significance (OR 1.85, 95% CI 0.94–3.55; p = 0.068). Relative to trios, singletons showed lower odds with wide uncertainty (OR 0.73, 95% CI 0.38–1.32; p = 0.309). CNV-called status was not independently associated with outcome (OR 1.36, 95% CI 0.93–1.98; p = 0.106). Prior genetic testing was associated with higher odds of a positive result (OR 1.88, 95% CI 1.13–3.08; p = 0.014) *(table 6)*. A sensitivity model including duos (full cohort, n = 817) yielded essentially the same direction and magnitude of effects, supporting the overall consistency of these exploratory findings.

Table 6. Exploratory multivariable logistic model for baseline diagnostic yield

Predictor (reference)	**OR (95% CI)**	**p**
Male (vs female)	0.98 (0.69–1.38)	0.901
Agnostic (vs panel)	1.85 (0.94–3.55)	0.068
Singleton (vs trio)	0.73 (0.38–1.98)	0.106
CNV called (vs not)	1.36 (0.93–1.98)	0.014
Prior genetic testing (yes vs no)	1.88 (1.13-3.08)	0.014

4.9 Representative Clinical Cases

To further illustrate the clinical relevance of the findings described above, selected representative cases are presented below.

Clinical Case 1. A female patient was referred at 5 years and 11 months for progressive neurological deterioration. Early development was unremarkable. At 2 years, genetic testing for suspected ectodermal dysplasia (hypotrichosis, nail abnormalities) was negative. From 3 years of age, she developed progressive motor impairment with frequent falls and lower limb involvement. Extensive metabolic, infectious, and neurophysiological investigations were initially non-diagnostic. ES performed externally was reported as negative, except for a non-conclusive *KIF1A* variant.

During admissions at our center subsequent evaluations showed rapid neurological worsening with pyramidal signs and neuroradiological progression *(Figure 23; Figure 24a)*. Repeat metabolic testing revealed elevated urinary 3-hydroxyisovaleric acid and increased C5OH acylcarnitine.

Despite the recent negative exome result, urgent reanalysis and resequencing were undertaken. Raw sequencing data (FASTQ files) were retrieved for direct comparison. Using an alternative bioinformatic pipeline, two compound heterozygous variants in the *BTD* gene were identified: a maternally inherited missense variant and a paternally inherited in-frame deletion, consistent with AR inheritance. Phenotype-driven prioritization (Exomiser) and manual review supported their pathogenic relevance. This case exemplifies reverse phenotyping, as the molecular findings prompted re-evaluation of the clinical presentation, leading to recognition of an atypical late-onset form of biotinidase deficiency with progressive neurological involvement. Prompt initiation of biotin supplementation halted neurological deterioration and resulted in marked clinical and neuroradiological improvement *(Figure 24b)*, preventing potential visual impairment. Targeted familial testing identified the same biallelic variants in the patient's sibling, who was neurologically asymptomatic except for sparse scalp hair. Early diagnosis enabled immediate biotin treatment, preventing disease manifestation.

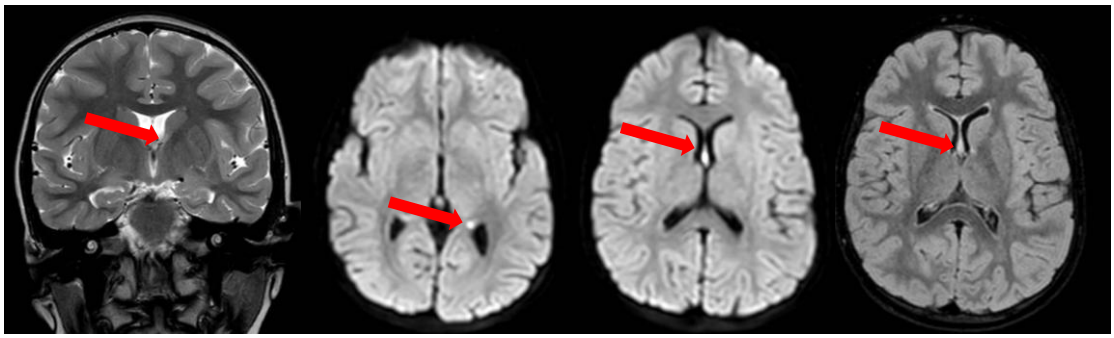
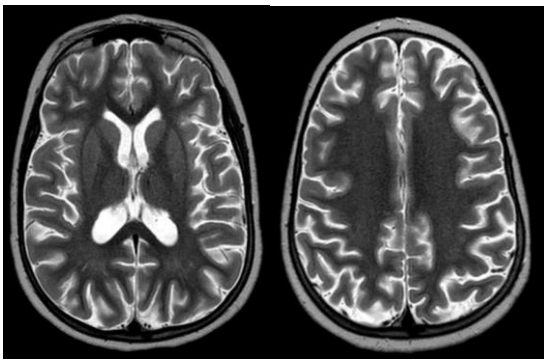


Figure 23. 5.11 years. Left amygdalo-hippocampal swelling, mildly hyperintense on T2-weighted and FLAIR sequences, with diffusion restriction on DWI extending to the ipsilateral fornix and hippocampal tail, consistent with excitotoxic changes (red arrows).

24a.



24b.

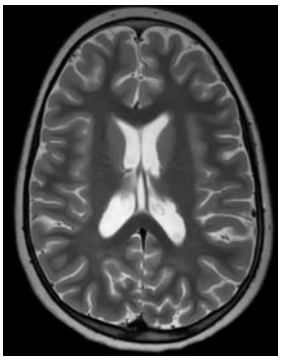


Figure 24. 24a. 6 ys Brain MRI demonstrating cerebral and cerebellar atrophy with bilateral white matter thinning and ex vacuo ventricular enlargement. Persistent diffusion restriction of the fornix columns and subtle T2/FLAIR hyperintensity of the optic chiasm are evident. **24b.** 6.2 ys. Post-treatment follow-up**.** Brain MRI demonstrates resolution of the prior diffusion restriction involving the fornix columns, near-complete resolution of the subtle optic chiasm T2/FLAIR hyperintensity, and overall improvement in cerebral trophism.

Clinical case 2. A male patient was longitudinally followed until 13 years of age for a leukoencephalopathy characterized by intracerebral calcifications, neurodevelopmental abnormalities, mild hemiparesis, skin dyschromia, and subtle dysmorphic features. Early development was initially unremarkable, with onset of motor clumsiness and language regression around 2 years of age. Serial brain imaging showed periventricular calcifications and diffuse white matter abnormalities, while prior infectious workup and targeted genetic testing for basal ganglia calcification–associated genes (*PDGFRB*, *PDGFB*, *SLC20A2*, *XPR1*) were negative.

At a subsequent re-evaluation, trio exome sequencing identified two rare *NOTCH3* variants in compound heterozygosity, inherited from each parent, both absent from population databases and initially classified as VUS. Given the reported association of *NOTCH3* with autosomal recessive vascular leukoencephalopathy, these findings were initially considered potentially contributory.

Updated clinical information prompted neuroradiological assessment of family members, revealing overlapping findings in the father and younger brother, the latter additionally presenting with developmental delay, macrocephaly, and mild dysmorphic features.
Exome reanalysis subsequently identified an additional heterozygous *COL4A2* variant shared among affected family members. Given its known association with cerebral small vessel disease and its segregation with the phenotype, this variant was considered likely pathogenic and explanatory of the clinical presentation.
This case highlights the importance of iterative exome reanalysis integrated with evolving clinical and familial data.

