Human Genomics

RESEARCH

Open Access

Diagnostic yield of long-read sequencing in short-read-negative patients: a systematic review and meta-analysis

Ali Babazade[1], M. Fatih Mulayim[1], Yusuf Bahap[1], Mehmet Ali Ergun[1] and Meral Yirmibes Karaoguz[1*]

Abstract

Background Long-read sequencing (LRS) is increasingly used for rare disease patients undiagnosed by short-read sequencing (SRS), but its additional yield has not been synthesized at the individual-patient level or across both targeted and genome-wide approaches, and the phenotypes and rescue mechanisms driving added diagnoses remain mostly uncharacterized.

Methods We assessed the LRS contribution across rare diseases by performing a systematic literature search in PubMed, Scopus, Web of Science, EMBASE and Google Scholar (to 31 October 2025), including studies that applied LRS to patients with monogenic rare disease undiagnosed after SRS. Studies were categorized into genome-wide cohort studies (Tier 1), targeted studies (Tier 2), and case reports/series (Tier 3), and random-effects meta-analysis was performed for Tier 1 studies with prespecified subgroup analyses. Patient phenotypes were encoded as Human Phenotype Ontology terms and hierarchically aggregated to test phenotype-specific associations with rescue.

Results Across 11 Tier 1 studies (538 patients), the pooled rescue rate combining definitive and possible diagnoses was 16.4% (95% CI: 6.9–29.0%); restricting to definitive rescue only, the pooled yield was 12.0% (95% CI: 3.3–25.3%). Across 15 Tier 2 targeted studies (165 SRS-negative patients), the rescue rate was 37.6%, reaching 100% in *MUC1*-associated ADTKD (12/12) and Canavan disease (2/2 SRS-negative and 8/8 second-hit cases) but only 3.9% in NF1 (2/51). Across 119 second-hit cases (Tier 1 + Tier 2), 34.5% were rescued. Among SRS-negative patients in the primary meta-analysis cohort, yield declined as prior testing deepened (ES-only 23.5% vs ES+GS 6.8%). Structural variants accounted for ~49% of genome-wide rescues, though ~29% of rescues reflected reanalysis findings rather than LRS-specific capability. In exploratory phenotype analysis, spasticity, microcephaly, and intellectual disability were among the phenotypes most strongly associated with rescue, a signal that persisted within neurological cases.

Conclusion LRS provides a meaningful rescue for patients with rare diseases undiagnosed after SRS, primarily through structural variant detection. Re-evaluation of existing SRS data and prioritization of patients directly after negative ES may optimize diagnostic yield, and phenotype-informed prioritization is a promising direction that warrants prospective evaluation.

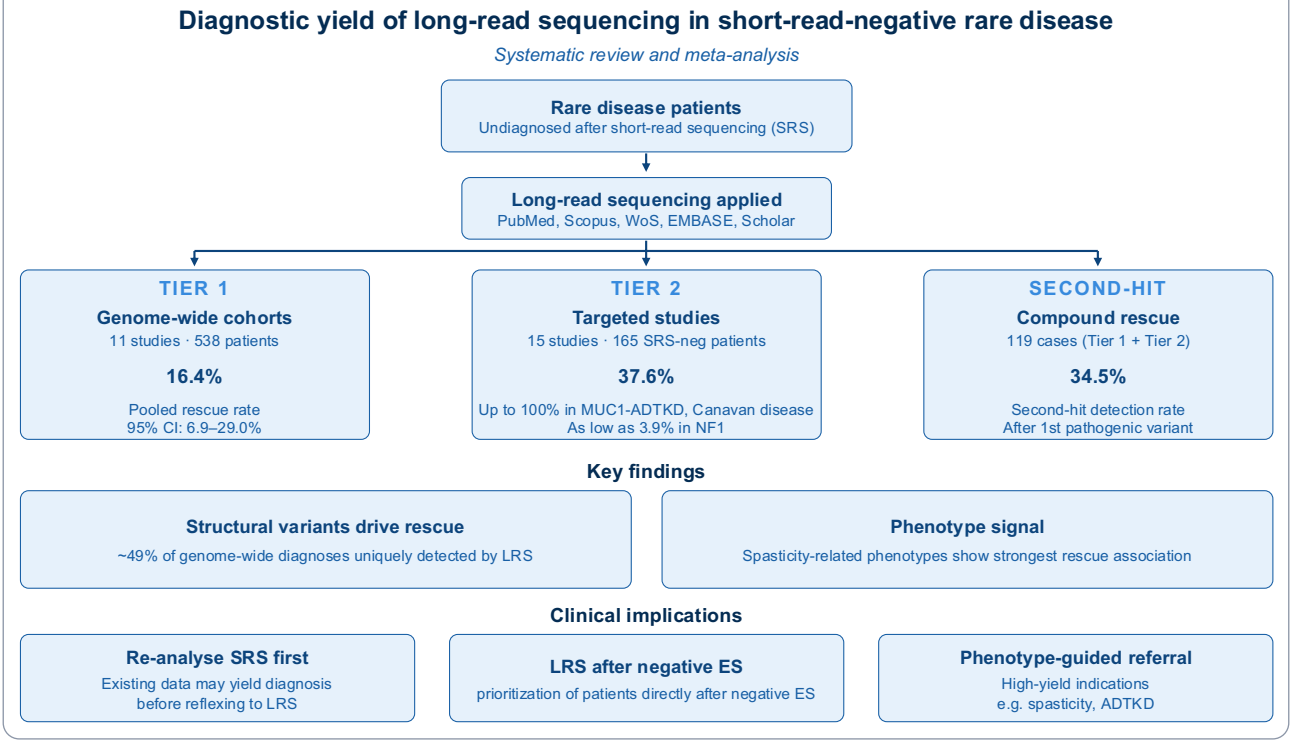
*Correspondence:
Meral Yirmibes Karaoguz
karaoguz@gazi.edu.tr; meral25@yahoo.com

Full list of author information is available at the end of the article

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Keywords Long-read sequencing, Structural variant, Short-read sequencing, Systematic review, Rare disease, Mendelian disease

Graphical Abstract



Introduction

Currently, short-read sequencing (SRS), including both exome sequencing (ES) and genome sequencing (GS), is the most commonly used genetic diagnostic test for suspected monogenic rare diseases. However, the diagnostic yield often plateaus at approximately 30–40%, leaving many patients without a molecular diagnosis [1, 2]. It is believed that a substantial proportion of these unresolved cases are due to variant types and genomic regions that SRS has limited sensitivity to detect, including structural variants (SV), repeat expansions (RE), and variants in complex or poorly mappable loci [3].

Long-read sequencing (LRS) technologies, primarily from Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT), offer significant technical advantages that can overcome these limitations; the generation of reads spanning thousands of bases improves detection of structural variants and repeat expansions, enables resolution of repetitive or GC-rich regions, and allows direct haplotype phasing without parental samples [3]. As such, LRS is being increasingly applied as a diagnostic rescue strategy for patients who have remained undiagnosed after SRS [4, 5]. Despite these advantages, clinical adoption remains limited, largely because of higher costs and infrastructure requirements compared with standard SRS [3, 6]. Studies to date have employed a range of approaches, from genome-wide LRS to targeted strategies including amplicon sequencing, adaptive sampling, and CRISPR-Cas9 enrichment, leading to considerable variability in reported diagnostic yields [3–7]. Although LRS has been systematically reviewed in specific disease groups such as inherited retinal dystrophies [8], the only broad synthesis to date evaluated long-read whole-genome sequencing in 646 previously unresolved cases and reported an overall additional yield of 4.5%, with structural variants making up the largest share of diagnoses [9]. That review, however, was limited to genome-wide long-read sequencing and did not include targeted assays, pool diagnostic yield into a meta-analytic estimate, stratify yield by prior short-read testing, or rank phenotypes by expected benefit [9]. It also drew only on cohort studies, omitting the individual case reports that document many early long-read rescues. Taken together, no synthesis has yet combined targeted and genome-wide long-read data, pooled diagnostic yield across studies, and linked specific phenotypes to their expected benefit—leaving clinicians without an evidence base for deciding whom to prioritize for long-read testing.

To address these gaps, we performed a systematic review and published individual patient data (IPD)

meta-analysis of LRS used in SRS-negative rare Mendelian disease cases. By analyzing patient data, our aims were to: (1) provide a pooled estimate of the additional diagnostic yield of LRS; (2) identify the phenotypic categories that predict a greater likelihood of achieving a diagnosis; and (3) evaluate the specific rescue mechanisms contributing to resolving these diagnostic odysseys.

Materials and methods

Search strategy and study selection

Systematic literature searches were performed across five electronic sources (PubMed/MEDLINE, Embase, Scopus, Web of Science, and Google Scholar) on 31 October 2025. Google Scholar was searched as a supplementary source (first 200 results by relevance; results may differ over time and by location). Records were imported into EndNote and deduplicated by automated title/author/year matching with manual review. The search strategy consisted of three concept blocks combined using Boolean AND logic: (1) long-read sequencing technologies (e.g., nanopore sequencing, PacBio, HiFi, SMRT, and third-generation sequencing); (2) diagnostic context (e.g., diagnosis, undiagnosed, unsolved); and (3) genetic disease (e.g., rare diseases, Mendelian, monogenic, genetic disorders, germline). Full search queries are available in Supplementary Table S1.

This review was conducted and reported in accordance with the PRISMA 2020 statement and the PRISMA-IPD extension, following the registered protocol (PROSPERO CRD420251275558; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251275558). Individual patient data were obtained by extraction from the included publications and their supplementary materials rather than as raw datasets from the original investigators; as provided for under PRISMA-IPD (item 10), we therefore describe these as published IPD. Data were extracted independently and in duplicate by two reviewers, and IPD integrity was assessed through internal consistency and completeness checks in place of investigator-confirmed data. The following study types were excluded: STR-only studies (LRS data analyzed exclusively for short tandem repeats), population-level carrier screening studies, technical benchmarks without patient-level diagnostic outcomes, non-English publications, and conditions diagnosed routinely by a disease-specific assay rather than short-read sequencing, in which no SRS-negative population arises (thalassemia, congenital adrenal hyperplasia, hemophilia; Supplementary Methods S2.1). No restriction was placed on publication status. Peer-reviewed articles and full-data grey literature, including preprints and theses, were eligible provided they met all inclusion criteria and reported extractable patient-level data; conference abstracts without full data were excluded.

Tier classification

The studies were classified into three tiers based on the LRS approach and study design. Tier 1 included genome-wide LRS cohort studies, Tier 2 included targeted LRS studies using gene-specific or panel-based approaches, and Tier 3 comprised individual case reports/series or small case series (n<5) applying LRS in rare disease patients.

Two reviewers (A.B. and M.F.M.) independently screened records using EndNote, with disagreements resolved by a third reviewer (M.Y.K.). After title, abstract screening and full-text assessment, 37 studies were included as Tier 1 and Tier 2, and 59 included as Tier 3. In total, 96 studies met inclusion criteria. A detailed list of excluded studies with exclusion reasons is provided in Supplementary Table S2a (Tier 1 and 2) and Supplementary Table S2b (Tier 3 case reports and series).

