Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Real-World Single-Center Study of Primary Immunodeficiency in Chinese Children: Diagnostic Efficacy and Genomic Landscape Revealed by Whole-Exome Sequencing Combined with Copy Number Variant Analysis

Qiang Zhang
Guangxi Maternal and Child Health Hospital
Qi Yang
Guangxi Maternal and Child Health Hospital
Sheng Yi
Guangxi Maternal and Child Health Hospital
Zailong Qin
Guangxi Maternal and Child Health Hospital
Jihong Zhang
Sheng Jing Hospital of China Medical University
Jingsi Luo
1270319085@qq.com

Guangxi Maternal and Child Health Hospital

Research Article

Keywords: Inborn errors of immunity, whole-exome sequencing, Pediatric, gene, primary immunodeficiency diseases

Posted Date: June 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9323897/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Clinical Immunology on October 10th, 2026. See the published version at https://doi.org/10.1007/s10875-026-02080-4.

Abstract

Background

Inborn Errors of Immunity (IEI) are genetically heterogeneous diseases with nonspecific clinical presentations, often resulting in delayed diagnosis. While whole-exome sequencing (WES) is widely used for diagnosing IEI, its real-world performance in the pediatric Chinese population remains understudied.

Methods

This study retrospectively analyzed 243 children aged 0 to 15 years suspected of having IEI at a single center. All underwent WES combined with copy-number variant (CNV) detection. Candidate variants were initially screened using the TGex platform and validated via Sanger sequencing. Variants were classified based on the ACMG/AMP 2015 guidelines, with only those listed as pathogenic or likely pathogenic in the IUIS 2024 IEI gene list considered diagnostic.

Results

In this region with a concentrated population of the Zhuang ethnic minority, a total of 77 children (31.69%) clinically suspected of having IEI received molecular diagnoses. 34 genes harbored pathogenic variants, with DiGeorge syndrome (22q11.2 deletion) being the most frequent diagnosis. Trio-based WES (proband plus both parents) significantly increased the diagnostic yield compared to proband-only sequencing (60.0% vs. 29.8%, $p < 0.001$). CNVs accounted for 27.27%(21 cases)of all diagnoses, emphasizing the importance of integrating CNV analysis into routine WES workflows. IEI phenotypes evolved with age: infants under six months primarily exhibited developmental immunodeficiencies, while older children more frequently presented with immune dysregulation and cancer-predisposing syndromes. Common manifestations included recurrent respiratory infections, cytopenia, and noninfectious features like congenital heart disease and neurodevelopmental delays, supporting IEI as a systemic disorder. Interestingly, 67.5% of patients displayed only 1–3 symptoms, yet even those with a single symptom achieved a diagnostic rate of 45%. Notably, the positivity rates among different bioinformatics analysts varied up to fivefold, underscoring the critical role of accurate variant interpretation in diagnostic outcomes.

Conclusions

IEI manifests as a dynamic, age-stratified spectrum, necessitating tailored diagnostic strategies for various life stages. For high-risk children, integrating WES, CNV analysis, and trio sequencing should be prioritized as initial testing. Standardizing variant interpretation is essential for equitable and

reproducible diagnosis. These findings offer a scalable, phenotype-driven framework for genomic diagnostics in resource-limited, ethnically diverse settings.

INTRODUCTION

IEI, formerly termed primary immunodeficiency diseases (PIDs), encompass over 508 distinct genetic disorders affecting innate and adaptive immunity, as classified by the International Union of Immunological Societies (IUIS) in 2024 [1, 2]. IEI is estimated to affect approximately 1% of the global population [3], yet it remains profoundly underdiagnosed—particularly in resource-limited settings such as China. This diagnostic gap is especially pronounced among Chinese children: national epidemiological data suggest an overall IEI incidence of ~ 1 in 5,000 live births. Given China's annual birth cohort of 16–18 million infants, this implies 3,000–4,000 incident IEI cases each year [4]. However, the proportion of clinically recognized cases remains as low as ~ 5% of this estimated burden [5], highlighting a substantial diagnostic gap where the vast majority of affected children remain unidentified. Among those diagnosed, nearly half lack a definitive molecular or genetic confirmation—particularly in primary care and resource-limited settings [6]. This massive underdiagnosis stems from the high phenotypic heterogeneity and complex genetic etiology of IEI, which often present with atypical or nonspecific manifestations, leading to frequent misdiagnosis and delayed identification. Compounding these challenges are limited access to specialized immunological testing and insufficient awareness among frontline clinicians. Consequently, delayed diagnosis exposes patients to recurrent life-threatening infections, autoimmune complications, and malignancies, with mortality rates reaching 10–12.7% in undiagnosed cohorts [7]. Moreover, affected children often endure diagnostic delays of several months to years, incurring annual pre-diagnosis medical costs exceeding ¥100,000 (≈ US$14,000) [8]. Chinese IEI patients thus face a triple burden: low diagnostic rates, poor access to targeted therapies, and substantial financial strain—making early, precise diagnosis and intervention critically urgent [5].

WES has emerged as a transformative tool for diagnosing genetically heterogeneous conditions. By capturing ~ 1% of the genome that harbors > 85% of known disease-causing variants, WES enables simultaneous interrogation of hundreds of IEI genes at a fraction of the cost of sequential single-gene testing [9, 10]. While WES significantly improves molecular confirmation for referred patients, it is important to acknowledge that it operates within the existing healthcare framework and does not automatically resolve upstream barriers such as initial clinical recognition or referral access. However, real-world diagnostic yields, optimal patient selection criteria, and the clinical utility of WES in diverse ethnic populations—particularly in Asia—remain inadequately characterized, especially regarding its performance in early-onset pediatric IEI.