Clinical case 3. An 11-year-old male patient was referred for genetic evaluation due to intellectual disability, cleft palate, and facial dysmorphisms. Developmental delay predominantly affected language, with regression in early childhood. Electroencephalogram showed epileptiform abnormalities, while brain MRI was unremarkable. Previous testing for FRAXA and DiGeorge syndrome was negative; array-CGH had been performed elsewhere without available documentation.
Given the clinical presentation, trio-based WES was performed and identified a heterozygous de novo variant in *SATB2*, consistent with the clinical phenotype. Implementation of a CNV detection algorithm on exome data revealed a deletion involving *PMP22* (chr.17p12), maternally inherited and consistent with hereditary neuropathy with liability to pressure palsies, supporting a dual molecular diagnosis enabled by integrated SNV and CNV analysis.
In line with pre-test consent, analysis also identified a paternally inherited variant in *RYR2* associated with catecholaminergic polymorphic ventricular tachycardia type 1, a potentially life-threatening but preventable arrhythmogenic condition. This secondary finding prompted cardiological evaluation of the proband and at-risk relatives, as well as cascade testing of first-degree relatives, with relevant preventive implications.

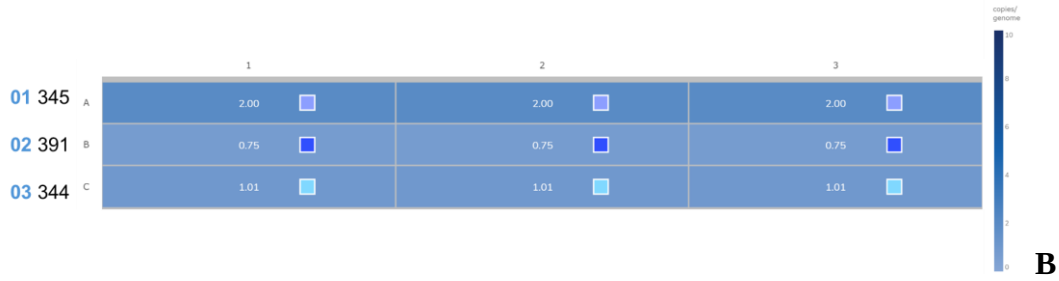
Clinical case 4. A 10-year-old male patient was referred for evaluation of a neurodevelopmental disorder associated with macrocephaly and corpus callosum abnormalities. Developmental delay was noted from early childhood, with electroencephalographic evidence of epileptiform activity and nonspecific corpus callosum changes on brain imaging. At our clinical assessment, he presented with severe

communication impairment, hyperkinetic behavior, mild dysmorphic features, and pigmentary skin anomalies.

Previous array-CGH identified a maternally inherited duplication on chromosome 20, considered non-pathogenic.

Trio-based WES, performed in our laboratory with a pipeline optimized for CNV detection, identified compound heterozygosity in *EMC10* consisting of a maternally inherited pathogenic single-nucleotide variant and a paternally inherited exon-level deletion (ex 1) predicted to be pathogenic and confirmed by digital PCR *(Figure 25)*. Pathogenic variants in *EMC10* are associated with AR neurodevelopmental disorder with dysmorphic features and epilepsy (OMIM #619264). Identification of the exon-level CNV was essential to establish compound heterozygosity and achieve molecular diagnosis.

A



B

Figure 25. (A) CNV analysis by digital PCR (QIAGEN QIAcuity) highlights a heterozygous deletion of *EMC10* in the proband (344), inherited from the father (391) and absent in the healthy control (mother, 345). This result is also evident in the heatmap (B).

Clinical case 5. A 10-year-old boy, presented with hypotonia, neurodevelopmental delay, stridor, and feeding difficulties with failure to thrive, later developing persistent gait imbalance, fine motor impairment, ataxia with frequent falls, and hyposthenia. Additional

features included facial dysmorphisms, palmar keratoderma, distal arthrogryposis, and skeletal abnormalities (clubfoot, scoliosis, pectus excavatum). Laboratory (including CK and muscle biopsy), neurophysiological, and neuroimaging studies were unremarkable. Previous genetic testing, including methylation analysis for Prader–Willi and Angelman syndromes, was negative. Array-CGH identified a paternally inherited microduplication on chromosome 15, classified as a VUS.

Given the clinical presentation, trio-base WES identified two novel *PIEZO2* variants: a maternally inherited nonsense variant and a paternally inherited intronic variant, initially classified as a VUS. Variants were confirmed by Sanger sequencing. In light of the strong phenotypic concordance with PIEZO2-related disease, functional studies were performed. Minigene assays demonstrated aberrant splicing caused by the intronic variant, leading to a premature termination codon and predicted nonsense-mediated mRNA decay (NMD). Studies in patient-derived fibroblasts treated with NMD inhibitors confirmed degradation of transcripts from both alleles *(Figure 26)*, supporting reclassification of the intronic variant as pathogenic. Biallelic pathogenic variants in *PIEZO2* cause AR distal arthrogryposis with impaired proprioception and touch (OMIM #617146). The integration of genomic and functional evidence established a definitive molecular diagnosis consistent with the patient's phenotype.

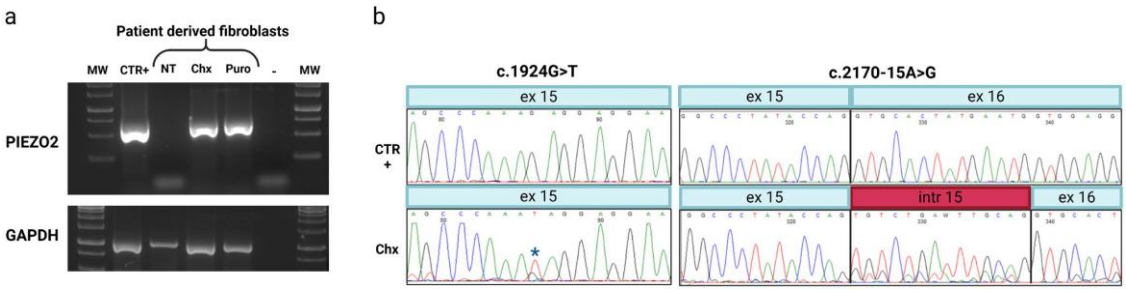


Figure 26. The identified variants target the *PIEZO2* mRNA to NMD in patient's primary fibroblasts. (a) PCR products obtained by the amplificationof cDNA derived from patient's fibroblasts treated with 100 µg/mL of Cycloheximide (Chx) or 200 µg/mL of Puromycin (Puro). CTR+ cDNAderived from RNA extracted from human dermal fibroblasts used as positive control. GAPDH was used as housekeeping gene to check the efficiencyof the cDNA synthesis. (b) Details of Sanger sequencing of the cDNA, focusing on exon 15 and of the region spanning from exons 15 and 16. Blue asterisk indicates the variant inherited from the mother and the red area represent the insertion of 14 nt of intron 15 into the transcript due to thesplicing variant inherited from the father. MW, Molecular Weight 1 Kb. Image from Bellardita et al, 2025.

The clinical and the functional characterization of the intronic *PIEZO2* variant has been previously published by our laboratory group in: *"Functional Characterization of a Novel Intronic Variant in PIEZO2 in a Recessive Form of Distal Arthrogryposis With Impaired Proprioception and Touch (DAIPT)"*, published in Mol Genet Genomic Med. 2025;13(8):e70126. https://doi.org/10.1002/mgg3.70126. [35]

5 Discussion

In this single-center, real-world cohort of 817 pediatric patients evaluated in a tertiary neurogenetics setting, the overall baseline diagnostic yield of ES, irrespective of analytical strategy, was 23.5%. Although slightly lower than yields (27–43%) reported in selected cohorts and meta-analyses [6-8], this likely reflects the unselected and operationally heterogeneous nature of routine clinical practice, including variable pre-test probability and broad referral patterns. Our results therefore provide a pragmatic benchmark of ES performance under real-world conditions rather than an optimized research estimate.

A central finding of this exploratory analysis is the significantly higher diagnostic yield achieved with agnostic exome-wide interpretation compared with panel-restricted analysis (29.6% vs 13.4%, $p < 0.001$). Although robust in magnitude, comparative yield differences must be interpreted within the constraints of phenotype-driven allocation and real-world workflow implementation inherent to this exploratory setting, particularly the relative enrichment of epilepsy/EE cases within the panel-based cohort. In highly heterogeneous neurological disorders and NDDs, restricting interpretation to predefined gene sets may reduce analytical burden but risks missing relevant diagnoses, particularly in multisystemic or atypical presentations and within a rapidly expanding gene–disease landscape [12-14]. Conversely, exome-wide interrogation systematically captures pathogenic variants across a broad and evolving genomic spectrum, consistent with guideline-supported implementation of ES/GS as first- or second-tier testing [3-4].