Study grouping and data extraction

The 22 studies in Tier 1 were assigned to predetermined study groups based on their designs and population characteristics. The primary group for the meta-analysis consisted of Group A, which included 11 studies utilizing a sequential SRS-to-LRS design with ≥5 SRS-negative patients per cohort, contributing 538 SRS-negative patients. The remaining 11 Tier 1 studies were analyzed separately due to methodological heterogeneity, specifically parallel or first-line LRS designs, second-hit/missing-allele cohorts, mixed cohorts whose SRS-negative subcohort fell below the n=5 threshold, expert-preselected cohorts, SV-only analyses, and autosomal recessive (AR) suspected cohorts (full rationales in Supplementary Table S10).

Patient-level data were extracted from each study into a standardized sheet (Dataset S1). For each patient, the following were recorded: prior SRS testing and outcomes (ES, GS, gene panel), LRS platform and sequencing chemistry, LRS family structure (singleton, duo, trio, or larger), patient phenotype, LRS outcome, rescue mechanism(s), rescued variant type and size, and ACMG classification.

Outcome definitions

SRS sequencing outcomes were classified as negative (no pathogenic or likely pathogenic variant identified), positive (molecular diagnosis established), inconclusive (incomplete findings), second-hit missing (one pathogenic allele identified in an autosomal recessive gene with the second allele undetected), or phasing unknown (variants detected but phase not resolved).

Long-read sequencing outcomes were classified as: Definitive Rescue (pathogenic or likely pathogenic variant establishing a molecular diagnosis), Possible Rescue (a variant of uncertain significance or candidate-gene finding judged clinically relevant by the original authors, not a confirmed diagnosis), Refined (additional characterization of known variants without establishing a new molecular diagnosis, such as breakpoint resolution, phasing confirmation), Missed (a clinically relevant variant established by prior or parallel testing that LRS did not identify), and Not Rescued. The primary outcome was the combined rescue rate, defined as the proportion of SRS-negative patients achieving Definitive or Possible Rescue (full outcome-category definitions are provided in Supplementary Methods S3.1; the combined outcome was not prespecified, and the definitive-only rate is reported alongside it throughout). Refined and Missed cases were retained in the denominator of any analysis frame for which the patient was eligible and counted as non-rescues; no Missed case occurred in the primary meta-analysis cohort.

Rescue mechanisms for combined rescue cases were defined as follows: SV detection as identification of variants larger than 50 base pairs; full-length sequencing as analysis of regions not covered by prior SRS; dark region resolution as analysis of genomic regions where SRS failed to map, or detect variants reliably; phasing as determination of whether variants were in cis or trans; methylation detection as assessment of methylation status contributing to diagnosis; repeat expansion detection as identification/characterization of pathogenic short tandem repeat expansions; reanalysis findings as diagnoses made from LRS data where the variant was identified through updated gene-disease knowledge or annotations rather than through LRS-specific technical capabilities; and mosaicism detection as identification of variants present in only a subset of cells.

Phenotype categorization

Where Human Phenotype Ontology (HPO) based phenotypic data were available, they were extracted directly; for patients for whom clinical data was given in narrative form, two clinicians (A.B. and M.F.M.) independently encoded each patient's clinical features into HPO codes, with 86.9% exact and 97.5% overall inter-rater concordance (details in Supplementary Methods S6.10). Terms directly reported and inferred from clinical descriptions were merged per patient. A bottom-up aggregation algorithm traversed from each patient's specific terms upward through the hierarchy, designating any ancestor node with ≥5 patients as a phenotype category (minimum depth of 2, excluding the HPO root and 'Phenotypic abnormality'). Each patient was assigned to all qualifying ancestor categories simultaneously.

Statistical analysis

Proportions were pooled using random-effects meta-analysis with the Freeman–Tukey double arcsine transformation; between-study variance ($\tau 2$) was estimated by restricted maximum likelihood (REML). Heterogeneity was evaluated by Cochran's Q statistic and the I2 statistic, and a 95% prediction interval was calculated. Leave-one-out sensitivity analysis was performed to evaluate the influence of individual studies. Publication bias was assessed by funnel plot inspection and Egger's regression test. Risk of bias was assessed for each study with an adapted QUADAS-2 tool (Supplementary Methods S4; full instrument and mapping to the standard tool in Supplementary Tables S3a and S3b) [10].

Confidence intervals for proportions were determined using the Wilson score method. Fisher's exact test was applied for categorical comparisons with Benjamini–Hochberg FDR correction at $q = 0.05$. A Cochran–Armitage trend test assessed whether rescue rates declined with increasing prior testing depth. Subgroup meta-analyses were performed by predominant prior testing type and LRS platform, and rescue rate was additionally compared by sequenced family members.

For phenotypic analysis, Fisher's exact test compared rescue rates inside versus outside each HPO category (≥ 5 patients). Odds ratios with 95% confidence intervals were computed using the log-OR method with Haldane-Anscombe correction, with Benjamini–Hochberg correction (FDR 5%). Because HPO categories are nested and therefore non-independent, this correction was repeated as a sensitivity analysis using the dependence-robust Benjamini-Yekutieli procedure (Supplementary Methods S6.8; Dataset S2). Because nesting also causes the same patients to appear under multiple ancestor terms, we did not enumerate all significant categories in the main text; instead, within each ancestor–descendant lineage we report the single most specific term whose association was estimable (finite odds ratio, i.e. not the result of complete separation), reproducible (at least 10 patients from at least 3 independent studies), and robust across the study-level confounding analyses, with the full set provided in Dataset S2 (Supplementary Methods S6.9). Two sensitivity analyses were performed: (1) exclusion of 156 patients lacking HPO annotation; and (2) a within-neurology analysis restricting the denominator to the 272 neurological patients (Supplementary Methods S6.7). To assess study-level confounding, we performed three analyses on the significant categories. We quantified single-study concentration (the share of each category's patients, and of its rescued patients, from the largest single study). We repeated the Fisher's exact and Benjamini–Hochberg pipeline after removing the single-disease-enriched cohorts. And we refitted each category with a random study intercept (rescue ~

category + (1|study); lme4::glmer), using a Bayesian fit (blme::bglmer) where all in-category patients were rescued and the frequentist odds ratio is not finite.

Tier 2 studies were analyzed by narrative synthesis. Tier 3 case reports/series were analyzed descriptively. All statistical analyses and plots were conducted in R (version 4.4).

Results

Study selection and cohort overview

A total of 96 studies eligible for review were found through the systematic literature search conducted during this project. These studies included 22 Tier 1, 15 Tier 2, and 59 Tier 3 articles (Fig. 1).

In total, 22 Tier 1 studies collectively recruited 923 patients with different rare disease phenotypes, of whom 684 (74.1%) had prior negative SRS results [4, 5, 11–30]. The primary meta-analysis cohort (Group A) was composed of 11 studies which used a sequential SRS-to-LRS approach, contributing 538 SRS-negative patients [4, 5, 12, 13, 17, 20, 24, 25, 27–29]. The remaining 11 Tier 1 studies and their 146 SRS-negative patients were analyzed separately because of methodological heterogeneity, including parallel or first-line LRS designs, targeted disease subgroups, second-hit/missing-allele cohorts, mixed-design cohorts in which the SRS-negative subcohort fell below the n=5 inclusion threshold, expert-preselected cohorts, limited-scope SV-only analysis, and AR-suspected cohorts (full per-study rationales: Supplementary Table S10) [11, 14–16, 18, 19, 21–23, 26, 30].

Tier 2 comprised 15 targeted LRS studies across nine disease areas and included 412 patients whose individual

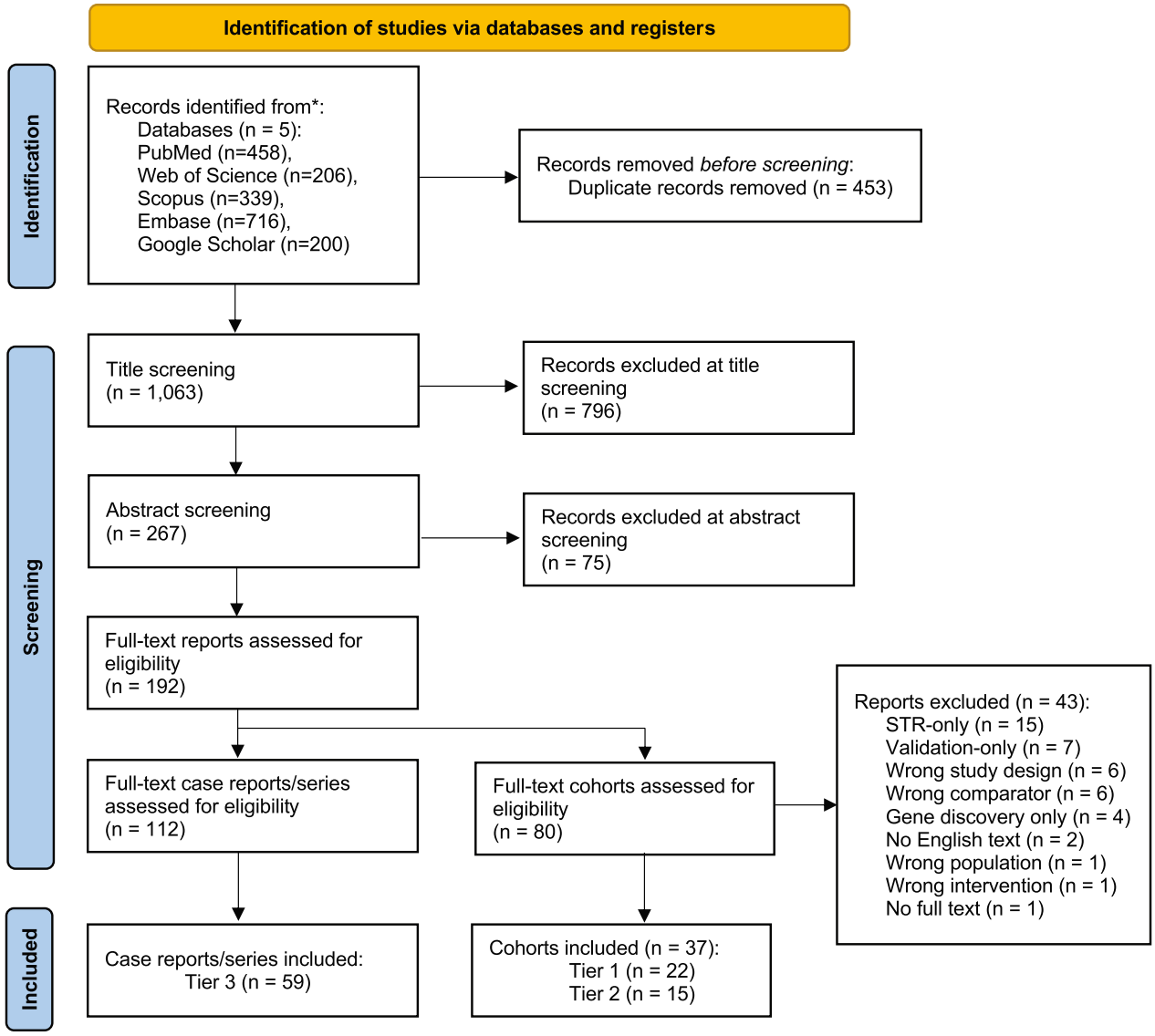


Fig. 1 Study selection (PRISMA review protocol). Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram

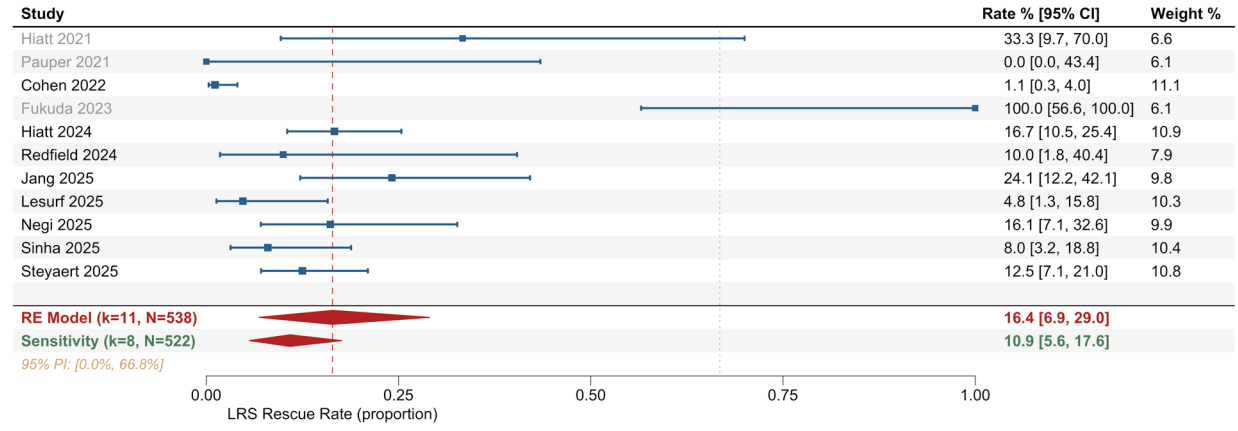
participant data were available [31–45]. Among the patients, 252 met criteria for inclusion in the analysis frame (165 SRS-negative, 37 second-hit/missing allele, 30 inconclusive, 20 phasing-unknown). The SRS-negative patients were analyzed as the primary Tier 2 cohort; second-hit cases were combined with the Tier 1 second-hit studies in a cross-tier analysis (see below). Tier 3 comprised 79 patients from 59 studies applying LRS in rare disease patients [46–104].

Primary meta-analysis of genome-wide LRS rescue

Among 538 SRS-negative patients across 11 Group A studies, 55 achieved diagnostic rescue (36 definitive, 19 possible). There was substantial variability in study-level rescue rates, ranging from 0% (Pauper et al. [13], 0/5) to 100% (Fukuda et al. [25], 5/5), reflecting heterogeneity in cohort composition, phenotype mix, and prior testing depth. The crude patient-weighted rate of rescue across the 11 Group A studies was 10.2% (55/538; 95% Wilson CI: 7.9–13.1%). The Freeman-Tukey double arcsine random-effects model estimated that the pooled rate of rescue among the 11 studies included in Group A was 16.4% (95% CI: 6.9–29.0%), with considerable heterogeneity (I^2 = 85.2%, Q = 67.5, df = 10, $p < 0.001$, $\tau^2 = 0.053$). The 95% prediction interval was 0–66.8% (Fig. 2a).

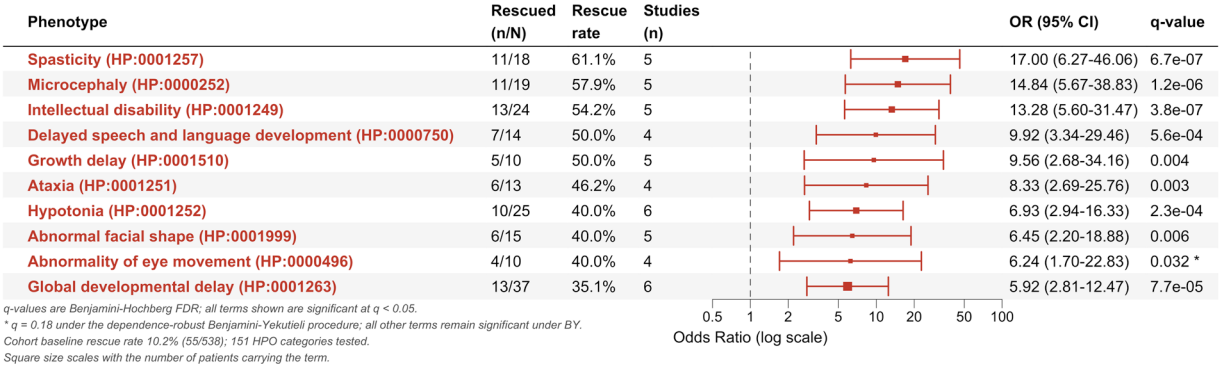
This divergence reflects the balanced weighting inherent to random-effects models. Because there was such a large amount of variation in the rate of rescue among the 11 studies in Group A ($\tau^2 = 0.053$), the weights assigned by the random-effects model were relatively uniform

a



Heterogeneity: I2 = 85.2%, Q = 67.48 (df = 10, p < 0.001), tau2 = 0.0526 (arcsine scale); Egger's regression p = 0.043.
Dashed red line, pooled estimate (16.4%); dotted gold line, upper bound of the 95% prediction interval (66.8%).
Small studies (N <= 6; Hiatt 2021; Pauper 2021; Fukuda 2023) shown in grey and excluded from the sensitivity analysis (k = 8, N = 522).

b



q-values are Benjamini-Hochberg FDR; all terms shown are significant at q < 0.05.
** q = 0.18 under the dependence-robust Benjamini-Yekutieli procedure; all other terms remain significant under BY.*
Cohort baseline rescue rate 10.2% (55/538); 151 HPO categories tested.
Square size scales with the number of patients carrying the term.

Fig. 2 Pooled LRS rescue rates and phenotype associations. **a** Forest plot of LRS rescue rates across 11 genome-wide cohort studies (N = 538). Random-effects pooled estimate: 16.4% (95% CI: 6.9–29.0%; I^2 = 85.2%); sensitivity analysis excluding small studies (N ≤ 6; Hiatt 2021; Pauper 2021; Fukuda 2023, shown in grey): 10.9% (95% CI: 5.6–17.6%; k = 8, N = 522). Dashed red line, pooled estimate; dotted gold line, upper bound of the 95% prediction interval (0.0–66.8%). **b** Non-redundant HPO phenotype associations with LRS rescue. Among the 65 categories significant after Benjamini–Hochberg correction, we show the most specific term in each ancestor–descendant lineage that was estimable, reproducible, and robust to study-level confounding (Supplementary Methods S6.9; Dataset S2). Terms are ordered by odds ratio and shown with rescued patients (n/N), rescue rate, number of contributing studies, odds ratio with 95% CI on a log scale, and q-value; square size scales with the number of patients carrying the term. n = 55 rescued of 538. CI, confidence interval; FDR, false discovery rate; HPO, Human Phenotype Ontology; LRS, long-read sequencing; OR, odds ratio

across all studies (6.1–11.1%). For example, Cohen et al. (N = 176, 1.1% rescue) received 11.1% of the total weight, versus its 32.7% proportional share of Group A patients, as expected under high τ2 [12].

Excluding the three smallest studies (N ≤ 6) resulted in a pooled estimate of 10.9% (95% CI: 5.6–17.6%; k = 8, PI: 0.1–35.3%, I2 = 82.2%). This pooled estimate of approximately 11%, based upon studies with sufficient sample sizes, may represent a more practical estimate for use when planning future LRS implementation. The definitive-only analysis yielded 12.0% (95% CI: 3.3–25.3%; I2 = 82.9%), indicating that the inclusion of possible rescue cases did not substantially alter the overall conclusion. Additionally, in a post hoc sensitivity analysis excluding the three preprints (Jang et al., Lesurf et al., Redfield et al.), the pooled rescue rate was 18.8% (95% CI 4.8–39.1%; k = 8, N = 457; I2 = 88.5%), consistent with the primary estimate and indicating that the inclusion of non-peer-reviewed sources did not inflate the pooled yield.

Leave-one-out analysis showed moderate sensitivity: removing Fukuda et al. [25] (100% rescue) produced the largest downward shift to 11.4%, while removing Cohen et al. [12] produced the largest upward shift to 18.9% (Supplementary Table S5). Egger's regression test was suggestive of small-study effects (intercept = 3.16, p = 0.043 Supplementary Table S7; funnel plot in Supplementary Figure S1) [105]. Restricting to the seven Group A studies rated low risk of bias on all four adapted QUADAS domains yielded a pooled rescue of 13.5% (95% CI: 7.8–20.6%; I2 = 40.4%; n = 7, N = 238), consistent with the primary estimate and with markedly reduced heterogeneity (Supplementary Table S6). Rescue rate did not differ by LRS family structure (singleton 10.8%, duo 10.8%, trio 9.2%), and this held after adjustment for study of origin and prior testing depth (Supplementary Table S6).

Analyses of meta-analysis excluded studies

A total of 11 studies (296 patients) excluded from the primary meta-analysis were evaluated separately (Supplementary Table S11). Three parallel/first-line studies (n = 144) were excluded due to design heterogeneity. Two parallel-design studies reported SRS-negative rescue rates of 0% (0/66) [30] and 7.1% (1/14) [22]. The third study, by Kamolvisit et al. applied LRS as a first-line test without prior SRS in 18 patients, achieving diagnosis in 11/18 (61.1%), and is therefore non-comparable to the SRS-negative framework [21]. The expert-selected movement disorder cohort by Sorrentino et al. reported a whole-cohort rescue rate of 92.9% (13/14; only one patient was SRS-negative); however, this rate reflects ascertainment bias from a preselected patient cohort with certain suspicions before LRS [23]. The AR-suspected cohort by AlAbdi et al. had a headline rescue rate of 38.2% (13/34); however, approximately half (n = 7) of these rescues were attributable to reanalysis, with an LRS-unique rescue rate of 17.6% (6/34) [18]. In the SV-only analysis by Shah et al. the rescue rate was 4.3% (1/23) [11]. Five Tier 1 cases were classified as Missed, in which a clinically relevant variant established by prior or parallel testing was not identified by LRS, reflecting low coverage (n = 2), pipeline misclassification of a complex indel (n = 1), or failure to detect (n = 1) or prioritize (n = 1) a variant (Supplementary Table S16). None occurred in the primary meta-analysis cohort.