Here, we report a retrospective, single-center study of 243 children with suspected IEI who underwent WES at Guangxi Maternal and Child Health Hospital between 2020 and 2025. Integrating deep phenotyping, high-throughput Sequencing, and CNV bioinformatics, we delineate the genomic landscape, age-dependent phenotypic trajectories, and clinical impact of WES in a representative Chinese pediatric

cohort. Our results provide actionable insights for early diagnosis, precision management, and optimized diagnostic strategies for IEI in resource-limited settings.

MATERIALS AND METHODS

Study design and participants

This retrospective, single-center study was conducted at Guangxi Maternal and Child Health Hospital (Nanning, China) between January 2020 and September 2025. We enrolled 243 pediatric patients (aged 10 days to 15 years) who were clinically suspected of having IEI based on established “red flag” criteria. Inclusion required at least one of the following: (1) recurrent or severe infections (≥ 8 episodes of respiratory infection per year or ≥ 2 infections requiring intravenous antibiotics), particularly those caused by opportunistic pathogens (e.g., *Pneumocystis jirovecii*, disseminated BCG infection, or persistent cytomegalovirus [CMV] infection); (2) atypical infection patterns (e.g., necrotizing pneumonia, deep abscesses, or hemophagocytic lymphohistiocytosis); (3) multi-system involvement (e.g., cytopenia, chronic diarrhea, hepatosplenomegaly, or mucocutaneous lesions); or (4) laboratory evidence of immune dysregulation (e.g., hypogammaglobulinemia, lymphocyte subset imbalance, or persistently elevated inflammatory markers without identifiable infection). All patients underwent preliminary immunological screening (including immunoglobulin quantification, lymphocyte subtyping, and complement assays) prior to genetic testing. The study was approved by the hospital’s Ethics Committee and conducted in accordance with the Declaration of Helsinki, with written informed consent obtained from all guardians.

Whole-exome sequencing and bioinformatic analysis

Genomic DNA was extracted from peripheral blood leukocytes. Whole-exome capture was performed using the Agilent SureSelect Clinical Research Exome V2 Kit, followed by sequencing on the Illumina HiSeq 2500 platform (paired-end 150-bp reads). Sequencing quality was rigorously monitored. Across the cohort, the mean sequencing depth was 254.8× (median: 251.3×; range: 143.7–369.2×). Target region coverage was > 99.3% at ≥ 10× (mean: 99.45% ± 0.07%), > 99.2% at ≥ 20× (mean: 99.38% ± 0.08%), and > 99.2% at ≥ 30× (mean: 99.33% ± 0.07%). The average on-target capture efficiency was 91.82% ± 0.28% (range: 90.95–92.44%). All samples passed predefined quality thresholds (mean depth ≥ 100×, ≥ 95% targets at ≥ 10×) and were included for downstream analysis. Raw reads were aligned to the human reference genome (GRCh37/hg19; this version was selected to maintain consistency with the capture kit's probe design and annotation BED files provided by the manufacturer) using BWA-MEM (v0.7.17). Variant calling was conducted with the Genome Analysis Toolkit (GATK v4.1), and annotation was performed using TGex (LifeMap Sciences). In addition, this study employed a coverage depth-based CNV detection method to identify CNVs from WES data. First, raw reads were processed using fastp (version 0.19.4) to remove adapter sequences, trim low-quality bases, and filter out poor-quality reads. Subsequently, the cleaned reads were aligned to the human reference genome (GRCh37/hg19) using BWA-MEM (v0.7.17), and the resulting BAM files were sorted and marked for PCR duplicates with Picard

Tools (v2.27.5). For each sample, read depth per target (exon) region was calculated using GATK DepthOfCoverage (v4.4.0.0). To account for differences in sequencing yield across samples, the raw read depth was normalized by the sample's average target coverage. Within each sequencing batch, a reference depth profile was established by computing the mean and standard deviation of normalized coverage across all samples for each target region. Candidate CNVs were identified by detecting genomic segments spanning at least three consecutive exons whose normalized coverage significantly deviated from the batch-specific baseline (defined as |Z-score| > 3, corresponding to a statistical significance of approximately $p < 0.003$). Adjacent aberrant segments were merged, and final calls were annotated as deletions or duplications based on the direction of depth change. This approach enables sensitive and specific detection of clinically relevant CNVs, including recurrent 22q11.2 deletions validated in our cohort. For candidate pathogenic CNVs detected in proband-only WES, we performed trio validation using the Illumina HumanCytoSNP-12 BeadChip array on the proband and both parents to clarify their inheritance patterns. Variants were initially classified according to the ACMG/AMP 2015 guidelines using InterVar to generate preliminary classifications. Subsequently, clinical geneticists performed expert manual review, integrating phenotype–genotype correlations, consistency with expected inheritance patterns, and the latest literature evidence to assign final classifications[11]. Only pathogenic or likely pathogenic variants in IEI-associated genes (per the IUIS 2024 classification) were considered diagnostic. All candidate pathogenic variants identified by TGex underwent validation via Sanger sequencing to confirm authenticity. Regardless of whether trio-WES was initially performed, parental samples were collected to determine the inheritance status of these variants. Primers for validation were designed using Oligo7 (v7.60, Molecular Biology Insights) and synthesized by Sangon Biotech (Shanghai, China). PCR amplification was performed under standard conditions (initial denaturation at 95°C for 5 min; 35 cycles of 95°C for 30 s, 58°C for 30 s, and 72°C for 45 s; final extension at 72°C for 10 min), followed by capillary sequencing on an ABI 3730xl Genetic Analyzer (Thermo Fisher Scientific).