Panel performance varied substantially across phenotypic subgroups, underscoring that diagnostic yield is closely linked to phenotype specificity and panel design [12,18]. Nevertheless, targeted approaches may remain appropriate in narrowly defined clinical scenarios, particularly when the phenotype is highly specific, the relevant gene set is stable and well characterized, turnaround time is critical, or resource constraints limit immediate exome-wide interpretation [12,18]. In this context, accurate and detailed phenotyping remains essential. Neonatal epilepsies, for example, may present with distinctive semeiology and electroclinical patterns that can meaningfully guide test selection and justify a focused, condition-specific panel approach. Similarly, highly suggestive neuroimaging findings may orient toward a restricted set of candidate genes supporting a targeted testing strategy. In line with this concept, our findings—particularly

for the mTOR panel—reinforce the value of phenotype-driven approaches in appropriately selected cases.

However, given the generally lower diagnostic yield of panels compared with exome-wide analysis, panel-first strategies are best embedded within tiered diagnostic pathways, in which negative panel results prompt escalation to unrestricted exome interpretation [18,20].

Systematic reanalysis emerged as a clinically meaningful extension of genomic testing, in line with the ACMG framework [33] and empirical evidence of incremental yield [34]. Reanalysis achieved a 27.2% incremental diagnostic yield among initially non-diagnostic cases (22/81), corresponding to a 2.7% absolute increase without additional primary sequencing. Beyond panel-to-exome escalation, periodic bioinformatic reinterpretation and clinically driven reassessment independently contributed to diagnostic gain. In the agnostic-only strict-exome subset, reanalysis yielded 24.3% (9/37; absolute increase 1.8%), predominantly driven by periodic bioinformatic reanalysis (55.6%) and updated clinical reassessment (33.3%).

Beyond its role as a single diagnostic test, ES should be viewed as a longitudinal, dynamic diagnostic platform rather than a static sequencing event. The incremental yield achieved through reanalysis—via panel-to-exome escalation and periodic bioinformatic reinterpretation—demonstrates that diagnostic value accumulates over time without additional primary sequencing, shifting the paradigm from “test result” to “evolving genomic resource”.

This framework integrates expanding gene–disease knowledge, refinement of variant classification systems, and phenotypic evolution with improved clinical annotation. Accordingly, ES represents a renewable diagnostic asset whose interpretative potential matures alongside scientific progress and longitudinal follow-up, transforming initially negative findings into temporally open cases.

Beyond its biological relevance, structured reanalysis has organizational impact. Implementing predefined intervals (e.g., 12–24 months), automated reannotation pipelines, and clinician-triggered reassessment reduces interpretative variability, promotes equitable access to updated knowledge, optimizes laboratory resources, and avoids unnecessary repeat sequencing. From a health-system perspective, periodic reinterpretation is a high-value, cost-efficient strategy that maximizes diagnostic return on prior sequencing investment, particularly in tertiary pediatric neurogenetics settings characterized by prolonged diagnostic trajectories [36].

Family configuration was another key determinant of performance. Although duos showed an apparently higher yield (33.3%, considering the overall cohort at baseline), this likely reflects small sample size and clinical selection bias, including overrepresentation of proband–mother pairs. More robustly, trio-based testing outperformed singleton analysis (29.1% vs 15.0%), consistent with the major contribution of de novo variants to pediatric NDDs [4] and prior evidence of improved interpretative resolution in trios [27-28]. The predominance of AD diagnoses (73.1%), particularly AD de novo variants (40.4%), biologically supports this observation. Trio analysis strengthens de novo classification under ACMG/AMP criteria [10], improves segregation assessment, and facilitates detection of complex mechanisms such as compound heterozygosity and mosaicism. Despite non-random allocation of analytic strategies, the data strongly support trio-based ES whenever feasible.

Beyond yield, trio ES may be cost-effective when downstream costs are considered: fewer unresolved VUS, reduced need for reflex segregation testing, fewer iterative "diagnostic odyssey" steps, and shorter time-to-diagnosis can offset higher upfront costs. Health-economic evaluations in rare disease pathways show that earlier ES/GS access can be cost-saving versus prolonged conventional testing, and that per-patient costs differ substantially between singleton and trio sequencing, supporting phenotype-driven prioritization of trio testing as a rational and resource-conscious implementation strategy [37-38].

Conversely, singleton ES may remain appropriate in clearly autosomal recessive phenotypes, particularly in consanguineous families, in resource-limited settings, or in late-onset, well-characterized monogenic disorders with a low expected contribution of de novo variants. Accordingly, while family-based sequencing should be prioritized when biologically and logistically feasible, implementation strategies may reasonably incorporate phenotype-driven triage to optimize resource allocation without compromising diagnostic effectiveness [39].

The spectrum of identified genes recapitulated the established architecture of NDDs, characterized by marked locus heterogeneity and convergence on synaptic, transcriptional, chromatin-remodeling, and metabolic pathways [3-4]. Recurrent diagnoses in *PRRT2*, *SCN2A*, *GRIN2A*, *CACNA1A*, and *DEPDC5* reflected the combined output of both analytical frameworks rather than a single workflow. The agnostic approach captured substantially greater gene diversity, reinforcing its advantage in complex phenotypes. Reanalysis also shifted part of the diagnostic burden toward AR

conditions, particularly compound heterozygous configurations, further supporting periodic reinterpretation as recommended by ACMG [33].

Variant classification was robust: approximately 88% of diagnoses were supported by P or LP variants, with only a minority based on VUS interpreted as causative in the presence of strong genotype–phenotype concordance and convergent in silico evidence. This conservative framework enhances reliability and aligns with ACMG standards [10]. Notably, trio-based analysis was associated with a lower baseline VUS rate compared with singletons (21.5% vs 26.7%), supporting the interpretative advantage conferred by segregation data in reducing uncertainty. Moreover, the marked reduction of VUS after reanalysis in the trio subgroup (3.2%) highlights the dynamic nature of variant classification, as reinterpretation over time frequently enables VUS reclassification or resolution.

Phenotype–genotype correlations further contextualized yield. Brain MRI abnormalities, developmental delay, dysmorphism, and hypotonia were significantly enriched among diagnostic cases, consistent with higher monogenic detection rates in syndromic or structurally defined phenotypes [7-8]. Conversely, epilepsy and ASD were more frequent among non-diagnostic cases, reflecting broader etiological heterogeneity and possible contributions from regulatory, mosaic, or polygenic mechanisms not fully captured by ES [12]. Within the agnostic subset, the highest yields were observed in hypotonia, developmental delay, dysmorphism, and EE, phenotypes enriched for highly penetrant de novo variants [4].

An additional strength of ES lies in integrated multi-variant-class analysis. Read-depth–based CNV calling provided a 5.6% incremental diagnostic contribution within the CNV-called subset, consistent with previous reports [30-32]. CNVs were decisive in three contexts: CNV-only diagnoses; compound heterozygous CNV/SNV configurations; and dual molecular diagnoses. These findings highlight the complementary nature of structural and sequence-level variants within a unified exome dataset. Nevertheless, ES is not the gold-standard for CNV detection; read-depth algorithms remain coverage-dependent with estimated sensitivity around 90%. Accordingly, all CNVs were confirmed by orthogonal methods (array-CGH or MLPA), ensuring diagnostic rigor.

Despite its strong performance, ES has intrinsic technical and interpretative limitations, including incomplete coverage and challenges in variant classification. Complementary approaches—such as functional assays—may therefore be required to confirm

pathogenicity in selected cases, as emphasized in recent genomic literature [40]. This integrative strategy was exemplified in representative case 5.

Secondary findings were reported in 6.4% of individuals, encompassing cardiovascular, oncologic, hematologic, metabolic, and other actionable conditions, in accordance with ACMG recommendations [11]. Their detection underscores both the preventive potential and ethical complexity of exome-wide testing, particularly in trio configurations involving unaffected parents. Structured consent and multidisciplinary governance proved essential for responsible disclosure.