Effect of prior testing depth

Of 453 Group A cases (excluding 85 cases with ambiguous prior testing: 77 Steyaert et al. [5] and 8 Negi et al. [17]), the proportion of cases rescued decreased with the depth of prior testing performed: 23.5% of ES-only cases (16/68), 12.8% of GS-only cases (23/179), and 6.8% of ES+GS cases (14/205); a single panel-only case was also rescued. There was a statistically significant enrichment of cases which had been tested with only ES (OR = 2.88, p-adj = 0.005) and a statistically significant depletion of cases which had been tested with both ES and GS (OR = 0.39, p-adj = 0.005), but no difference for GS-only (p-adj = 0.553). A statistically significant dose–response relationship was identified using the Cochran–Armitage trend test, though this should be interpreted with caution given single-study influence (see below).

The majority of this gradient can be attributed to Cohen et al., who were responsible for 164 of 205 ES+GS cases with a rescue rate of only 1.2% (2/164) [12]. Among the 41 non-Cohen ES+GS cases, there was a 29.3% (12/41) rescue rate, indicating that the lower ES+GS rescue rate reflects Cohen et al. having done an unusually large amount of prior testing rather than dual-platform SRS itself causing the lower rescue rate.

Rescue mechanisms and variant characteristics

Among the 55 rescued Group A cases, the most common rescue mechanism was structural variant (SV) detection at 49.1% (27/55), comprising deletions (n = 17), inversions (n = 4), duplications (n = 3), complex SVs (n = 2), and one mobile element insertion. Rescued SVs ranged from 57 bp to 126 Mb. The next most common rescue mechanism was reanalysis (n = 16), followed by repeat expansion (n = 8), phasing and full-length sequencing (n = 5 each), dark region resolution (n = 2) and methylation detection (n = 4). Some patients had multiple rescue mechanisms. There were no mosaicism-related rescues in the Group A cohort (Fig. 3). Across the full 923 Tier 1 patients, ME–mediated variants were identified in 6 cases (0.7%; 3 ME insertions and 3 ME-mediated rearrangements).

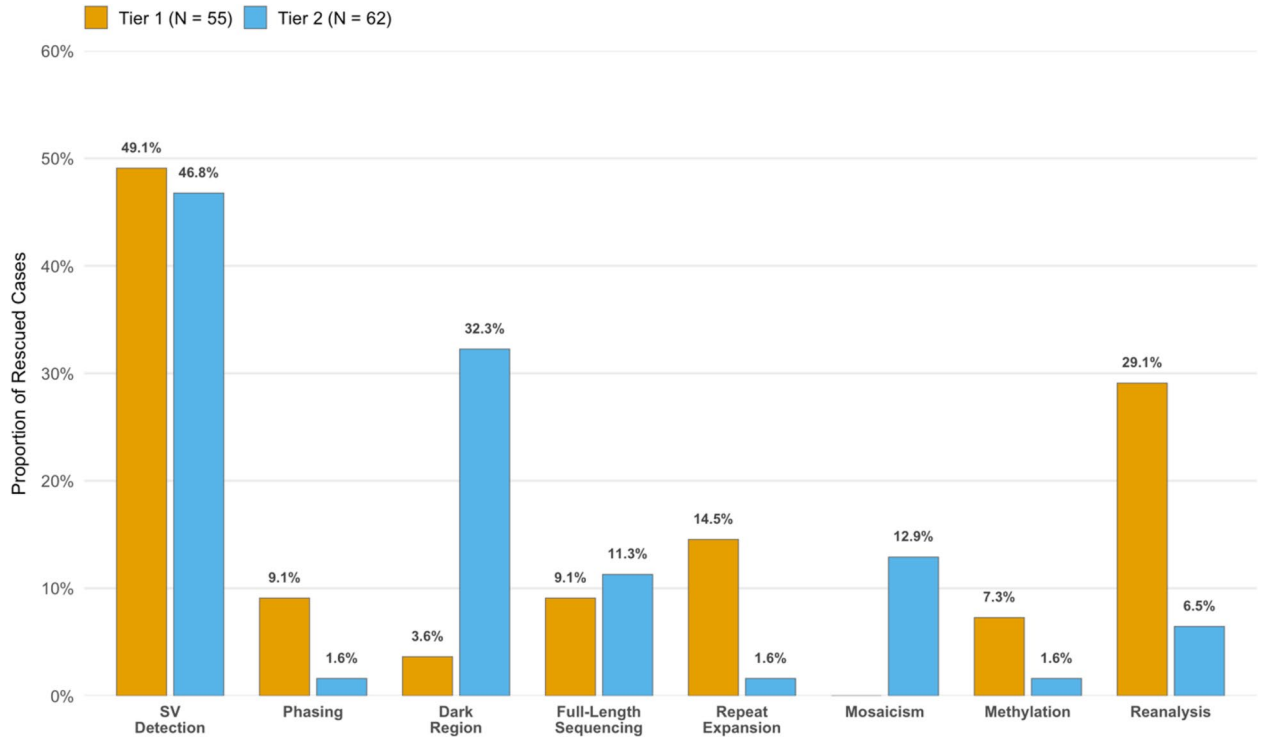


Fig. 3 Distribution of rescue mechanisms across Tier 1 and 2. Proportion of rescued cases attributed to each mechanism category, stratified by evidence tier: Tier 1, genome-wide LRS SRS-negative patients (Group A, 55 rescued); Tier 2, targeted LRS in SRS-negative patients (62 rescued). Multiple mechanisms may be annotated per case; percentages do not sum to 100%

Mechanisms of rescue varied among the different prior testing categories. SV detection dominated ES-only rescues (12/16 or 75.0%), because ES is incapable of detecting non-coding structural variants. Reanalysis was concentrated in GS-only cases (13/23 or 56.5%), indicating that many of these rescues reflected updated knowledge, as opposed to LRS-specific technical superiority. Excluding reanalysis, a majority of the rescued cases had one mechanism for their rescue (n = 30), while 6 cases had two mechanisms and 3 cases had three. Of the remaining 16 cases, 15 were rescued through reanalysis alone, and one was not attributable to any of the defined mechanisms (a 1 bp *SHANK3* frameshift deletion missed by SRS through stochastic loss of alternate-allele reads).

Phenotype associations with LRS rescue

Of 538 patients, 382 (71.0%) had HPO-coded phenotypes, yielding 288 unique terms aggregated into 151 categories (Dataset S2). The "neurological" phenotype (HP:0000707 "Abnormality of the nervous system") was the most represented organ system (272 patients, 50.6%). Overall, 55 patients (10.2%) achieved LRS rescue (36 definitive, 19 possible), with nervous system affected patients accounting for 43 of 55 rescues (78.2%) (Dataset S2).

Phenotype-specific rescue rates and phenotype sensitivity analyses

Of the 151 categories tested, 65 were significant after Benjamini–Hochberg correction. Because HPO categories are nested, these largely reflected the same patients counted at different levels of the ontology, so for each lineage we report only the most specific term with an estimable, reproducible, and confounding-robust association (Supplementary Methods S6.9; full results in Dataset S2). Ten phenotypes met these criteria (Fig. 2b), led by spasticity (61.1% rescued, OR 17.00, 95% CI 6.27–46.06), microcephaly (57.9%, OR 14.84), and intellectual disability (54.2%, OR 13.28), each five- to six-fold above the 10.2% cohort baseline and supported by five independent studies. By contrast, several well-represented categories yielded no rescues, including myopathy (0/20) and renal abnormalities (0/16).

Excluding patients without HPO phenotypes reduced FDR-significant categories from 65 to 55, with odds ratios decreasing ~31% across all categories. Notably, the two broadest nervous system categories lost significance ($q > 0.1$), indicating the broad neurological signal was partly inflated by the comparison group. In the within-nervous system analysis (baseline rescue rate 15.8%), 34 of 51 sub-categories remained significant. Specific phenotypes—spasticity (OR = 10.90), abnormal brain morphology (OR = 7.23), and intellectual disability (OR =

8.59)—drove rescue rates well above the nervous system baseline. Conversely, neurodevelopmental delay showed lower rescue rate than other neurological patients (10.7% vs 24.3%, q = 0.006) (Supplementary Dataset S2).

Because phenotype composition is correlated with study design, it was tested whether these associations were driven by individual cohorts (Supplementary Dataset S2). First, for most categories the rescued cases were spread across several studies, but a minority were dominated by one, notably the reflex categories (all rescues from Fukuda et al.) and gait disturbance (five of eight rescues from the same cohort). Second, we screened Dataset S1 for single-disease-enriched cohorts (those enrolling one disease or organ) and repeated the analysis with these removed (Fukuda et al. [spastic paraplegia], Lesurf et al. [congenital heart disease], Redfield et al. [sensorineural hearing loss]); 59 of 65 categories remained significant. Cohorts defined by a broad phenotypic category, such as neurodevelopmental disorder, were not treated as single-disease-enriched and were retained. Third, a mixed-effects logistic regression with random study intercepts was concordant: all 65 categories retained a positive point estimate, 60 with a confidence interval excluding 1. Across all checks, 58 categories were robust by every criterion. The phenotypes reported in Fig. 2b, together with the nested terms they subsume (including spastic paraplegia and abnormal brain morphology), remained robust throughout, whereas the categories that did not (reflex abnormalities, gait disturbance, broad "abnormality of movement", cognitive impairment, atypical behavior, and short stature) were single-study-driven or non-specific and are best regarded as cohort-driven. Rescue was also more frequent in patients with more affected organ systems, but the number of systems was collinear with the number of reported HPO terms and neither remained associated with rescue when modelled together, indicating that this pattern reflects phenotyping depth rather than multisystem involvement itself (Supplementary Section S14, Table S18).

Tier 2—targeted LRS

Across 15 targeted LRS studies, 90 of 252 patients (165 SRS-negative, 37 second-hit/missing allele, 30 inconclusive, 20 phasing-unknown) were rescued using LRS (35.7%). Diagnostic rescue was achieved in 62 of 165 SRS-negative patients (37.6%, 95% CI: 30.5–45.2%); additionally, 6 of 20 phasing-unknown cases were rescued, while 14 of 30 inconclusive cases were refined; across all categories 27 of 252 patients had their diagnosis refined (10.7%). Rescue rates varied dramatically by disease context.

In SRS-negative subsets across studies, *MUC1* associated tubulointerstitial kidney disease (ADTKD) was rescued at 100% (12/12), with all variants being cytosine duplications in the SRS-refractory VNTR dark region [45]. Tuberous sclerosis complex (TSC) showed 53.8% rescue (14/26), predominantly through somatic mosaic structural variants [33]. Autosomal dominant polycystic kidney disease demonstrated 57.1% rescue (20/35) across two cohorts through pseudogene resolution in *PKD1* [39, 42]. The hereditary cancer cohort showed 25.8% rescue (8/31), spanning several genes and mechanisms including methylation detection in *MLH1* [41]. Notably, neurofibromatosis type 1 (NF1) yielded the largest negative finding: only 2/51 (3.9%) were rescued, suggesting most SRS-negative NF1 patients harbor variants beyond amplicon-based T-LRS [43].