To exclude sample mix-ups or non-parental relationships and meet the ACMG PS2 criterion (*De novo* (both maternity and paternity confirmed) in a patient with the disease and no family history), we performed short tandem repeat (STR) kinship analysis on all putative *de novo* variants identified by Sanger sequencing. Using the SifaSTR™ 23 plex kit, 23 core STR loci were amplified and analyzed by capillary electrophoresis. Allele transmission patterns between the proband and parents were compared, and the cumulative paternity index (CPI) was calculated to confirm biological parentage.This procedure ensures that all *de novo* variants reported in this study are established on reliable familial genetic backgrounds, effectively avoiding false-positive determinations due to sample mismatches or non-parental relationships.

The median turnaround time from sample receipt to clinical report was 32 days.

Statistical Analysis

Comparisons of categorical variables were performed using Fisher's exact test or the chi-square test. For key categorical outcomes (e.g., inheritance patterns, variant types), diagnostic yields and proportions are reported with 95% confidence intervals (CIs), calculated using the Wilson score method to ensure accuracy, particularly with small sample sizes or proportions near boundaries.

Age was categorized into four groups based on key biological milestones in immune system development: 0–6 months, 7 months–3 years, 4–8 years, and 9–15 years. Differences in distribution across age groups were assessed using the chi-square goodness-of-fit test. Pairwise comparisons between age groups were conducted using Fisher's exact test, with p-values adjusted by the Bonferroni correction to control the family-wise error rate.

Missing data were handled using complete case exclusion if the missing rate for any key variable (age, sex, core clinical phenotypes, sequencing quality metrics) exceeded 10%; otherwise, pairwise deletion was applied. In unadjusted comparisons, a two-sided $p < 0.05$ was considered statistically significant. All statistical analyses were performed using IBM SPSS Statistics, Version 19.0 (IBM Corp., Armonk, NY, USA).

RESULTS

Diagnostic Yield and Genetic Architecture of Pediatric IEI

In a cohort of 243 children clinically suspected of IEI, WES yielded a molecular diagnosis in 77 cases (31.69%, 95% CI: 26.2–37.8%) (Fig. 1A), identifying 88 variants across 34 IEI-associated genes. Among cases with a defined mode of inheritance, autosomal disorders predominated (57 of 77 confirmed cases, 74.03%, 95% CI: 63.3–82.5%), while X-linked diseases accounted for 20 cases (25.97%, 95% CI: 17.5–36.7%) (Fig. 1B). The cohort exhibited a marked male predominance, with 53 male patients (68.83%; male-to-female ratio of 2.21:1), consistent with the expected sex distribution of X-linked inborn errors of immunity. This finding underscores the importance of maintaining a high index of suspicion for X-linked immunodeficiencies in male children during clinical evaluation. Among variant types, sequence variation was most prevalent (73.86%, 65 variants, 95% CI: 63.8–81.9%), with missense comprising 43.18% (38 variants, 95% CI: 33.3–53.6%) of all variants. Critically, structural variation contributed to 23.86% (21 variants, 95% CI: 16.2–33.7%) of diagnoses (Fig. 1C), including 18 cases of 22q11.2 deletion causing DiGeorge syndrome, highlighting the necessity of integrating robust CNV detection into routine WES pipelines to avoid missing nearly one-quarter of pathogenic lesions. The majority of patients were hospitalized (88.89%, 216 cases, 95% CI: 84.3–92.3%) and, largely due to cost constraints, underwent proband-only sequencing (93.83%, 228 cases, 95% CI: 90.1–96.2%). In stark contrast, although only 15 families (6.17%, 95% CI: 3.8–9.9%) underwent trio-based WES (proband plus both parents) (Fig. 1A), its diagnostic yield was significantly higher than that of singleton WES (60.0%, 95% CI: 35.8–80.2% vs. 29.8%, 95% CI: 24.3–36.1%; $p < 0.001$). This advantage is attributable primarily to the direct and efficient identification of *de novo* variants (such as pathogenic variants in genes like KRAS and RAD51) by trio-WES, without the need for additional verification steps. In comparison, for singleton WES, the

pathogenicity assessment of some variants relied on subsequent Sanger sequencing of parental samples and STR kinship analysis. For instance, in this study, the KRAS c.38G > A / p. (Gly13Asp) variant was initially classified as a "variant of uncertain significance" (VUS); only after confirmation as a *de novo* variant was it upgraded to "pathogenic" (PS2 evidence) according to the ACMG/AMP guidelines.