Finally, this findings highlights the bidirectional relationship between laboratory and clinic. Genomic findings frequently prompted reverse phenotyping, refining clinical interpretation through targeted re-examination. As illustrated by Clinical Case 1, molecular results do not merely confirm clinical hypotheses—they can reshape them. This iterative clinic–laboratory dialogue represents a defining feature of contemporary genomic medicine and amplifies the diagnostic impact of ES beyond static data generation.

6 Limitations and Strengths

This real-world, single-center exploratory analysis has inherent methodological limitations. First, the analysis design may introduce selection bias; allocation to analytical strategies and family configurations was not randomized but reflected routine clinical implementation, limiting comparability across groups and precluding causal inference. CNV calling was introduced later during the study period and applied to a numerically unbalanced subset, reducing internal uniformity of the cohort. In addition, subgroup sizes were generally small, and temporal analyses were not performed, as results were designed to provide cumulative four-year estimates rather than date-stratified outcomes.

The multivariable logistic regression model assessing baseline diagnostic yield should be interpreted as exploratory and hypothesis-generating. Effect estimates were characterized by limited precision, with wide confidence intervals, and duos were excluded from the primary model, reducing statistical power for subgroup analyses. Although sensitivity analyses including the full cohort (n = 817) demonstrated directional consistency and supports the robustness of findings, these should be interpreted cautiously in light of the wide confidence intervals, exploratory design and limited statistical power.

Nevertheless, several strengths enhance the validity of the findings: the large pediatric cohort, the high proportion of trio-based analyses enabling robust inheritance modeling, and the application of both panel-restricted and agnostic workflows within a consistent institutional and interpretative framework. The integrated assessment of SNVs, indels, and CNVs—combined with systematic orthogonal validation—ensured analytical rigor while maximizing diagnostic completeness. Structured ACMG-based variant classification, detailed phenotype–genotype correlation, and formal evaluation of reanalysis pathways further support biological coherence and interpretative robustness.

7 Future Perspectives

The evolution of genomic diagnostics in pediatric neurodevelopmental disorders is progressively moving toward an integrated, multilayered framework that extends beyond conventional ES. RNA sequencing has demonstrated increasing clinical utility in resolving splice-altering and regulatory variants and in improving diagnostic yield in previously unsolved Mendelian conditions [41-42]. Long-read genome sequencing further enhances detection of complex structural variants, repeat expansions, and cryptic genomic rearrangements that remain challenging for short-read technologies, with documented increases in diagnostic yield in neurodevelopmental cohorts [43]. In parallel, genome-wide DNA methylation profiling and disorder-specific episignatures are emerging as robust complementary tools for refining variant interpretation and resolving genetically ambiguous cases [44-45]. Finally, systematic escalation to WGS in exome-negative individuals—supported by recent large-scale analyses and best-practice recommendations—enables comprehensive interrogation of non-coding regions, deep intronic variants, and mosaic events not captured by coding-focused approaches [7]. Together, these advances delineate a transition toward truly integrated genomic medicine, in which ES represents not an endpoint, but the core of a dynamic and continuously expandable diagnostic platform capable of progressively reducing the unsolved fraction in heterogeneous pediatric disorders.

8 Conclusions

Taken together, our findings indicate that pediatric neurological and NDDs in a tertiary-care setting are predominantly driven by highly heterogeneous, largely dominant (often de novo) monogenic mechanisms that are most effectively captured through trio-based, exome-wide interrogation. Compared with panel-restricted strategies, agnostic ES emerges as an expandable diagnostic platform, capable of integrating SNVs, indels, CNVs, iterative reinterpretation, and clinically actionable secondary findings within a single dataset. In this framework, its added value extends beyond diagnostic yield, residing in its structural capacity to enable dynamic, multidimensional, and clinically integrated genomic diagnostics over time, shifting the paradigm from predefined testing strategies toward platform-based personalized genomics applied to the individual patient.

9 References

1. **Marwaha S, Knowles JW, Ashley EA.** A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. *Genome Med*. 2022;14(1):23. https://doi.org/10.1186/s13073-022-01026-w.
2. **Boyle CA, Boulet S, Schieve LA, et al.** Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics* 2011;127 (6), 1034–1042. https://doi.org/10.1542/peds.2010-2989.
3. **Manickam K, McClain MR, Demmer LA, et al; ACMG Board of Directors.** Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the ACMG. *Genet Med*. 2021;23(11):2029–2037. https://doi.org/10.1038/s41436-021-01242-6.
4. **Srivastava S, Love-Nichols JA, Dies KA, 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. https://doi.org/10.1038/s41436-019-0554-6.
5. **Srivastava S, Cohen JS, Vernon H, et al.** Clinical whole exome sequencing in child neurology practice. *Ann Neurol.* 2014;76(4):473–483. https://doi.org/10.1002/ana.24251.
6. **Neuens S, Soblet J, Penninckx A, et al.** Diagnostic yield of clinical exome sequencing in 868 children with neurodevelopmental disorders. *Eur J Med Genet*. 2025 Aug:76:105030. https://doi.org/10.1016/j.ejmg.2025.105030.
7. **Pandey R, Fluetsch Brennan N, Trachana K, et al.** A meta-analysis of diagnostic yield and clinical utility of genome and exome sequencing in pediatric rare and undiagnosed genetic diseases. *Genet Med.* 2025;27(6):101398. https://doi.org/10.1016/j.gim.2025.101398.
8. **Slavotinek A, Rego S, Sahin-Hodoglugil N, et al.** Diagnostic yield of pediatric and prenatal exome sequencing. NPJ Genom Med. 2023 May 26;8(1):10. https://doi.org/10.1038/s41525-023-00353-0.
9. **Lan X, Tang X, Weng W, et al.** Diagnostic Utility of Trio–Exome Sequencing for Children With Neurodevelopmental Disorders. *JAMA Netw Open*. 2025 Mar 3;8(3):e251807. https://doi.org/10.1001/jamanetworkopen.2025.1807.

10. **Richards S, Aziz N, Bale S, et al.** Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the ACMG and AMP. *Genet Med.* 2015;17(5):405–424. https://doi.org/10.1038/gim.2015.30.
11. **Miller DT, Lee K, Gordon AS, et al.** Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2021 update: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021 August ; 23(8): 1391–1398. https://doi.org/10.1038/s41436-021-01171-4.
12. **Souche E, Beltran S, Brosens E, et al. (ESHG/EuroGentest).** Recommendations for the use of genome-wide sequencing in clinical diagnostics. *Eur J Hum Genet.* 2022;30:1017–1025. https://doi.org/10.1038/s41431-022-01113-x.
13. **Sheidley BR, Malinowski J, Bergner AL, et al.** Genetic testing for the epilepsies: A systematic review. *Epilepsia.* 2022;63(2):375–387. https://doi.org/10.1111/epi.17141.
14. **Sedláčková L, Sterbova K, Vlckova M, et al.** Yield of exome sequencing in patients with developmental and epileptic encephalopathies and inconclusive targeted gene panel. *Eur J Paediatr Neurol.* 2024;46:109–116. https://doi.org/10.1016/j.ejpn.2023.10.006**.**
15. **Miller DT, Lee K, Abul-Husn NS, et al.** ACMG SF v3.2 list for reporting of secondary findings in clinical exome and genome sequencing. *Genet Med.* 2023. https://doi.org/0.1038/s41436-023-01471-6.
16. **Kohler S, Gargano M, Matentzoglu N, et al.** The Human Phenotype Ontology in 2021. *Nucleic Acids Research*, 2021, Vol. 49, Database issue D1207–D1217. https://doi.org/10.1093/nar/gkaa1043.
17. **Sánchez Suárez A, Martínez Menéndez B, Escolar Escamill E, et al.** Whole exome sequencing and panel-based analysis in 176 children with neurodevelopmental disorders. *Genes (Basel).* 2024;15(10):1310. https://doi.org/10.3390/genes15101310.
18. **Molina-Ramírez LP, Kyle C, Ellingford JM, et al.** Personalised virtual gene panels reduce interpretation workload and maintain diagnostic rates of proband-only clinical exome sequencing for rare disorders. *J Med Genet.* 2021;59(4):393–400. https://doi.org/10.1136/jmedgenet-2020-107303**.**