Among the 62 SRS-negative rescued cases, mechanism analysis (categories not mutually exclusive) revealed SV detection in 29 (46.8%), dark region resolution in 20 (32.3%), mosaicism in 8 (12.9%), full-length sequencing in 7 (11.3%), reanalysis in 4 (6.5%), and repeat expansion, phasing, and methylation detection in 1 case each (1.6%). LRS platform comparison showed similar rescue rates for ONT (14/38, 36.8%) and PacBio (47/126, 37.3%), indicating no platform difference when restricted to this referral category (one mixed-platform case excluded). ME involvement was identified in 20 of 412 Tier 2 patients (4.9%).

Missing second-hit

Tier 1 LRS rescued 19/82 cases (23.2%) and Tier 2 targeted LRS rescued 22/37 cases (59.5%). Tier 1 rescues were driven largely by *PRKN* structural variants in Parkinson disease (7/48) [15, 16]. Additionally, of 15 retinal disease second-hit cases, 3 were rescued [26]. In Tier 2, complete rescue was achieved in Canavan disease (8/8, all ME insertions in *ASPA*) [32, 40]. Combined, 41/119 second-hit cases were rescued (34.5%). The mechanism profile was dominated by structural variant detection (n = 31) and phasing (n = 27), followed by full-length sequencing (n = 6), and dark-region resolution (n = 2).

Tier 3—case reports/series

Tier 3 comprised 79 patients from 59 case reports/series, spanning SRS-negative (n = 42), inconclusive (n = 22), phasing-unknown (n = 7), and second-hit missing (n = 8) cases. Across the SRS-negative subset, 40/42 cases were rescued, 1 was refined, and 1 was not rescued. Among the 40 rescued cases, the rescue mechanisms were: SV detection (n = 31), full-length sequencing (n = 6), repeat expansion (n = 5), dark region resolution (n = 3), and mosaicism (n = 1); mechanisms are non-exclusive (Supplementary Table S15).

Discussion

This systematic review provides the first meta-analytic synthesis of LRS diagnostic yield in SRS-negative patients through a published individual-patient data meta-analysis. Overall, the random effects meta-analysis revealed a pooled rescue rate of 16.4% (95% CI: 6.9–29.0%) with a 95% prediction interval of 0–66.8%. Restricting to definitive rescues, which corresponds to the registered primary outcome, gives 12.0% (95% CI: 3.3–25.3%). Excluding small studies, the estimated pooled rescue rate is 10.9%, and restricting to studies at low risk of bias yields 13.5% (I^2=40.4%). We consider these sensitivity-derived estimates to be the most pragmatic for planning the prospective implementation of LRS. Heterogeneity between the studies analyzed was considerable (I^2 = 85.2%), reflecting genuine differences in cohort composition and prior testing depth rather than statistical noise. Sensitivity analysis based on leave-one-out further supported the idea that heterogeneity was due to real differences in study cohorts. The results of the current study indicate that LRS is a meaningful, yet not universally applicable rescue strategy, and the yield is highly dependent upon the clinical context.

Prior SRS testing depth emerged as a potential correlate of LRS rescue rate. Statistical analysis showed an apparent dose–response relationship between the degree of prior testing and the rescue rate (ES-only 23.5%, GS-only 12.8%, ES+GS 6.8%). However, the very low rescue rate in the combined ES+GS group was largely driven by a very large cohort by Cohen et al. that underwent an exceptionally thorough prior evaluation including linked-read GS [12]. Among the 41 non-Cohen ES+GS cases, the rescue rate was 29.3% (12/41)—substantially higher than the gradient implies. The robust finding could therefore be an enrichment of rescues among ES-only patients, where ES does not interrogate non-coding regions and has limited capability for SV detection—variant classes that made up three-quarters of rescues from the exome-only group. Further studies are required to establish whether individuals who have received exome-only testing represent an appropriate priority group for LRS referral.

Beyond prior testing depth, disease context was a major determinant of LRS yield. Genome-wide cohorts (Tier 1) showed lower rescue rates than targeted disease-specific studies (Tier 2), which preferentially enrolled patients with either suspected SRS-refractory variant classes or strong phenotype-to-gene correspondence. This supports the idea that when the causative variant class is known to be SRS-refractory, focused interrogation can be highly efficient. For example, ADTKD due to *MUC1* cytosine duplications in an SRS-refractory VNTR region showed near complete rescue [45]; as did Canavan disease due to SVA element insertions in *ASPA* [32, 40]—two examples of variant classes completely invisible to SRS. Tuberous sclerosis complex showed rescue in over half of cases, mostly through identification of somatic SV mosaicism [33]; and autosomal dominant polycystic kidney disease benefited from pseudogene resolution in *PKD1* [39, 42]. The largest negative finding was associated with neurofibromatosis type 1, where fewer than 4% of patients were rescued [43]; this suggests that amplicon-based targeted LRS alone may have limited utility for most SRS-negative NF1 patients. These disease-specific patterns support the use of first-line targeted LRS for ADTKD and Canavan disease, with consideration of TSC and ADPKD at an earlier stage.

The LRS rescue patterns shown above also raise a broader question: what variant classes and rescue mechanisms account for the additional diagnoses achieved by LRS? Across genome-wide studies, structural variant detection was the single most frequent rescue mechanism, accounting for nearly half of all rescues and showing consistency across evidence tiers: 49.1% in Tier 1, and 46.8% in Tier 2 (with SV detection also dominant in Tier 3 case reports/series, albeit publication-biased). Together, these patterns indicate that SV detection is a principal mechanism by which LRS rescue diagnoses missed by SRS in this referral population.

Other rescue mechanisms are deserving of individual discussion as well since they can have different implications for clinical application of LRS. Reanalysis was the second most common type of rescue mechanism to be identified, representing approximately 29% of all genome-wide rescues (16/55; Supplementary Table S17). This reflects the benefit of revisiting patient data and updated gene–disease knowledge rather than LRS-specific technical capability, a distinction critical for patient counseling. Reanalysis was concentrated among patients with prior GS. This would support the interpretation that the original SRS data were technically adequate but interpreted under earlier, less complete annotation at the time. Given that 27.3% of Tier 1 rescues were achievable through reanalysis alone (15/55; Supplementary Table S17), re-evaluation of existing SRS data before LRS referral may reduce unnecessary testing. RE, phasing, full-length sequencing, and methylation detection each accounted for smaller but clinically meaningful fractions of rescues. Additionally, in Tier 2, mosaic SV drove rescue exclusively in the TSC cohort. Because these variants arose at low allele fractions, long reads spanning the breakpoint detected them more directly than short reads, which must infer them from limited supporting signal. Detection nonetheless depends on analytic sensitivity as much as platform, and mosaic-aware short-read pipelines could in theory resolve comparable variants. In total, 70.9% of the rescues (39/55; Supplementary Table S17) required genuinely LRS-specific capabilities.

A finding that deserves particular attention is the persistent occurrence of ME-mediated SVs (ME insertions and Alu/Alu-mediated rearrangements) across evidence tiers, identified in 0.7% (6/923) of Tier 1 and 4.9% (20/412) of Tier 2 patients. For context, prior literature estimates the contribution of diagnostic ME insertions at ~0.03% of unselected clinical exomes [106–108] and that of Alu/Alu-mediated rearrangements at ~0.3% of new human genetic diseases [109], though the latter has not been re-estimated at the cohort level since. Taken at face value, our rates could be interpreted as exceeding both prior estimates, particularly in Tier 2; however, this comparison should not be over-interpreted, as our denominators consist of patients undiagnosed or partially diagnosed by SRS—a population intrinsically enriched for variant classes that short-read technology fails to detect—whereas the literature comparators derive from unselected clinical or population cohorts. The true population-level contribution of ME-mediated pathogenesis therefore remains uncertain and will require denominator-matched LRS cohorts to resolve. Our findings nonetheless suggest that this contribution has been underestimated, and that broader LRS adoption may revise upward the estimated burden of MEs in rare disease.

Beyond the molecular mechanisms through which LRS achieves rescue, the clinical phenotype itself appears to inform diagnostic success. The cohort's substantial nervous-system enrichment (half of patients carrying "Abnormality of the nervous system" and accounting for 78.2% of rescues) initially raised the concern that the apparent neurological signal reflected cohort composition rather than LRS-amenable biology. Sensitivity and within-neurology analyses argue against this: the broadest nervous-system categories lost significance after excluding phenotype-unannotated patients, while specific sub-phenotypes (spasticity, microcephaly, and intellectual disability) remained robustly elevated against a neurological baseline of 15.8%. By contrast, seizure and broad movement abnormalities did not differ from baseline. These findings suggest that the LRS benefit within neurology is concentrated among patients with specific phenotypes rather than being a generalized feature of neurological presentation, and they nominate these sub-phenotypes as candidates for phenotype-guided LRS referral. Further cohorts with broader phenotypic representation will be needed to characterize LRS yield for other organ-system categories.

There are also several limitations to this study. Our synthesis relied on individual patient data extracted from published reports, which could in principle introduce selective inclusion. In practice the risk appears limited: the included cohorts reported outcomes for essentially all sequenced patients, so the 538 Group A patients represent near-complete cohorts (483 not rescued and 55 rescued) rather than a subset enriched for positive findings, and the pooled estimate is therefore unlikely to be materially inflated by selective reporting of solved cases. The phenotype analyses, by contrast, are more vulnerable to differential reporting detail. Rescued cases were phenotyped more completely than not-rescued cases: HPO-codable clinical data were available for 54 of 55 rescued patients (98.2%) but only 328 of 483 not-rescued patients (67.9%). Because rescued patients contributed more granular terms, and hence membership in more phenotype categories, phenotype-rescue associations may be partly inflated by ascertainment of clinical detail rather than true biological enrichment. The same mechanism accounts for the apparent association between the number of affected organ systems and rescue (Supplementary Section S14). The pre-specified sensitivity analysis restricted to phenotype-annotated patients bounds this effect, reducing FDR-significant categories from 65 to 55 and lowering odds ratios by approximately 31%, with the two broadest nervous-system categories losing significance; residual differential-annotation bias nonetheless cannot be excluded. We therefore interpret the phenotype associations as exploratory and hypothesis-generating rather than as validated triage criteria.