Further analysis revealed that the 77 confirmed cases harbored variants in 34 distinct genes, spanning major categories of the IUIS 2024 classification (Tabel 1). The spectrum included severe combined immunodeficiencies (IL2RG, JAK3, RAG1), predominantly antibody deficiencies (BTK, CD19), phagocyte defects (CYBB, ELANE), autoinflammatory disorders (NLRC4, TNFAIP3, MVK), immune dysregulation syndromes (FOXP3, WAS, PIK3CD/PIK3R1, FAS), bone marrow failure or cancer-predisposing conditions (GATA2, RUNX1, SBDS) and Combined immunodeficiencies with associated or syndromic features (ATM, PMS2). Notably, DiGeorge syndrome due to 22q11.2 deletion emerged as the single most frequent diagnosis (18 cases, 23.38%). A striking proportion of cases (42 cases, 54.55%, 95% CI: 43.5–65.2%) carried *de novo* variants—particularly in KRAS (6 out of 6 cases), NRAS (3/3), BCL11B (2/2)—emphasizing that IEI often presents as a sporadic disorder without family history. Of the 88 variants, 66 were previously reported and 22 represented novel findings, expanding the known mutational landscape of IEI. Collectively, these results demonstrate that pediatric IEI in this population is characterized by extensive genetic and phenotypic heterogeneity, with a substantial burden of *de novo* and structural variants, necessitating comprehensive genomic analysis and a high index of clinical suspicion even in the absence of familial recurrence.

Age-Stratified Clinical Evolution

In this study, we integrated multidimensional analyses of patient age distribution, molecular diagnoses, and disease-type trajectories to systematically delineate, the dynamic age-dependent evolution of IEI in children. The age stratification into four groups—0–6 months, 7 months–3 years, 4–8 years, and 9–15 years—was based on key biological milestones in human immune system development and the typical clinical presentation windows of the relevant disorders. As shown in Fig. 2, the majority of genetically confirmed cases (38 of 77 cases, 49.35%, 95% CI: 38.5–60.3%) clustered within the first six months of life—a peak that aligns precisely with the waning of maternal antibodies and the critical window of immune system maturation—highlighting the neonatal period as a pivotal opportunity for early detection of severe developmental immunodeficiencies such as severe combined immunodeficiency (SCID) and DiGeorge syndrome. This temporal–phenotypic relationship is vividly illustrated in the Sankey diagram (Fig. 3): infants aged 0–6 months were almost exclusively diagnosed with classical developmental disorders, including T-B+ SCID and phagocyte defects; between 7 months and 3 years, the disease spectrum began to shift toward antibody deficiencies and autoinflammatory conditions; and by adolescence (9–15 years), cases diverged markedly into complex syndromic IEI characterized by cancer predisposition and chronic inflammation, such as activated PI3Kδ syndrome (APDS), GATA2 deficiency, and DNA repair disorders. The heatmap in Fig. 4 quantifies this transition, revealing that T-B+ SCID (11 cases) and DiGeorge syndrome (4 cases) dominated the 0–6-month group, whereas the 9–15-year group—though smaller in absolute numbers—exhibited the greatest diversity and complexity of

diagnoses, including multiple gain-of-function–driven immune dysregulation syndromes. The overall distribution of confirmed IEI cases across the four age groups was non-uniform ($\chi^2 = 32.25$, df = 3, $p < 0.0001$), supporting an age-stratified phenotypic evolution. Pairwise comparisons using Fisher's exact test revealed nominal differences between the 9–15-year cohort and younger groups ($p = 0.038–0.043$), although these did not remain significant after Bonferroni correction for multiple testing (adjusted $\alpha = 0.0083$). No differences were observed among patients aged 0–8 years (all $p > 0.05$; Table 2). These findings, interpreted alongside the strong overall test and consistent clinical trends, suggest a progressive shift in IEI phenotypes with advancing age.

Phenotypic Spectrum and Clinical Red Flags

Our analysis of 243 highly suspicious pediatric cases reveals that while the clinical manifestations of IEI are profoundly heterogeneous, they converge on a core set of “red flag” symptoms that serve as critical diagnostic anchors (Fig. 5). Recurrent or severe respiratory infections emerged as the paramount clinical indicator (76.54%, 186 of 243 suspected cases), establishing it as the single most predictive feature for IEI diagnosis. This is followed by laboratory-based hallmarks: cytopenia (26.34%, 64 of 243 suspected cases) and immunoglobulin abnormalities (20.58%, 50 of 243 suspected cases). These findings underscore the necessity for clinicians to maintain a high index of suspicion for IEI when encountering children presenting with these cardinal signs.

Crucially, our data demonstrate that IEI manifests as systemic disorders extending far beyond classical infectious phenotypes. A significant proportion of patients exhibited non-infectious features, including growth retardation (18.11%), congenital heart disease (17.28%), gastrointestinal symptoms (13.17%), and neurodevelopmental disorders (8.64%). The high prevalence of these diverse systemic manifestations highlights the broad pathophysiological impact of IEI and necessitates a comprehensive, multi-system evaluation in suspected cases.

A striking finding emerged from the analysis of symptom burden. Contrary to the intuitive assumption that diagnostic yield increases with the number of concurrent symptoms, our analysis revealed an inverted U-shaped relationship (Fig. 6). The diagnostic positivity peaked at 46.15% for patients presenting with five distinct symptoms, after which it declined sharply among those with more complex, multi-system presentations. Notably, even patients exhibiting only a single, highly suggestive symptom demonstrated a substantial diagnostic yield of 45.0%. This suggests that a focused, rather than exhaustive, clinical signature can be sufficient to justify genetic investigation. This counterintuitive “less is more” approach offers a new perspective on conventional wisdom and provides a practical framework for clinicians: early, targeted WES testing should be prioritized not only for children with extensive multisystem involvement, but critically, for those exhibiting a single, highly specific alarm sign, thereby enabling earlier diagnosis and intervention in this vulnerable population.