19. **Lee K, Abul-Husn NS, Amendola LM, et al (ACMG).** ACMG SF v3.3 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2025 Aug;27(8):101454. https://doi.org/10.1016/j.gim.2025.101454.
20. **Van der Schoot V, Haer-Wigman L, Feenstra I, et al.** Lessons learned from unsolicited findings in clinical exome sequencing of 16,482 individuals. *Eur J Hum Genet*. 2022 Feb;30(2):170-177. https://doi.org/10.1038/s41431-021-00964-0.
21. **Kore P, Wilson MW, Tiao G, et al**. Improved allele frequencies in gnomAD through local ancestry inference. *Nat Commun*. 2025 Oct 6;16(1):8734. https://doi.org/10.1038/s41467-025-63340-2.
22. **Yuan X, Wang J, Dai B.** Evaluation of phenotype-driven gene prioritization methods for Mendelian diseases. *Briefings in Bioinformatics.* 2022;23(2):bbac019. https://doi.org/10.1093/bib/bbac019.
23. **ClinGen Variant Classification Guidance (SVI WG).** *Clinical Genome Resource.*Last updated: July 2025.
24. **Walker LC, de la Hoya M, Wiggins GAR, et al.** Using the ACMG/AMP framework to capture evidence related to predicted and observed impact on splicing: Recommendations from the ClinGen SVI Splicing Subgroup. *Am J Hum Genet*. 2023 Jul 6;110(7):1046-1067. https://doi.org/10.1016/j.ajhg.2023.06.002
25. **Bridges Y, de Souza V, Cortes KG.** Towards a standard benchmark for phenotype-driven variant and gene prioritisation algorithms: PhEval - Phenotypic inference Evaluation framework. *bioRxiv*[Preprint]. 2025 Feb 20:2024.06.13.598672. Originally published 2024 Jun 16. [Version 2] https://doi.org/10.1101/2024.06.13.598672.
26. **Seaby EG, Thomas NS, Hunt D, et al.** A Panel-Agnostic Strategy 'HiPPo' Improves Diagnostic Efficiency in the UK Genomic Medicine Service. *Healthcare (Basel).* 2023;11(24):3179. https://doi.org/10.3390/healthcare11243179.
27. **Sanchis-Juan A, Megy K, Stephens J, et al.** Genome sequencing and comprehensive rare-variant analysis of 465 families with neurodevelopmental

disorders. *Am. J. Hum. Genet.* 2023. 110 (8), 1343–1355. https://doi.org/10.1016/j.ajhg.2023.07.007.

28. **Smedley D, Smith KR, Martin A, et al.** 100,000 genomes Pilot on rare-disease diagnosis in health care – preliminary report. *N. Engl. J. Med.* 2021. 385 (20), 1868–1880. https://doi.org/10.1056/NEJMoa2035790.
29. **López FV, Ashton JJ, Cheng G.** A systematic analysis of contemporary whole exome sequencing capture kits to optimise high-coverage capture of CCDS regions**.** *NAR Genom Bioinform.* 2025 Sep 1;7(3):lqaf115. https://doi.org/10.1093/nargab/lqaf115.
30. **Babadi M et al, Fu JM, Lee SK, et al.** GATK-gCNV enables discovery of rare copy number variants from exome sequencing data. *Nat Genet*. 2023 September; 55(9): 1589–1597. https://doi.org/10.1038/s41588-023-01449-0.
31. **Guo Q, Zhang Y, Gui X, et al.** Evaluation of CNV detection tools for prenatal diagnosis using amniotic fluid clinical whole-exome sequencing data. *Genomics*. 2026 Jan 8;118(2):111196. https://doi.org/10.1016/j.ygeno.2026.111196.
32. **Louw N, Carstens N, Lombard Z, et al.** Incorporating CNV analysis improves the yield of exome sequencing for rare monogenic disorders—an important consideration for resource-constrained settings. Front Genet. 2023 Dec 14;14:1277784. https://doi.org/10.3389/fgene.2023.1277784.
33. **Deignan JL, Chung WK, Kearney HM, et al.** Points to consider in the reevaluation and reanalysis of genomic test results: a statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2019 Jun;21(6):1267-1270. https://doi.org/10.1038/s41436-019-0478-1.
34. **Malgrem H, Kvarnung M, Gustafsson, et al.** Diagnostic yield of 1000 trio analyses with exome and genome sequencing in a clinical setting. *Front Genet.* 2025 Jun 20:16:1580879. https://doi.org/10.3389/fgene.2025.1580879.
35. **Bellardita M, Ferruccio R, Menta L, et al.** Functional Characterization of a Novel Intronic Variant in *PIEZO2* in a Recessive Form of Distal Arthrogryposis With Impaired Proprioception and Touch (DAIPT)", *Mol Genet Genomic Med*. 2025;13(8):e70126. https://doi.org/10.1002/mgg3.70126.
36. **Best S, Fehlberg Z, Richards C, et al.** Reanalysis of genomic data in rare disease: current practice and attitudes among Australian clinical and laboratory

genetics services. *Eur J Hum Genet*. 2024 Nov;32(11):1428-1435. https://doi.org/10.1038/s41431-024-01633-8.

37. **Regier DA, Loewen R, Chan B, et al.** Real-world diagnostic outcomes and cost-effectiveness of genome-wide sequencing for developmental and seizure disorders: Evidence from Canada. *Genet Med*. 2024 Apr;26(4):101069. https://doi.org/10.1016/j.gim.2024.101069.
38. **Lavelle TA, Feng X, Keisler M, et al.** Cost-effectiveness of exome and genome sequencing for children with rare and undiagnosed conditions. *Genet Med*. 2022 Jun;24(6):1349-1361. https://doi.org/10.1016/j.gim.2022.03.005.
39. **Tan TY, Lunke S, Chong B, et al.** A head-to-head evaluation of the diagnostic efficacy and costs of trio versus singleton exome sequencing analysis. *Eur J Hum Genet*. 2019 Dec;27(12):1791-1799. https://doi.org/10.1038/s41431-019-0471-9.
40. **Tabet D, Parikh V, Mali P, et al.** Scalable Functional Assays for the Interpretation of Human Genetic Variation. *Annu Rev Genet*. 2022 Nov 30:56:441-465. https://doi.org/10.1146/annurev-genet-072920-032107.
41. **Zhao S, Macakova K, Sinson JC, et al.** Clinical validation of RNA sequencing for Mendelian disorder diagnostics**.** *The American Journal of Human Genetics***.** 2025;112(4):779–792. https://doi.org/10.1016/j.ajhg.2025.02.006.
42. **Murdock DR, Dai H, Burrage LC, et al.** Transcriptome-directed analysis for Mendelian disease diagnosis overcomes limitations of DNA testing. *J Clin Invest*. 2021;131(1):e141500. https://doi.org/10.1172/JCI141500.
43. **Hiatt SM, Lawlor JMJ, Handley LH, et al.** Long-read genome sequencing and variant reanalysis increase diagnostic yield in neurodevelopmental disorders**.** *Genome Research***.** 2024;34(11):1747–1762. https://doi.org/10.1101/gr.279227.124.
44. **LaFlamme CW, Rastin C, Sengupta S, et al.** Diagnostic utility of DNA methylation analysis in genetically unsolved pediatric epilepsies and CHD2 episignature refinement**.** *Nature Communications***.** 2024;15(1):6524. https://doi.org/10.1038/s41467-024-50159-6.
45. **Husson T, Lecoquierre F, Nicolas G, et al.** Episignatures in practice: independent evaluation of published episignatures for the molecular diagnostics of ten neurodevelopmental disorders. *European Journal of Human Genetics***.** 2024;32:190–199. https://doi.org/10.1038/s41431-023-01474-x.