Both the outcome and rescue classification system and the inferred HPO phenotypes were designed and conducted by the review authors and involve subjective judgment. Patient-level phenotype tests are susceptible to study-level confounding because phenotype composition is correlated with study of origin; our sensitivity analyses showed the core neurological phenotypes were robust while a minority of signals were cohort-driven. Given the small number of contributing studies (n = 11), we treat these associations as exploratory and hypothesis-generating. In addition, because the HPO categories are nested, each patient contributes to all ancestor categories of their annotated terms, so the category-level tests are not independent; Benjamini–Hochberg correction, which assumes independence or positive regression dependence, may therefore overstate the number of significant categories. A dependence-robust Benjamini-Yekutieli sensitivity analysis left the core findings unchanged, with nine of the ten reported terms remaining significant. Furthermore, the cohort's strong neurological enrichment further limits statistical power to detect rescue signals in non-neurological phenotypes, although it does enable a more granular dissection of which neurological sub-phenotypes drive the observed benefit. Because neurological presentations predominate in clinical sequencing referrals, our cohort is enriched for nervous-system phenotypes (50.6% of patients, 78.2% of rescues). This skews the data: the pooled rescue rate is weighted toward neurological cohorts, phenotype associations are measured

against a nervous-system-heavy baseline (so enrichment reflects specific phenotypes, not neurological status itself), and non-neurological systems are too sparsely represented to assess. Additionally, a single large cohort by Cohen et al. contributed a third of the primary meta-analysis denominator with uniquely exhaustive prior testing, exerting outsized influence on both the pooled rescue rate and the ES+GS subgroup estimate [12]. Egger's test suggested small-study effects (p = 0.043); these cannot be cleanly separated from study-design heterogeneity in this literature, and we cannot exclude a contribution from publication bias. Three of the eleven primary-analysis studies, and one study contributing only to the excluded-study synthesis (Ek et al.), were preprints that had not completed peer review. These unreviewed sources carry additional uncertainty, and their extracted patient-level data have not been subject to peer review; a sensitivity analysis excluding all preprints left the pooled estimate slightly higher (18.8% vs 16.4%), indicating that their inclusion did not inflate the pooled yield, but residual risk cannot be fully excluded. In addition, technical parameters including sequencing depth and coverage were incompletely reported across the studies, precluding formal depth–yield analysis. Tier 3 case reports/series data reflect extreme publication bias and are reported narratively only. We also restricted eligibility to English-language reports, excluding two Chinese-language studies on this basis. Given that only two studies were affected, and that 11 of the 37 Tier 1 and Tier 2 studies (29.7%) came from groups in China, Japan, and South Korea publishing in English-language journals, we expect any impact on the pooled estimates to be minimal, although residual bias from non-English reports not indexed in the searched databases cannot be excluded. Furthermore, short-read analysis pipelines also differed across the included studies and were frequently not reported in detail. We therefore cannot fully separate genuine LRS-specific rescues from variants that were present in the short-read data but missed by a suboptimal pipeline, although the reanalysis category captures part of this effect.

We believe that future studies should focus on collecting more detailed prospective data from patients undergoing both SRS and LRS testing, with standardized SRS and LRS protocols to reduce the heterogeneity that limits current meta-analytic precision. Formal mixed-effects phenotype–outcome models on larger patient datasets would enable more rigorous identification of phenotype-based triage criteria. The role of systematic periodic reanalysis of existing SRS data as a complement or alternative to LRS referral requires evaluation, given that a substantial minority of rescues were attributable to updated annotation rather than LRS-specific technology.

Conclusions

This systematic review indicates that LRS can provide a meaningful rescue for patients with rare diseases undiagnosed after SRS, primarily through structural variant detection. Re-evaluation of existing SRS data before LRS referral and consideration of patients directly after negative ES may help optimize diagnostic yield while reducing unnecessary testing. Specific clinical phenotypes may further inform prioritization but require prospective evaluation before they can guide referral.

Abbreviations

LRS	Long-read sequencing
SRS	Short-read sequencing
ES	Exome sequencing
GS	Genome sequencing
SV	Structural variant
RE	Repeat expansion
ME	Mobile element
STR	Short tandem repeat
VNTR	Variable number tandem repeat
PacBio	Pacific Biosciences
ONT	Oxford Nanopore Technologies
SMRT	Single-molecule real-time
ADTKD	Autosomal dominant tubulointerstitial kidney disease
ADPKD	Autosomal dominant polycystic kidney disease
TSC	Tuberous sclerosis complex
NF1	Neurofibromatosis type 1

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40246-026-01039-1.

Additional file 1.

Additional file 2.

Additional file 3.

Additional file 4.

Acknowledgments

We are grateful to R Gene Biosciences and Genoks for covering the article processing charge of this publication. Neither company had any role in the design of the study, the collection, analysis and interpretation of data, or the writing of the manuscript.

Author contributions

Conceptualization: A.B., M.Y.K., M.A.E.; Data Curation: A.B., M.F.M; Formal Analysis: A.B., M.A.E.; Investigation: A.B., M.F.M., Y.B.; Methodology: A.B., M.A.E., M.Y.K.; Project Administration: M.Y.K., A.B., M.A.E.; Software: A.B.; Validation: A.B., M.F.M.; Visualization: A.B.; Writing-original draft: A.B.; Writing-review & editing: A.B., M.F.M., M.A.E., Y.B., M.Y.K.; Supervision: M.Y.K., M.A.E.

Funding

This study was supported by the Gazi University Scientific Research Projects Coordination Unit under grant number TGA-2025-9923, within the scope of the Research Universities Support Program (ADEP) of the Council of Higher Education (YÖK), Türkiye.

Data availability

All data analyzed in this study were extracted from previously published articles cited in the References. The complete individual patient dataset (Dataset S1) and phenotype analysis dataset (Dataset S2) generated during this study are provided as Supplementary Materials. All analysis scripts used in this study are publicly available at Zenodo (https://doi.org/10.5281/zenodo.20127016).

Declarations

Ethics approval and consent to participate

This study is a systematic review and meta-analysis of previously published data and did not involve direct interaction with human participants or animal subjects. Therefore, ethical approval and informed consent were not required.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Medical Genetics, Faculty of Medicine, Gazi University, Ankara 06500, Türkiye

Received: 5 June 2026 / Accepted: 31 August 2026

References

1. 100,000 Genomes Project Pilot Investigators. 100,000 Genomes pilot on rare-disease diagnosis in health care-preliminary report. N Engl J Med. 2021;385(20):1868–80.
2. Chung CCY, Hue SPY, Ng NYT, Doong PHL, Chu ATW, Chung BHY. Meta-analysis of the diagnostic and clinical utility of exome and genome sequencing in pediatric and adult patients with rare diseases across diverse populations. Genet Med. 2023;25(9):100896.
3. Del Gobbo GF, Boycott KM. The additional diagnostic yield of long-read sequencing in undiagnosed rare diseases. Genome Res. 2025;35(4):559–71.
4. Hiatt SM, Lawlor JMJ, Handley LH, Latner DR, Bonnstetter ZT, Finnila CR, et al. Long-read genome sequencing and variant reanalysis increase diagnostic yield in neurodevelopmental disorders. Genome Res. 2024;34(11):1747–62.
5. Steyaert W, Sagath L, Demidov G, Yépez VA, Esteve-Codina A, Gagneur J, et al. Unraveling undiagnosed rare disease cases by HiFi long-read genome sequencing. Genome Res. 2025;35(4):755–68.
6. Oehler JB, Wright H, Stark Z, Mallett AJ, Schmitz U. The application of long-read sequencing in clinical settings. Hum Genom. 2023;17(1):73.
7. Iyer SV, Goodwin S, McCombie WR. Leveraging the power of long reads for targeted sequencing. Genome Res. 2024;34(11):1701–18.
8. Ibrahim M, Chebly A, El Shamieh S. Long-read sequencing in inherited retinal dystrophies: a systematic review. Ophthalmol Sci. 2026;6(3):101069.
9. Ibrahim AA, Fakhro KA, Abdallah AM. Diagnostic yield of long-read sequencing for rare diseases: a systematic review. Front Genet. 2026;17:1809097.
10. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36.
11. Shah J, Mondal D, Jain D, Mhatre P, Patel K, Iyer A, et al. Long read whole genome sequencing-based discovery of structural variants and their role in aetiology of non-syndromic autism spectrum disorder in India. BMC Med Genomics. 2025;18(1):131.
12. Cohen ASA, Farrow EG, Abdelmoity AT, Alaimo JT, Amudhavalli SM, Anderson JT, et al. Genomic answers for children: dynamic analyses of >1000 pediatric rare disease genomes. Genet Med. 2022;24(6):1336–48.
13. Pauper M, Kucuk E, Wenger AM, Chakraborty S, Baybayan P, Kwint M, et al. Long-read trio sequencing of individuals with unsolved intellectual disability. Eur J Hum Genet. 2021;29(4):637–48.
14. Chu Y, Zhang C, Pan Z, Peng J. Application of long-read sequencing in the diagnosis of Duchenne/Becker muscular dystrophy: unveiling complex structural variations and deep intronic mutations. Ital J Pediatr. 2025;51(1):204.
15. Daida K, Cogan G, Tesson C, Lesage S, Schaake S, Balck A, et al. Screening of hidden pathogenic structural variants in PRKN. Mov Disord. 2025;40(9):2009–11.
16. Daida K, Yoshino H, Malik L, Baker B, Ishiguro M, Genner R, et al. The utility of long-read sequencing in diagnosing early onset Parkinson's disease. Ann Neurol. 2025;97(4):753–65.
17. Negi S, Stenton SL, Berger SI, Canigiula P, McNulty B, Violich I, et al. Advancing long-read nanopore genome assembly and accurate variant calling for rare disease detection. Am J Hum Genet. 2025;112(2):428–49.
18. AlAbdi L, Shamseldin HE, Khouj E, Helaby R, Aljamal B, Alqahtani M, et al. Beyond the exome: utility of long-read whole genome sequencing in exome-negative autosomal recessive diseases. Genome Med. 2023;15(1):114.
19. Bruels CC, Littel HR, Daugherty AL, Stafki S, Estrella EA, McGaughy ES, et al. Diagnostic capabilities of nanopore long-read sequencing in muscular dystrophy. Ann Clin Transl Neurol. 2022;9(8):1302–9.
20. Sinha S, Rabea F, Ramaswamy S, Chekroun I, El Naofal M, Jain R, et al. Long read sequencing enhances pathogenic and novel variation discovery in patients with rare diseases. Nat Commun. 2025;16(1):2500.
21. Kamolvisit W, Cheawsamoot C, Chetruengchai W, Kor-Anantakul P, Thangpong R, Srichomthong C, et al. Singleton rapid long-read genome sequencing as first tier genetic test for critically ill children with suspected genetic diseases. Eur J Hum Genet. 2025;33(6):768–74.
22. Smits DJ, Ferraro F, Drost M, van der Linde HC, de Graaf BM, van Bever Y, et al. Nanopore long-read sequencing for the critically ill facilitates ultrarapid diagnostics and urgent clinical decision making. Eur J Hum Genet. 2026;34(1):108–18.
23. Sorrentino U, Pavlov M, Mirza-Schreiber N, Brugger M, Brunet T, Tsoma E, et al. Integrating long-read nanopore sequencing for precision resolution of genomic variants in dystonia. Mov Disord. 2026;41(1):70–83.
24. Hiatt SM, Lawlor JMJ, Handley LH, Ramaker RC, Rogers BB, Partridge EC, et al. Long-read genome sequencing for the molecular diagnosis of neurodevelopmental disorders. Hum Genet Genomics Adv. 2021;2(2):100023.
25. Fukuda H, Mizuguchi T, Doi H, Kameyama S, Kunii M, Joki H, et al. Long-read sequencing revealing intragenic deletions in exome-negative spastic paraplegias. J Hum Genet. 2023;68(10):689–97.
26. Sano Y, Koyanagi Y, Wong JH, Murakami Y, Fujiwara K, Endo M, et al. Likely pathogenic structural variants in genetically unsolved patients with retinitis pigmentosa revealed by long-read sequencing. J Med Genet. 2022;59(11):1133–8.
27. Jang SS, Kim S, Lee S, Kim SY, Moon J, Kim J, et al. Pangenome-based identification of cryptic pathogenic variants in undiagnosed rare disease patients. medRxiv. 2025. https://doi.org/10.1101/2025.07.08.25330875.
28. Lesurf R, Jain A, Hanafi N, Mitina A, Yin Y, Kolla VA, et al. Long-read genome sequencing increases genomic yield in congenital heart disease. medRxiv. 2025. https://doi.org/10.1101/2025.05.14.25327523.
29. Redfield SE, Shao W, Sun T, Pastolero A, Rowell WJ, French CE, et al. Long-read sequencing increases diagnostic yield for pediatric sensorineural hearing loss. medRxiv. 2024. https://doi.org/10.1101/2024.09.30.24314377.
30. Ek M, Kvarnung M, Ten Berk Boer E, La Fleur L, Ljöstad L, Lyander A, et al. Long-read genome sequencing enhances diagnostics of pediatric neurological disorders. Res Sq. 2025. https://doi.org/10.21203/rs.3.rs-6863124/v1.
31. Chaudhry A, Trinder M, Vesely K, Cermakova L, Jackson L, Wang J, et al. Genetic identification of homozygous familial hypercholesterolemia by long-read sequencing among patients with clinically diagnosed heterozygous familial hypercholesterolemia. Circ Genom Precis Med. 2023;16(2):e003887.
32. Dominguez Gonzalez CA, Bell KM, Rajagopalan R, de Silva MG, Lemes A, Zabala C, et al. Deep intronic SVA_E insertion identified as the most common pathogenic variant associated with Canavan disease: a diagnostic blind spot. Neurol Genet. 2025;11(5):e200291.
33. Duan J, Pan S, Ye Y, Hu Z, Chen L, Liang D, et al. Uncovering hidden genetic variations: long-read sequencing reveals new insights into tuberous sclerosis complex. Front Cell Dev Biol. 2024;12:1415258.
34. Miller DE, Lee L, Galey M, Kandhaya-Pillai R, Tischkowitz M, Amalnath D, et al. Targeted long-read sequencing identifies missing pathogenic variants in unsolved Werner syndrome cases. J Med Genet. 2022;59(11):1087–94.
35. Miller DE, Sulovari A, Wang T, Loucks H, Hoekzema K, Munson KM, et al. Targeted long-read sequencing identifies missing disease-causing variation. Am J Hum Genet. 2021;108(8):1436–49.
36. Nakamichi K, Huey J, Sangermano R, Place EM, Bujakowska KM, Marra M, et al. Targeted long-read sequencing enriches disease-relevant genomic regions of interest to provide complete Mendelian disease diagnostics. JCI Insight. 2024;9(20):e183902.
37. Rodilla C, Núñez-Moreno G, Benitez Y, Romero R, Fernández-Caballero L, Mínguez P, et al. Cas9-targeted-based long-read sequencing for genetic screening of RPE65 locus. Front Genet. 2024;15:1439153.
38. Rudaks LI, Stevanovski I, Yeow D, Reis ALM, Chintalaphani SR, Cheong PL, et al. Targeted long-read sequencing as a single assay improves the diagnosis of spastic-ataxia disorders. Ann Clin Transl Neurol. 2025;12(4):832–41.
39. Sun Q, Xu P, Mao A, Huang S, Li J, Chen L, et al. Targeted long-read sequencing enables higher diagnostic yield of ADPKD by accurate PKD1 genetic analysis. NPJ Genom Med. 2025;10(1):22.