Variability in Variant Interpretation and Quality Control Measures

In the diagnosis of IEI based on WES, the bioinformatics interpretation pipeline plays an important role in determining the final diagnostic efficacy. To systematically assess factors contributing to variability in diagnostic outcomes, this study examined interpretation patterns across the top five clinicians by referral volume and the top five bioinformatics analysts by case load.

The results revealed that although diagnostic rates for cases referred by the top five clinicians showed variation (0.0%–50.0%), this difference did not reach statistical significance ($\chi^2 = 8.35$, df = 4, $p = 0.079$), likely due to limited case numbers per clinician (≤ 8 cases) and heterogeneous case complexity. In contrast, the top five bioinformatics analysts collectively interpreted 73.7% (179/242 cases) of the cohort's cases. Their individual diagnostic rates exhibited notable variation (11.54%–53.33%), and this difference was statistically significant ($\chi^2 = 22.87$, df = 4, $p = 0.00013$; see Supplementary Table S1 for detailed data).

This finding indicates that operational differences in variant interpretation, rather than clinical referral patterns, represent a significant contributor to variability in diagnostic yield. Although we acknowledge limitations inherent to this analysis, such as non-randomized case assignment and potential confounding by case complexity, the substantial magnitude of variability suggests that subjective factors in variant classification—including differential application of ACMG/AMP criteria—can lead to inconsistent outcomes even within the same institutional framework and technological platform. Consequently, establishing standardized and reproducible interpretation pipelines is paramount. This insight has shifted our quality improvement focus from the vague notion of clinical judgment to the optimizable interpretation pipeline, driving measures to minimize human bias and enhance diagnostic consistency.

DISCUSSION

Diagnostic Yield and Methodological Advantages

This study presents a systematic analysis of the clinical manifestations and genetic variant spectrum in a pediatric cohort with IEI from Guangxi, a Zhuang ethnic minority-concentrated region in southern China. The integrated WES and bioinformatics-based CNV analysis pipeline employed in our research achieved an overall diagnostic yield of 31.69% in this suspected pediatric IEI cohort. This result is consistent with international reports (25%–40%) [12, 13] and further establishes WES as a first-tier diagnostic tool for high-risk children. This diagnostic performance is significantly superior to traditional sequential testing strategies: for instance, the standalone diagnostic yield of chromosomal microarray analysis (CMA) as a first-tier test for IEI is only about 4.5% [14]; targeted gene panel sequencing for specific IEI phenotypes maintains a diagnostic yield mostly in the range of 15–25% [15], yet it is limited by its predefined gene content; as for techniques such as Sanger sequencing or FISH, they are primarily used for specific variant confirmation or research scenarios due to their limited throughput. Our integrated WES/CNV pipeline enables simultaneous, broad screening for pathogenic variants across a wide range of genes, detecting both point variants and copy number variants, thereby achieving higher

diagnostic efficiency. Crucially, our data indicate that diagnostic success depends not merely on sequencing depth or the size of a gene panel, but more fundamentally on clinical context, analytical rigor, and standardized bioinformatic pipelines. Recent studies have also confirmed that WES combined with CNV analysis can increase the overall diagnostic yield for genetic disorders to 30–40% or even higher, particularly for diseases like IEI [16]. Therefore, the data from this study further support the use of comprehensive WES/CNV analysis as a first-tier molecular diagnostic tool for suspected IEI, to replace or prioritize over traditional stepwise strategies. This approach holds promise for significantly shortening the time to diagnosis and alleviating the patients' "diagnostic odyssey."

Age-Stratified Clinical Evolution

The significant overall age distribution ($\chi^2 = 32.25$, $p < 0.0001$), combined with consistent clinical trends across age groups, supports the concept of IEI as a dynamic, age-stratified spectrum rather than a static collection of monogenic disorders. While pairwise comparisons did not retain statistical significance after multiple-testing correction—likely due to limited sample sizes in older age groups the observed phenotypic trajectories align with known principles of immune ontogeny and longitudinal observations from European cohorts [17], evolving from 'immune developmental failure' in infancy to 'immune dysregulation and malignant transformation' in adolescence. This exploratory analysis provides a hypothesis-generating framework for age-tailored diagnostic strategies, offering clinicians a practical, age-informed basis for risk stratification and stage-specific interventions, warranting validation in larger, multicenter cohorts. Notably, DiGeorge syndrome (22q11.2 deletion) emerged as the single most frequent diagnosis, surpassing classic monogenic disorders such as X-linked agammaglobulinemia (BTK) or severe combined immunodeficiency (IL2RG) that predominate in Western series [18, 19]. This discrepancy may reflect population-specific genetic architecture, founder effects, or environmental modifiers, highlighting the limitations of extrapolating Eurocentric diagnostic priorities to Asian populations. Similar regional variations are documented in other genetic conditions, such as G6PD deficiency in Southeast Asia [20], underscoring the need for locally curated genomic reference databases.