10 Supplementary Material

10.1 Supplementary Methods

Sequencing and Bioinformatic Processing. Genomic DNA was extracted from peripheral blood using standard laboratory protocols; in selected probands, alternative tissue sources were used when clinically indicated. Massive parallel sequencing was performed according to validated clinical laboratory procedures.

Read alignment was conducted using Burrows–Wheeler Aligner (BWA), and variant calling was performed using Genome Analysis Toolkit (GATK)-based pipelines following Best Practices guidelines, (https://gatk.broadinstitute.org/). In parallel, data were processed using a commercial analysis platform (Sophia DDM v4).

The ILL1XG1G5_CNV_exome pipeline was used for identification of single nucleotide variants (SNVs), small insertions/deletions (indels), and exome-based copy number variants (CNVs).

All variants were referenced to the GRCh38/hg38 genome assembly and annotated according to RefSeq transcript records (NM_ accession prefix).

Variant Filtering and Annotation. Variants were filtered based on allele frequency (≤0.01% in population databases, including gnomAD v2.1.1), segregation consistency within the family, evolutionary conservation (GERP score), and predicted functional impact using in silico tools (SIFT, PolyPhen-2, MutationTaster, AlphaMissense).

Clinical annotation included interrogation of ClinVar and additional curated databases. Variant classification was performed according to ACMG/AMP criteria [Richards R et al, 2015] using dedicated interpretation software (Franklin by Genoox).

Targeted Sanger sequencing was performed for validation of identified variants and for segregation analysis when required.

CNV Detection and Validation. CNVs were identified from exome data using a read-depth–based algorithm integrated within the exome analysis pipeline.

CNVs were validated using independent confirmatory techniques, including array comparative genomic hybridization (array-CGH) or, in selected cases, multiplex ligation-dependent probe amplification (MLPA).

Data Coding and Variable Definitions

The analytical dataset was structured with one row per tested individual.

The variable *Exome* was used to describe WES-specific outcomes and included the following categories:

- WT (wild type),
- WT (SNV-),
- Positive,
- Positive (SNV-),
- Positive SNV+.

Individuals solved exclusively through panel-based analysis did not have WES-specific outcome coding and were excluded from strict exome-only yield calculations.
CNV-calling status was derived from the variable *Detection CNV*, coded as:

- NE (not performed),
- WT (CNV-),
- Positive (CNV+).

CNV-related analyses were restricted to cases with CNV calling performed (WT or Positive).

Reanalysis categories were not mutually exclusive and were recorded independently when multiple reanalysis types applied to the same case.

Statistical Software

All analyses were conducted using R (version 4.5.2; R Foundation for Statistical Computing, 2025).

10.2 Supplementary Tables

Supplementary Table S1. Genes included in the EE, CM, and MTOR pathway Panels

EE Panel	*ALDH7A1, ALG13, ARHGEF9, ATP6V0C, ATP6V1A, BRAT1, CACNA1A, CACNA1B, CAMK2A, CDKL5, CHD2, CHRNA4, CSNK2B, CUX2, DDC, DEPDC5, DNM1, EEF1A2, FGF12, FOXG1, GABRA1, GABRB2, GABRB3, GABRG2, GAD1, GNAO1, GRIA2, GRIN1, GRIN2A, GRIN2B, GRIN2D, HCN1, HCN2, HNRNPU, IQSEC2, ITPA, KCNA2, KCNB1, KCNQ2, KCNQ3, KCNT1, KCNT2, MECP2, MEF2C, NEXMIF, NPRL2, NTRK2, PACS2, PCDH19, PDE2A, PIGA, PLPBP, PNKP, PNPO, PRRT2, QARS1, SCN1A, SCN2A, SCN3A, SCN8A, SLC25A12, SLC25A22, SLC2A1, SLC35A2, SLC35A3, SLC6A1, SPTAN1, ST3GAL3, STXBP1, SYNGAP1, SZT2, SYN1, SYT1, TBC1D24, TMEM63b, TPP1, UBA5, VAMP2, WDR45, WWOX, YWHAG*.
CM Panel	*ACTB, ACTG1, ADGRG1, ARFGEF2, ATP6V0A2, B3GALNT2, B4GAT1, CNTNAP2, COL18A1, CRADD, DCHS1, DCX, DEPDC5, DYNC1H1, EML1, ERMARD, FAT4, FKRP, FKTN, FLNA, GMPPB, GPSM2, GRIN1, CRPPA, KIFBP, KIF2A, KIF5C, LAMA2, LAMB1, LAMC3, LARGE, MEF2C, NDE1, NEDD4L, NPRL2, NPRL3, NSDHL, OCLN, PAFAH1B1, PAX6, PIK3R2, POMGNT1, POMGNT2, POMK, POMT1, POMT2, RAB18, RAB3GAP1, RAB3GAP2, RELN, RTTN, SCN3A, SNAP29, SRPX2, TBC1D20, TMEM5, TMTC3, TUBA1A, TUBA8, TUBB, TUBB2B, TUBB3, TUBG1, VLDLR, WDR62*.
MTOR-pathway Panel	*AKT1, AKT3, CCND2, DEPDC5, KPTN, MTOR, NPRL2, NPRL3, PIK3CA, PIK3R2, PTEN, RHEB, RPS6, STRADA, TBC1D7, TSC1, TSC2*

Supplementary Table S2. Gene counts combined

gene	n_baseline	n_new_after_reanalysis	n_total
PRRT2	10	0	10
ANKRD11	2	2	4
GRIN2A	4	0	4
SCN2A	4	0	4
CACNA1A	3	0	3
DEPDC5	3	0	3
KCNQ2	3	0	3
MECP2	2	1	3
SATB2	3	0	3
SCN1A	3	0	3
SETD5	3	0	3
SORD	3	0	3
TSC1	3	0	3
ADAR	2	0	2
ARID2	2	0	2
ASXL3	2	0	2
ATL1	2	0	2
ATP6V0C	2	0	2
BTD	1	1	2
CHD2	2	0	2
CHMP1A	2	0	2
CSNK2B	2	0	2
CTNNB1	2	0	2
DEAF1	2	0	2
GABRA1	1	1	2
GDI1	2	0	2
GRIN2B	2	0	2
IKBKG	2	0	2
ITPR1	2	0	2
KIF1A	2	0	2

NR2F1	1	1	2
NSD1	2	0	2
PTCH1	1	1	2
PTEN	2	0	2
SLC1A4	2	0	2
SLC2A1	1	1	2
SLC5A6	2	0	2
SYNGAP1	2	0	2
TSC2	2	0	2
ZEB2	2	0	2
ACAN	1	0	1
ADAMTS17	1	0	1
ADCY5	1	0	1
ADGRG1	1	0	1
ALDH18A1	0	1	1
ALG1	1	0	1
ARID1B	1	0	1
ARSA	1	0	1
ARX	1	0	1
ATAD3A	1	0	1
ATP6V0A1	1	0	1
CACNA1A	1	0	1
CDC45	1	0	1
CEP57	1	0	1
CFTR	0	1	1
CHD8	0	1	1
CLCN2	1	0	1
CLCN4	1	0	1
CNOT3	1	0	1
COL4A1	1	0	1
COL4A2	0	1	1
CSF1R	1	0	1
CUX2	1	0	1
DDHD2	0	1	1

DHDDS	1	0	1
DNM1	1	0	1
DNMT3A	1	0	1
DYNC1H1	1	0	1
DYRK1A	1	0	1
ECHS1	1	0	1
EIF2AK2	1	0	1
EMC10	1	0	1
EP300	1	0	1
EPM2A	0	1	1
FBXW11	1	0	1
FGF8	1	0	1
FGFR3	1	0	1
FLNA	0	1	1
FOXG1	1	0	1
GABRG2	1	0	1
GH1	1	0	1
GNB1	1	0	1
GPSM2	1	0	1
HCN1	1	0	1
HCN4	1	0	1
HERC1	0	1	1
HIST1H1E	1	0	1
HPRT1	1	0	1
IGHMBP2	1	0	1
KARS1	1	0	1
KAT6A	1	0	1
KCNB1	1	0	1
KMT2A	1	0	1
KMT2B	1	0	1
MACF1	1	0	1
MAOA	1	0	1
MAPK1	1	0	1
MAST1	1	0	1