40. Weiß M, Selig M, Friedrich J, Wierczeiko A, Diederich S, Sigel H, et al. Deep intronic SVA_E retrotransposition as a novel factor in Canavan disease pathogenesis. Hum Gene Ther. 2025;36(17–18):1248–56.
41. Nakamura W, Hirata M, Oda S, Chiba K, Okada A, Mateos RN, et al. Assessing the efficacy of target adaptive sampling long-read sequencing through hereditary cancer patient genomes. NPJ Genom Med. 2024;9(1):11.
42. Xu D, Mao A, Chen L, Wu L, Ma Y, Mei C. Comprehensive analysis of PKD1 and PKD2 by long-read sequencing in autosomal dominant polycystic kidney disease. Clin Chem. 2024;70(6):841–54.
43. Zheng Y, Chen M, Zhang S, Peng Y, Wu X, Guo D, et al. Long-read sequencing for NF1 gene analysis: enhancing diagnostic accuracy for neurofibromatosis type 1. Hum Mol Genet. 2025;34(19):1628–37.
44. Viering D, Hureaux M, Neveling K, Latta F, Kwint M, Blanchard A, et al. Long-read sequencing identifies novel pathogenic intronic variants in Gitelman syndrome. J Am Soc Nephrol. 2023;34(2):333–45.
45. Tanaka Y, Nagano C, Sakakibara N, Okada E, Aoyama S, Kimura Y, et al. Phenotype and genotype of autosomal dominant tubulointerstitial kidney disease in a Japanese cohort. Clin Exp Nephrol. 2025;29(6):788–96.
46. Baumann AA, Knol LI, Arlt M, Hutschenreiter T, Richter A, Widmann TJ, et al. Long-read genome and RNA sequencing resolve a pathogenic intronic germline LINE-1 insertion in APC. NPJ Genom Med. 2025;10(1):30.
47. Bestetti I, Crippa M, Sironi A, Bellini M, Tumiatti F, Ballabio S, et al. Long-read sequencing reveals chromothripsis in a molecularly unsolved case of Cornelia de Lange syndrome. Front Genet. 2024;15:1358334.
48. Bjørnstad PM, Aaløkken R, Åsheim J, Sundaram AYM, Felde CN, Østby GH, et al. A 39 kb structural variant causing Lynch syndrome detected by optical genome mapping and nanopore sequencing. Eur J Hum Genet. 2024;32(5):513–20.
49. Chen L, Luo X, Wang H, Tian Y, Liu Y. Identifying inversions with breakpoints in the Dystrophin gene through long-read sequencing: report of two cases. BMC Med Genomics. 2024;17(1):227.
50. Damián A, Núñez-Moreno G, Jubin C, Tamayo A, de Alba MR, Villaverde C, et al. Long-read genome sequencing identifies cryptic structural variants in congenital aniridia cases. Hum Genomics. 2023;17(1):45.
51. Dardas Z, Marafi D, Duan R, Fatih JM, El-Rashidy OF, Grochowski CM, et al. Genomic balancing act: deciphering DNA rearrangements in the complex chromosomal aberration involving 5p15.2, 2q31.1, and 18q21.32. Eur J Hum Genet. 2025;33(2):231–8.
52. Di Letto P, Budillon A, Rahman SI, Del Vecchio Blanco F, Zanobio M, Scarpato M, et al. Nanopore sequencing solves an elusive case of Sotos syndrome. Am J Med Genet A. 2025;197(7):e64039.
53. DuBois M, Dixon K, Sherlaw-Sturrock C, Shen Y, Probst F, Clarke L, et al. Identification of a non-coding causative variant underlying Warsaw breakage syndrome using long-read based genomic sequencing and transcriptome analysis. Am J Med Genet A. 2026;200(2):432–7.
54. Fabian-Morales GE, Ordoñez-Labastida V, Rowell WJ, Lambert C, Fanslow C, Robertson A, et al. Resolving the diagnostic odyssey in inherited retinal dystrophies through long-read genome sequencing. Am J Med Genet A. 2025;197(10):e64139.
55. Fant A, Trova S, Monfrini E, Treves G, Musacchia F, Landuzzi F, et al. Compound heterozygous structural variants in cases with unsolved PRKN-associated Parkinson's disease. Mov Disord. 2025;40(12):2722–31.
56. Fernández-Suárez E, González-Del Pozo M, Méndez-Vidal C, Martín-Sánchez M, Mena M, de la Morena-Barrio B, et al. Long-read sequencing improves the genetic diagnosis of retinitis pigmentosa by identifying an Alu retrotransposon insertion in the EYS gene. Mob DNA. 2024;15(1):9.
57. Feshbach LH, Similuk M, Amendola LM, Lewis KL, Walkiewicz MA, Tokita M, et al. Long-read sequencing as a diagnostic tool for primary ciliary dyskinesia. Case Rep Genet. 2025;2025:5109434.
58. Geng C, Zhang C, Li P, Tong Y, Zhu B, He J, et al. Identification and characterization of two DMD pedigrees with large inversion mutations based on a long-read sequencing pipeline. Eur J Hum Genet. 2023;31(5):504–11.
59. Gupta P, Nakamichi K, Bonnell AC, Yanagihara R, Radulovich N, Hisama FM, et al. Familial co-segregation and the emerging role of long-read sequencing to re-classify variants of uncertain significance in inherited retinal diseases. NPJ Genom Med. 2023;8(1):20.
60. Hort Y, Sullivan P, Wedd L, Fowles L, Stevanovski I, Deveson I, et al. Atypical splicing variants in PKD1 explain most undiagnosed typical familial ADPKD. NPJ Genom Med. 2023;8(1):16.
61. Jacobson AL, AbuRayyan A, Gulsuner S, Slater H, Anasiz Y, Ahmad SM, et al. Long-read DNA and RNA sequencing for inherited polyposis and colorectal cancer: cryptic intronic variants and multiple mutational mechanisms. J Med Genet. 2025;62(10):617–23.
62. Jiang Y, Wu L, Huang S, Li P, Gao B, Yuan Y, et al. Study of complex structural variations of X-linked deafness-2 based on single-molecule sequencing. Biosci Rep. 2021;41(6):1.
63. Jin B, Yoon JG, Kim A, Moon J, Kim HJ. Late-onset Ataxia-Telangiectasia presenting with dystonia and tremor: the use of nanopore long-read sequencing solving the variant phase. Neurol Genet. 2024;10(2):e200141.
64. Laurent S, Vannier A, Gehrig C, Abramowicz M, Paoloni-Giacobino A, Cao Van H, et al. Improved variant detection using long-read sequencing and optical mapping: illustration in STRC-related hearing loss. Eur J Med Genet. 2025;73:104986.
65. Lei M, Liang D, Yang Y, Mitsuhashi S, Katoh K, Miyake N, et al. Long-read DNA sequencing fully characterized chromothripsis in a patient with Langer-Giedion syndrome and Cornelia de Lange syndrome-4. J Hum Genet. 2020;65(8):667–74.
66. Liu C, Wang X, Xu C, Liu X, Ke L, Li Y, et al. Long-read sequencing identifies ATXN3 repeat expansions, and transcriptomics reveals disease progression biomarkers and druggable targets for spinocerebellar ataxia type 3. BMC Neurol. 2025;25(1):370.
67. Llansó L, Stevanovski I, Morís G, Collet-Vidiella R, Segarra-Casas A, González-Quereda L, et al. Repeat expansions in PLIN4 cause autosomal dominant vacuolar myopathy with sarcolemmal features. Ann Clin Transl Neurol. 2025;12(10):2136–51.
68. Lopriore P, Legati A, Neuhofer CM, Lo Gerfo A, Kopajtich R, Barresi M, et al. An inherited mtDNA rearrangement, mimicking a single large-scale deletion, associated with MIDD and a primary cardiological phenotype. Mitochondrion. 2025;83:102037.
69. Mao B, Cai X, Lin N, Jiang Y, Hao N, Dai Y, et al. Congenital muscular dystrophies and myopathies: the leading cause of genetic muscular disorders in eleven Chinese families. BMC Musculoskelet Disord. 2025;26(1):51.
70. Mariscal J, Kay CN. Atypical presentation of Oguchi disease with severe cystoid macular edema and compound heterozygous SAG pathogenic variants. Ophthalmic Genet. 2026;47(1):89–92.
71. Martínez-Granero F, Martínez-Cayuelas E, Rodilla C, Núñez-Moreno G, Rodríguez de Alba M, Blanco-Kelly F, et al. Biallelic intragenic tandem duplication of CPLANE1 in Joubert syndrome: a case report. Clin Genet. 2023;103(4):448–52.
72. Melas M, Kautto EA, Franklin SJ, Mori M, McBride KL, Mosher TM, et al. Long-read whole genome sequencing reveals HOXD13 alterations in synpolydactyly. Hum Mutat. 2022;43(2):189–99.
73. Merker JD, Wenger AM, Sneddon T, Grove M, Zappala Z, Fresard L, et al. Long-read genome sequencing identifies causal structural variation in a Mendelian disease. Genet Med. 2018;20(1):159–63.
74. Miao H, Zhou J, Yang Q, Liang F, Wang D, Ma N, et al. Long-read sequencing identified a causal structural variant in an exome-negative case and enabled preimplantation genetic diagnosis. Hereditas. 2018;155:32.
75. Mizuguchi T, Okamoto N, Yanagihara K, Miyatake S, Uchiyama Y, Tsuchida N, et al. Pathogenic 12-kb copy-neutral inversion in syndromic intellectual disability identified by high-fidelity long-read sequencing. Genomics. 2021;113(1 Pt 2):1044–53.
76. Mizuguchi T, Suzuki T, Abe C, Umemura A, Tokunaga K, Kawai Y, et al. A 12-kb structural variation in progressive myoclonic epilepsy was newly identified by long-read whole-genome sequencing. J Hum Genet. 2019;64(5):359–68.
77. Mori T, Sakamoto M, Tayama T, Goji A, Toda Y, Fujita A, et al. A case of epilepsy with myoclonic atonic seizures caused by *SLC6A1* gene mutation due to balanced chromosomal translocation. Brain Dev. 2023;45(7):395–400.
78. Moriura Y, Nishio Y, Ichimura S, Noda H, Tanahashi Y, Yamamoto H, et al. De novo CDKN1C variant in Beckwith-Wiedermann spectrum with atypical complications. Hum Genome Var. 2025;12(1):9.
79. Park J, Dufke C, Fleszar Z, Schlotterbek M, Buena-Atienza E, Stühn LG, et al. Long-read sequencing identifies mosaic sequence variations in Friedreich's ataxia-GAA repeats. Int J Mol Sci. 2025;26(11):4969.
80. Perera BLA, Stewart R, Furuta Y, Ezell KM, Rives L, Nunley B, et al. First clinical diagnosis of FAME3 via commercial Long-Read sequencing reveals mosaic repeat expansion in MARCHF6 gene. Neurogenetics. 2025;26(1):61.
81. Qiu X, Jin X, Li J, Zhang Y, Ma J, Zheng F. Long-Read sequencing identified a PKD1 gene conversion in ADPKD rather than the false-positive exon deletion indicated by WES and MLPA. Hum Mutat. 2024. https://doi.org/10.1155/2024/7225526.