Phenotypic Spectrum and Clinical Red Flags

Our observation that non-infectious manifestations—congenital heart disease (17.28%), growth retardation (18.11%), and neurodevelopmental disorders (8.64%)—are highly prevalent supports the IUIS 2024 reclassification of IEI as systemic disorders [21, 22]. This aligns with growing evidence that immune-related genes often exert pleiotropic roles in development and homeostasis [23, 24]. For example, although PMS2 deficiency is fundamentally a DNA mismatch repair disorder, its inclusion in the IUIS classification reflects the well-documented immune dysregulation and lymphoma susceptibility resulting from defective immune surveillance in this syndrome. Importantly, 67.5% of patients presented with only 1–3 symptoms, and even single-symptom cases yielded a diagnostic rate of 45%, challenging the notion that extensive phenotyping is prerequisite for genomic testing. This resonates with prior studies showing that early WES in infants with isolated cytopenia or failure to thrive frequently yields actionable diagnoses [25], supporting a lowered threshold for WES in select high-risk scenarios,

particularly where immunological workup is limited. Finally, the elevated proportion of X-linked disorders (20 cases, 25.97%) and male predominance (M:F = 2.21:1)—exceeding European estimates (~ 15%) [26]—likely reflects both biological susceptibility and referral bias toward more severely affected boys. Regardless of etiology, male sex emerges as a robust clinical red flag, warranting its inclusion in IEI screening algorithms.

Value of Trio Sequencing, CNV Analysis, and Interpretation Standardization

Trio sequencing nearly doubled the diagnostic yield, highlighting its indispensable role in IEI genomics, especially for sporadic cases without a family history—a common scenario in our cohort (42 cases, 54.55% carried *de novo* variants). This finding aligns with the conclusion of a recent study on patients with suspected sporadic IEI, which also emphasized the crucial role of trio-WES in definitively identifying *de novo* variants and improving diagnostic accuracy [27]. Therefore, the results of this study support the adoption of trio-WES as a first-tier test for children with suspected IEI, which should be prioritized when feasible, considering the long-term cost-effectiveness of reducing diagnostic delays. CNVs accounted for 27.27% (21 cases)of diagnoses—emphasizing that omission of CNV analysis risks missing nearly one-quarter of pathogenic lesions. Moreover, we identify variability in variant interpretation as a previously underappreciated source of diagnostic inconsistency. While referral patterns among physicians showed no significant variation ($\chi^2 = 8.35$, $p = 0.079$), bioinformatic analysts exhibited notable differences in diagnostic rates (11.54%–53.33%; $\chi^2 = 22.87$, $p = 0.00013$). We acknowledge that this exploratory analysis may introduce confounding factors; nevertheless, this finding mirrors broader concerns raised by the Global Rare Diseases Registry regarding inconsistent application of ACMG/AMP guidelines and the lack of inter-laboratory calibration [28–30]. Rather than attributing variability to individual performance, we emphasize the systemic need for standardized training, blinded re-review, and centralized quality assurance—principles now implemented in consortia such as the Undiagnosed Diseases Network [31, 32]To address interpretation subjectivity, our center has established a systematic quality control framework: mandatory dual-independent review for all VUS cases, regular multidisciplinary calibration meetings, and ongoing competency training. These measures aim to enhance diagnostic consistency and equity, particularly in resource-limited settings where access to specialized expertise may be limited.

Limitations

Although this study achieved a molecular diagnostic yield of 31.69% through whole-exome sequencing combined with copy number variation analysis, 68.3% of patients remain without a definitive genetic diagnosis. This "diagnostic gap" may stem from various factors, including regulatory variants in non-coding regions, structural rearrangements, mosaic variants, novel disease-causing genes not yet cataloged in existing disease-gene databases, or complex genetic mechanisms that are challenging to capture with current bioinformatics pipelines. Several limitations warrant acknowledgment. First, this was a single-center retrospective study, which may introduce referral bias toward more severe or atypical cases, potentially overestimating the diagnostic yield compared to community-based populations.

Second, the proportion of trio-based sequencing was limited (6.2%) due to cost constraints, restricting the statistical power for subgroup analyses regarding inheritance patterns. Third, while we adhered to ACMG/AMP guidelines, functional validation was not performed for all novel variants, relying instead on computational predictions and phenotype-genotype correlations.

CONCLUSION

In conclusion, this study evaluated the diagnostic utility of WES combined with CNV analysis in an IEI cohort from a resource-limited region. Our findings demonstrate that integrating trio sequencing with CNV detection significantly enhances diagnostic efficiency, while the observed age-dependent phenotypic patterns support the implementation of age-stratified diagnostic strategies. Additionally, the data underscore the importance of establishing interpretation frameworks adapted to local genetic architectures and lowering the testing threshold for high-risk infants. Despite remaining challenges, these results provide empirical evidence for optimizing diagnostic workflows in similar settings. Future efforts should focus on integrating whole-genome sequencing, functional analyses, and periodic data reanalysis strategies to explore the underlying etiologies of undiagnosed cases, thereby further narrowing the diagnostic gap and expanding our understanding of the genetic landscape of IEI. Concurrently, expanding local variant databases and improving access to testing will advance diagnostic equity in underrepresented regions, ultimately improving clinical management for children with suspected IEI.

Data availability statement

Data generated in this study is either published in the article or supplementary data, or sensitive. For any inquiries, please contact corresponding author.

Declarations

Ethics approval and consent to participate

The studies involving humans were approved by Institutional Review Board and Ethics Committee of Guangxi Maternal and Child Health Hospital(Approval No.2017211).

Consent to Participate:

Written informed consent was obtained from all participants and/or their legal guardians prior to their inclusion in the study.

Consent to Publish:

All patients and/or their parents signed informed consent regarding publishing their data.