MBD5	1	0	1
MED13L	1	0	1
MORC2	1	0	1
MTOR	1	0	1
MYT1L	0	1	1
NANS	1	0	1
NARS1	1	0	1
NF1	1	0	1
NFIA	1	0	1
NOTCH3	1	0	1
NPR2	1	0	1
NPRL2	1	0	1
OPHN1	1	0	1
PCDH19	1	0	1
PHIP	1	0	1
PIEZO2	0	1	1
PIGN	1	0	1
PIK3CA	1	0	1
POGZ	1	0	1
POLG	0	1	1
POLR3A	1	0	1
PPARG	1	0	1
PQBP1	1	0	1
RAI1	1	0	1
RANBP2	1	0	1
RTTN	0	1	1
SCN8A	1	0	1
SEMA6B	1	0	1
SLC6A1	1	0	1
SMARCE1	1	0	1
SMPD4	1	0	1
SOX2	1	0	1
SOX4	1	0	1
SPART	1	0	1

STXBP1	1	0	1
SZT2	1	0	1
TANC2	1	0	1
TCF4	1	0	1
TMX2	1	0	1
TOR1AIP1	1	0	1
TUBB2B	0	1	1
VPS13B	1	0	1
WDR62	1	0	1
ZBTB11	0	1	1
ZC4H2	1	0	1
ZDHHC9	1	0	1
ZNF462	1	0	1
ZNF699	1	0	1

Supplementary Tables S3-S6

Table S3: Agnostic subset: percentage positive by clinical signs and symptoms at baseline and after reanalysis. Values are n/N (%).

Phenotype	Baseline	Reanalysis
Epilepsy	44/212 (20.8%)	47/212 (22.2%)
Brain MRI abnormalities	65/183 (35.5%)	66/183 (36.1%)
Developmental delay	64/152 (42.1%)	65/152 (42.8%)
Dysmorphism	59/141 (41.8%)	63/141 (44.7%)
Intellectual disability	47/140 (33.6%)	49/140 (35.0%)
Epileptic encephalopathy	27/69 (39.1%)	28/69 (40.6%)
Brain malformations	8/32 (25.0%)	8/32 (25.0%)
Hypotonia	31/60 (51.7%)	32/60 (53.3%)
Autism/ASD	7/58 (12.1%)	8/58 (13.8%)
Spasticity/spastic paraparesis	7/16 (43.8%)	7/16 (43.8%)
Visual impairment	11/25 (44.0%)	12/25 (48.0%)
Hearing loss	2/9 (22.2%)	2/9 (22.2%)
Neurodegenerative	1/3 (33.3%)	1/3 (33.3%)

Table S4: Agnostic subset: McNemar test for baseline vs reanalysis positivity by clinical signs and symptoms. Values are n (%).

Phenotype	n	WT→Positive (n, %)	p
Epilepsy	212	3 (1.4%)	0.248
MRI abnormalities	183	1 (0.5%)	1.000
Developmental delay	152	1 (0.7%)	1.000
Dysmorphism	141	4 (2.8%)	0.134
Intellectual disability	140	2 (1.4%)	0.480
Epileptic encephalopathy	69	1 (1.4%)	1.000
Brain malformations	32	0 (0.0%)	–
Hypotonia	60	1 (1.7%)	1.000
Autism/ASD	58	1 (1.7%)	1.000
Spasticity/spastic paraparesis	16	0 (0.0%)	–
Visual impairment	25	1 (4.0%)	1.000
Hearing loss	9	0 (0.0%)	–
Neurodegenerative	3	0 (0.0%)	–

Table S5: Panel subset: percentage positive by clinical signs and symptoms at baseline and after reanalysis. Values are n/N (%).

Phenotype	Baseline	Reanalysis
Epilepsy	31/207 (15.0%)	41/207 (19.8%)
Brain MRI abnormalities	10/65 (15.4%)	17/65 (26.2%)
Developmental delay	7/38 (18.4%)	11/38 (28.9%)
Dysmorphism	4/25 (16.0%)	9/25 (36.0%)
Intellectual disability	2/39 (5.1%)	6/39 (15.4%)
Epileptic encephalopathy	6/51 (11.8%)	7/51 (13.7%)
Brain malformations	2/25 (8.0%)	4/25 (16.0%)
Hypotonia	2/13 (15.4%)	4/13 (30.8%)
Autism/ASD	0/7 (0.0%)	0/7 (0.0%)
Spasticity/spastic paraparesis	0/2 (0.0%)	1/2 (50.0%)
Visual impairment	0/3 (0.0%)	1/3 (33.3%)
Hearing loss	0/0 (–)	0/0 (–)
Neurodegenerative	0/0 (–)	0/0 (–)

Table S6: Panel subset: McNemar test for baseline vs reanalysis positivity by clinical signs and symptoms. Values are n (%).

Phenotype	n	WT→Positive (n, %)	p
Epilepsy	207	10 (4.8%)	0.004
Brain MRI abnormalities	65	7 (10.8%)	0.023
Developmental delay	38	4 (10.5%)	0.134
Dysmorphism	25	5 (20.0%)	0.074
Intellectual disability	39	4 (10.3%)	0.134
Epileptic encephalopathy	51	1 (2.0%)	1.000
Brain malformations	25	2 (8.0%)	0.480
Hypotonia	13	2 (15.4%)	0.480
Autism/ASD	7	–	–
Spasticity/spastic paraparesis	2	–	–
Visual impairment	3 –	–	–

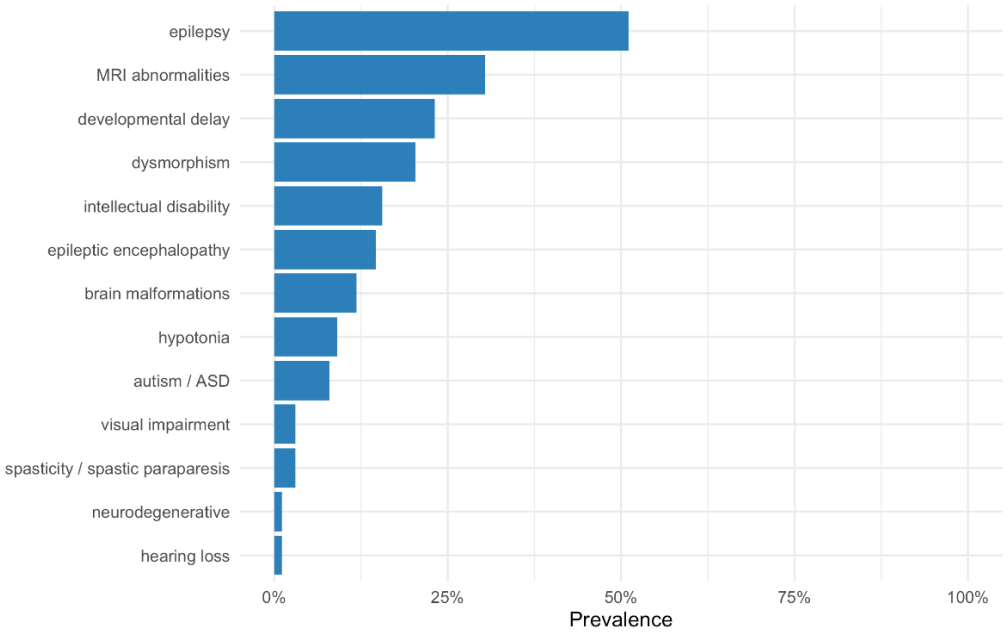
Supplementary Table S7. Secondary findings list

Sex	Incidental findings
F	*LDLR*
F	*RRM2B*
M	*ABCA4*
F	*MSH6* (maternal)
F	*HBB*
F	*G6PD*
F	*GLA*
F	*PROM1*
M	*RYR1* (maternal)
M	*TTN*
F	*G6PD*
F	*SMAD4*
M	*BRCA1*
M	*FLNC* (maternal)
M	*MEFV*
M	*HBB* (paternal)
F	*MYBPC3* (mother)
M	*HBB*
F	*HBB* (paternal); *G6PD* (mother)
F	*MIH7*
M	*HBB*
M	*HBB*
F	*HBB; G6PD*
F	*BLM* (biparental)
F	*SCN5A* (Not reported at the family's request))
F	*HBA1 e 2*
F	*HBB*
M	*SCN5A* (paternal)
M	*HBB* (maternal)
M	*HBB*
M	*TTN*