82. Reiner J, Pisani L, Qiao W, Singh R, Yang Y, Shi L, et al. Cytogenomic identification and long-read single molecule real-time (SMRT) sequencing of a Bardet-Biedl Syndrome 9 (BBS9) deletion. NPJ Genom Med. 2018;3:3.
83. Sakamoto M, Kurosawa K, Tanoue K, Iwama K, Ishida F, Watanabe Y, et al. A heterozygous germline deletion within USP8 causes severe neurodevelopmental delay with multiorgan abnormalities. J Hum Genet. 2024;69(2):85–90.
84. Shangguan S, Cui X, Li J, Li N, Liu R, Chen X. Alu-mediated FANCD2 exonic deletion contributes to Fanconi anaemia. Br J Haematol. 2025;207(3):1076–84.
85. Sjursen W, Svaasand EK, Gilde B, Ravi A, Korseth KM, Singh AK, et al. Lynch syndrome caused by SINE-VNTR-Alu-F retrotransposon insert in MSH6 confirmed after 20 years of testing: a case report and literature review. Hered Cancer Clin Pract. 2025;23(1):22.
86. Sneddon TP, Melville SA, Xiong M, Foss K, Cardwell EB, Rafferty K, et al. Long-read sequencing of a neurodevelopmental disorder patient reveals complex rearrangement involving the ARID1B gene. Am J Med Genet A. 2026;200(2):426–31.
87. Sonoda K, Ishihara H, Sakazaki H, Suzuki T, Horie M, Ohno S. Long-read sequence confirmed a large deletion including MYH6 and MYH7 in an infant of atrial septal defect and atrial arrhythmias. Circ Genom Precis Med. 2021;14(4):e003223.
88. Steigerwald C, Borsuk J, Pappas J, Galey M, Scott A, Devaney JM, et al. CLN2 disease resulting from a novel homozygous deep intronic splice variant in TPP1 discovered using long-read sequencing. Mol Genet Metab. 2023;140(4):107713.
89. Thewamit R, Khongkhatithum C, Thampratankul L, Kamolvisit W, Khongkrapan A, Wattanasirichaigoon D. Case report: Severe nonketotic hyperglycinemia in a neonate without apparent seizures but concomitant cleft palate and cerebral sinovenous thrombosis. Front Pediatr. 2023;11:1155035.
90. Wahlster L, Verboon JM, Ludwig LS, Black SC, Luo W, Garg K, et al. Familial thrombocytopenia due to a complex structural variant resulting in a WAC-ANKRD26 fusion transcript. J Exp Med. 2021;218(6):e20210444.
91. Wang GQ, Rui HL, Dong HR, Sun LJ, Yang M, Wang YY, et al. SMRT sequencing revealed to be an effective method for ADTKD-MUC1 diagnosis through follow-up analysis of a Chinese family. Sci Rep. 2020;10(1):8616.
92. Watson CM, Dean P, Camm N, Bates J, Carr IM, Gardiner CA, et al. Long-read nanopore sequencing resolves a TMEM231 gene conversion event causing Meckel-Gruber syndrome. Hum Mutat. 2020;41(2):525–31.
93. Wu X, Xu Q, Chen G, Huang J, Zhong Y, Tian L, et al. Identification of a cryptic unbalanced translocation Der(22)t(12;22)(q24.33;q13.33) in a large Chinese family with Phelan-McDermid syndrome by nanopore sequencing. Sci Rep. 2025;15(1):2656.
94. Xie Z, Liu C, Lu Y, Sun C, Liu Y, Yu M, et al. Exonization of a deep intronic long interspersed nuclear element in Becker muscular dystrophy. Front Genet. 2022;13:979732.
95. Xie Z, Sun C, Liu Y, Yu M, Zheng Y, Meng L, et al. Practical approach to the genetic diagnosis of unsolved dystrophinopathies: a stepwise strategy in the genomic era. J Med Genet. 2021;58(11):743–51.
96. Xie Z, Sun C, Zhang S, Liu Y, Yu M, Zheng Y, et al. Long-read whole-genome sequencing for the genetic diagnosis of dystrophinopathies. Ann Clin Transl Neurol. 2020;7(10):2041–6.
97. Yamada M, Watanabe D, Miya F, Shiramizu H, Inoue M, Kabata H, et al. Patients with hereditary hemorrhagic telangiectasia diagnosed by nanopore long read sequencer. Congenit Anom (Kyoto). 2025;65(1):e70019.
98. Yanagi K, Coker J, Miyana K, Aso S, Kobayashi N, Satou K, et al. Biallelic CC2D2A variants, SNV and LINE-1 insertion simultaneously identified in siblings using long-read whole-genome sequencing and haplotype phasing. J Hum Genet. 2023;68(6):431–5.
99. Yoon JG, Lee S, Park S, Jang SS, Cho J, Kim MJ, et al. Identification of a novel non-coding deletion in Allan-Herndon-Dudley syndrome by long-read HiFi genome sequencing. BMC Med Genomics. 2025;18(1):41.
100. Zamora-Cánovas A, de la Morena-Barrio B, Marín-Quilez A, Sierra-Aisa C, Male C, Fernández-Mosteirin N, et al. Targeted long-read sequencing identifies and characterizes structural variants in cases of inherited platelet disorders. J Thromb Haemost. 2024;22(3):851–9.
101. Zeng S, Zhang MY, Wang XJ, Hu ZM, Li JC, Li N, et al. Long-read sequencing identified intronic repeat expansions in *SAMD12* from Chinese pedigrees affected with familial cortical myoclonic tremor with epilepsy. J Med Genet. 2019;56(4):265–70.
102. Zhao A, Shen J, Ding Y, Sheng M, Zuo M, Lv H, et al. Long-read sequencing identified a novel nonsense and a de novo missense of *PPA2* in trans in a Chinese patient with autosomal recessive infantile sudden cardiac failure. Clin Chim Acta. 2021;519:163–71.
103. Zhao L, Hu C, Pan S, Wang D, Wang Y, Li X. Two novel deep intronic variants cause Duchenne muscular dystrophy by splice-altering mechanism. Neuromuscul Disord. 2024;45:104470.
104. Martin AB. Using long-read sequencing to diagnose rare neuromuscular disorders: a pilot study [Undergraduate honors thesis]. Winnipeg: University of Manitoba; 2025. http://hdl.handle.net/1993/39077. Accessed 15 April 2026.
105. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–34.
106. Torene RI, Galens K, Liu S, Arvai K, Borroto C, Scuffins J, et al. Mobile element insertion detection in 89,874 clinical exomes. Genet Med. 2020;22(5):974–8.
107. Wijngaard R, Demidov G, O'Gorman L, Corominas-Galbany J, Yaldiz B, Steyaert W, et al. Mobile element insertions in rare diseases: a comparative benchmark and reanalysis of 60,000 exome samples. Eur J Hum Genet. 2024;32(2):200–8.
108. Gardner EJ, Prigmore E, Gallone G, Danecek P, Samocha KE, Handsaker J, et al. Contribution of retrotransposition to developmental disorders. Nat Commun. 2019;10(1):4630.
109. Deininger PL, Batzer MA. Alu repeats and human disease. Mol Genet Metab. 1999;67(3):183–93.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.