Competing interests

The authors assert that the study was conducted without any existing commercial or financial associations that might be perceived as potential conflicts of interest.

Funding

The research received financial support from the Guangxi Science and Technology Program (21-220-22 and 2025LTLT06002), Guangxi Science and Technology Plan Project(GuiKe LT2600640039), Guangxi Zhuang Region Health Department（Z20190311, Z-A20230305, Z-A20220256）, Liaoning Province Applied Basic Research Program Project (2023JH2/101300044), the National Natural Science Foundation of China (82160172) and the Natural Science Foundation of Guangxi Zhuang Autonomous Region, China（2025GXNSFAA069024）.

Author Contribution

Conceptualization, Qiang Zhang, Jihong Zhang and Jingsi Luo; Funding acquisition, Qiang Zhang，Jihong Zhang , Qi Yang and Jingsi Luo; Investigation, Qi Yang, Sheng Yi; Project administration, Jihong Zhang and Jingsi Luo; Resources, Jingsi Luo; Supervision, Jingsi Luo; Validation, Zailong Qin; Writing – original draft, Qiang Zhang; Writing – review & editing, Qiang Zhang, Qi Yang, Sheng Yi, Zailong Qin, Jihong Zhang and Jingsi Luo.

Acknowledgement

We extend our sincere gratitude to all members for their participation in this study.

Data Availability

Data generated in this study is either published in the article or supplementary data, or sensitive. For any inquiries, please contact corresponding author.

References

1. Bousfiha A, Moundir A, Tangye SG, Picard C, Jeddane L, Al-Herz W, et al. The 2022 Update of IUIS Phenotypical Classification for Human Inborn Errors of Immunity. J Clin Immunol. 2022;42:1508–20. https://doi.org/10.1007/s10875-022-01352-z.
2. Poli MC, Aksentijevich I, Bousfiha AA, Cunningham-Rundles C, Hambleton S, Klein C, et al. Human inborn errors of immunity: 2024 update on the classification from the International Union of

Immunological Societies Expert Committee. J Hum Immun. 2025;1:e20250003. https://doi.org/10.70962/jhi.20250003.

3. Quinn J, Modell V, Orange JS, Modell F. Growth in diagnosis and treatment of primary immunodeficiency within the global Jeffrey Modell Centers Network. Allergy Asthma Clin Immunol. 2022;18:19. https://doi.org/10.1186/s13223-022-00662-6.
4. Mao HW, Zhao XD. [Futher strengthen the immunoglobulin G replacement therapy of primary immunodeficiency diseases]. Zhonghua Er Ke Za Zhi. 2017;55:1–3. https://doi.org/10.3760/cma.j.issn.0578-1310.2017.01.001.
5. Deng M, Mao H. Inborn errors of immunity in mainland China: the past, present and future. BMJ Paediatr Open. 2023;7:e002002. https://doi.org/10.1136/bmjpo-2023-002002.
6. Xia Y, He T, Luo Y, Li C, Lim CK, Abolhassani H, et al. Targeted next-generation sequencing for genetic diagnosis of 160 patients with primary immunodeficiency in south China. Pediatr Allergy Immunol. 2018;29:863–72. https://doi.org/10.1111/pai.12976.
7. Wang L-L, Jin Y-Y, Hao Y-Q, Wang J-J, Yao C-M, Wang X, et al. Distribution and clinical features of primary immunodeficiency diseases in Chinese children (2004–2009). J Clin Immunol. 2011;31:297–308. https://doi.org/10.1007/s10875-010-9493-3.
8. SHI LY, ZHANG F, SUN ZC, AN YF, LI J, LIU Z, et al. Investigation on Diagnosis, Treatment and Medical Insurance of Children with Primary Immunodeficiency Diseases in China. China J Pharm Econ. 2021;16:24–9. https://doi.org/10.12010/j.issn.1673-5846.2021.07.005.
9. Petersen B-S, Fredrich B, Hoeppner MP, Ellinghaus D, Franke A. Opportunities and challenges of whole-genome and -exome sequencing. BMC Genet. 2017;18:14. https://doi.org/10.1186/s12863-017-0479-5.
10. Moundir A, Aissaoui O, Akhrichi N, Allaoui A, Benhsaien I, Jouanguy E, et al. Application of whole-exome sequencing to predict inborn errors of immunity in pediatric severe infections and sepsis. Clin Exp Immunol. 2025;219:uxaf007. https://doi.org/10.1093/cei/uxaf007.
11. Li Q, Wang K, InterVar. Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet. 2017;100:267–80. https://doi.org/10.1016/j.ajhg.2017.01.004.
12. Rosina E, Pezzani L, Apuril E, Pezzoli L, Marchetti D, Bellini M, et al. Comparison of first-tier whole-exome sequencing with a multi-step traditional approach for diagnosing paediatric outpatients: An Italian prospective study. Mol Genet Genomic Med. 2024;12:e2316. https://doi.org/10.1002/mgg3.2316.
13. Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. NPJ Genom Med. 2018;3:16. https://doi.org/10.1038/s41525-018-0053-8.
14. Beers BJ, Similuk MN, Ghosh R, Seifert BA, Jamal L, Kamen M, et al. Chromosomal microarray analysis supplements exome sequencing to diagnose children with suspected inborn errors of immunity. Front Immunol. 2023;14:1172004. https://doi.org/10.3389/fimmu.2023.1172004.