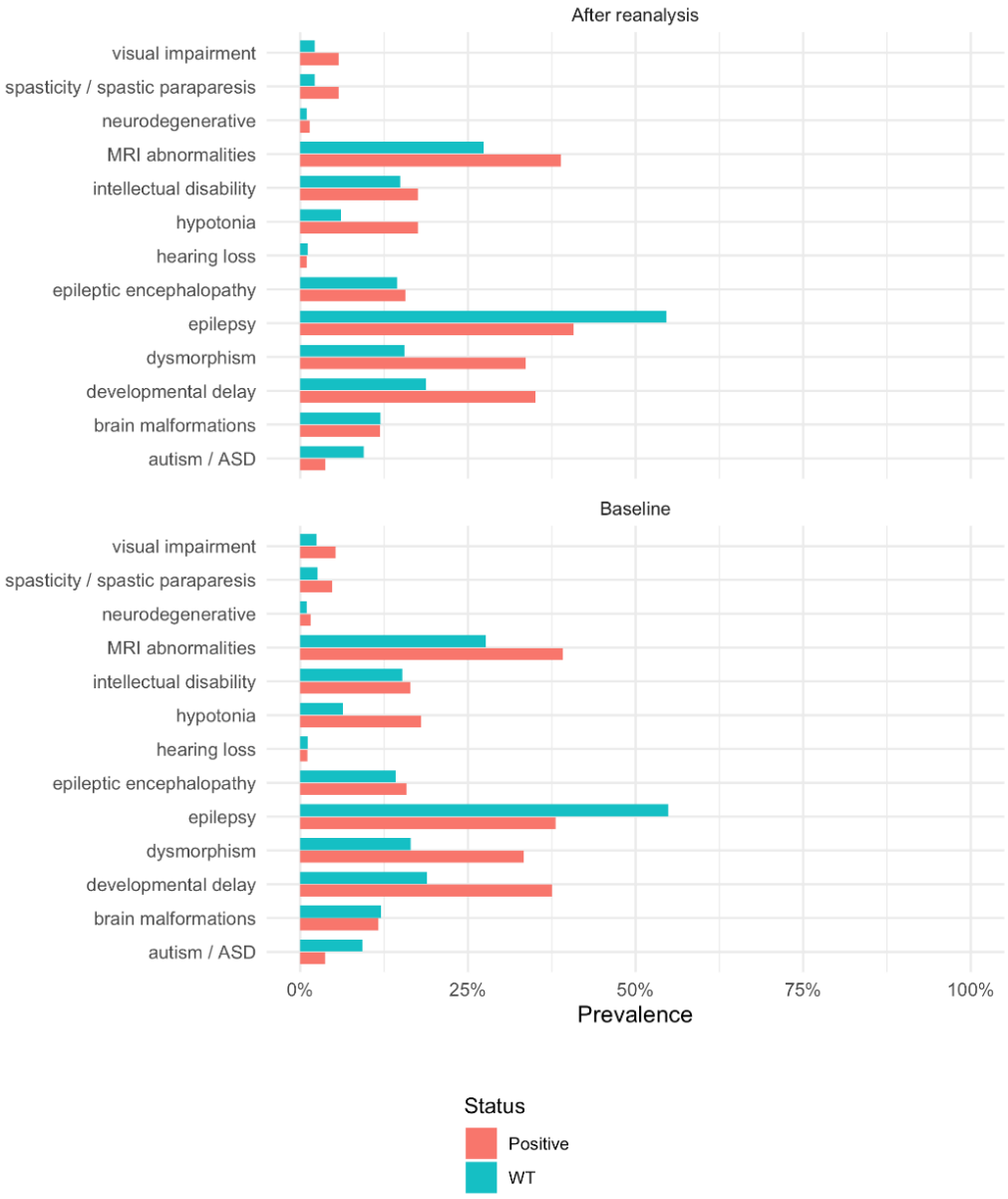
F	*HBB* (maternal)
M	*MYH7* (paternal); *HBB* (maternal)
M	*APOB* (maternal)
M	*HBB* (maternal); *MPZL2* (biparental); *TTN*
M	*LDLR*
M	*MYH7* (paternal)
M	*MYH7* (maternal)
M	*HBB*
M	*G6PD*
M	*TBC1D8B*
F	*KCNQ1* (biparental, incomplete penetrance)
F	*G6PD*
F	*LDLR* (paternal)
M	*HBA1*
M	*TTN* (paternal)
F	*LIPE* (paternal)
M	*MSH2; HFE*
F	*MUTYH* (maternal); *MUTYH* (paternal)
M	*HBB* (proband + mother)
M	*G6PD*
M	*RYR2* (paternal)

10.3 Supplementary Figures

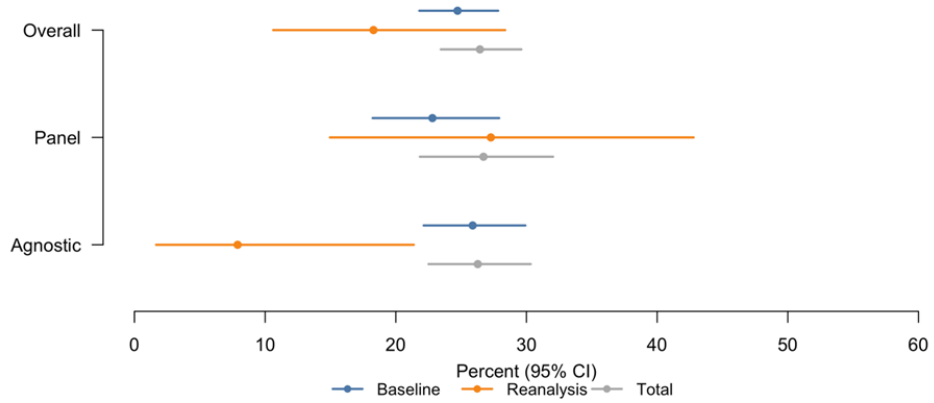
Supplementary Figure S1. Phenotype prevalence (overall cohort)



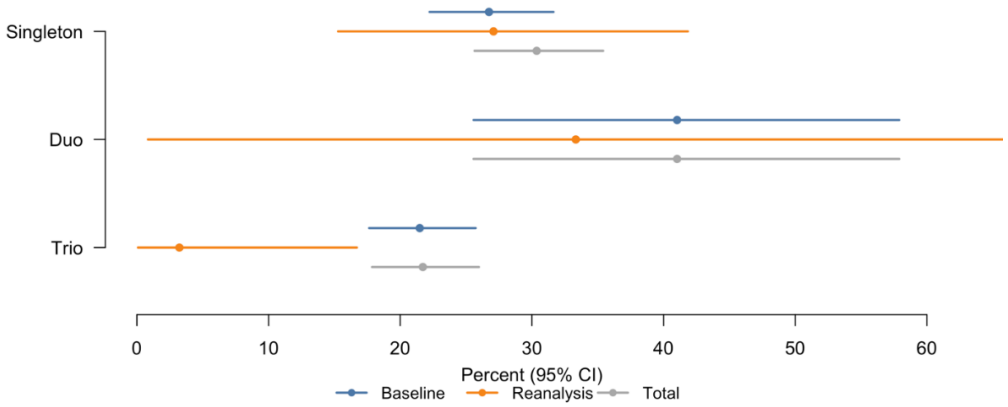
Supplementary Figure S2. Phenotype prevalence by diagnostic status (overall cohort)



Supplementary Figure S3. VUS reporting by approach (percent, 95% CI)



Supplementary Figure S4. VUS reporting by family configuration (percent, 95% CI)



Supplementary Figure S5. Secondary findings counts (per gene)