15. Chinen J, Lawrence M, Dorsey M, Kobrynski LJ. Practical approach to genetic testing for primary immunodeficiencies. Ann Allergy Asthma Immunol. 2019;123:433–9. https://doi.org/10.1016/j.anai.2019.08.016.
16. Heimall JR, Hagin D, Hajjar J, Henrickson SE, Hernandez-Trujillo HS, Tan Y, et al. Use of Genetic Testing for Primary Immunodeficiency Patients. J Clin Immunol. 2018;38:320–9. https://doi.org/10.1007/s10875-018-0489-8.
17. Thalhammer J, Kindle G, Nieters A, Rusch S, Seppänen MRJ, Fischer A, et al. Initial presenting manifestations in 16,486 patients with inborn errors of immunity include infections and noninfectious manifestations. J Allergy Clin Immunol. 2021;148:1332–e13415. https://doi.org/10.1016/j.jaci.2021.04.015.
18. Jiang C, He Y, Chen X, Xia F, Shi F, Xu X, et al. X-linked severe combined immunodeficiency complicated by disseminated bacillus Calmette-Guérin disease caused by a novel pathogenic mutation in exon 3 of the IL2RG gene: a case report and literature review. Front Immunol. 2024;15:1453046. https://doi.org/10.3389/fimmu.2024.1453046.
19. Aadam Z, Kechout N, Barakat A, Chan K-W, Ben-Ali M, Ben-Mustapha I, et al. X-Linked Agammagobulinemia in a Large Series of North African Patients: Frequency, Clinical Features and Novel BTK Mutations. J Clin Immunol. 2016;36:187–94. https://doi.org/10.1007/s10875-016-0251-z.
20. Tantular IS, Kawamoto F. Distribution of G6PD deficiency genotypes among Southeast Asian populations. Trop Med Health. 2021;49:97. https://doi.org/10.1186/s41182-021-00387-z.
21. Tangye SG, Al-Herz W, Bousfiha A, Cunningham-Rundles C, Franco JL, Holland SM, et al. Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. J Clin Immunol. 2022;42:1473–507. https://doi.org/10.1007/s10875-022-01289-3.
22. Cortesi M, Dotta L, Cattalini M, Lougaris V, Soresina A, Badolato R. Unmasking inborn errors of immunity: identifying the red flags of immune dysregulation. Front Immunol. 2024;15:1497921. https://doi.org/10.3389/fimmu.2024.1497921.
23. Williams AM, Ngo TM, Figueroa VE, Tate AT. The Effect of Developmental Pleiotropy on the Evolution of Insect Immune Genes. Genome Biol Evol. 2023;15:evad044. https://doi.org/10.1093/gbe/evad044.
24. Dubovik T, Starosvetsky E, LeRoy B, Normand R, Admon Y, Alpert A et al. Architecture of a multi-cellular polygenic network governing immune homeostasis [Internet]. bioRxiv; 2018 [cited 2025 Oct 30]. 256073. https://doi.org/10.1101/256073.
25. Scholz T, Blohm ME, Kortüm F, Bierhals T, Lessel D, van der Ven AT, et al. Whole-Exome Sequencing in Critically Ill Neonates and Infants: Diagnostic Yield and Predictability of Monogenic Diagnosis. Neonatology. 2021;118:454–61. https://doi.org/10.1159/000516890.
26. Lougaris V, Pession A, Baronio M, Soresina A, Rondelli R, Gazzurelli L, et al. The Italian Registry for Primary Immunodeficiencies (Italian Primary Immunodeficiency Network; IPINet): Twenty Years of

Experience (1999–2019). J Clin Immunol. 2020;40:1026–37. https://doi.org/10.1007/s10875-020-00844-0.

27. Hebert A, Simons A, Schuurs-Hoeijmakers JHM, Koenen HJPM, Zonneveld-Huijssoon E, Henriet SSV, et al. Trio-based whole exome sequencing in patients with suspected sporadic inborn errors of immunity: A retrospective cohort study. Elife. 2022;11:e78469. https://doi.org/10.7554/eLife.78469.
28. Amendola LM, Muenzen K, Biesecker LG, Bowling KM, Cooper GM, Dorschner MO, et al. Variant Classification Concordance using the ACMG-AMP Variant Interpretation Guidelines across Nine Genomic Implementation Research Studies. Am J Hum Genet Elsevier. 2020;107:932–41. https://doi.org/10.1016/j.ajhg.2020.09.011.
29. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho Y-Y, et al. Sherloc: a comprehensive refinement of the ACMG–AMP variant classification criteria. Genet Med. 2017;19:1105–17. https://doi.org/10.1038/gim.2017.37.
30. Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. Am J Hum Genet. 2016;98:1067–76. https://doi.org/10.1016/j.ajhg.2016.03.024.
31. Ramoni RB, Mulvihill JJ, Adams DR, Allard P, Ashley EA, Bernstein JA, et al. The Undiagnosed Diseases Network: Accelerating Discovery about Health and Disease. Am J Hum Genet. 2017;100:185–92. https://doi.org/10.1016/j.ajhg.2017.01.006.
32. Dr KS, Ca A, Hj B, Ja B, Am B. Effect of Genetic Diagnosis on Patients with Previously Undiagnosed Disease. The New England journal of medicine [Internet]. N Engl J Med. 2018. https://doi.org/10.1056/NEJMoa1714458. [cited 2025 Oct 30];379.

Tables

Tables are available in the Supplementary Files section.

Figures

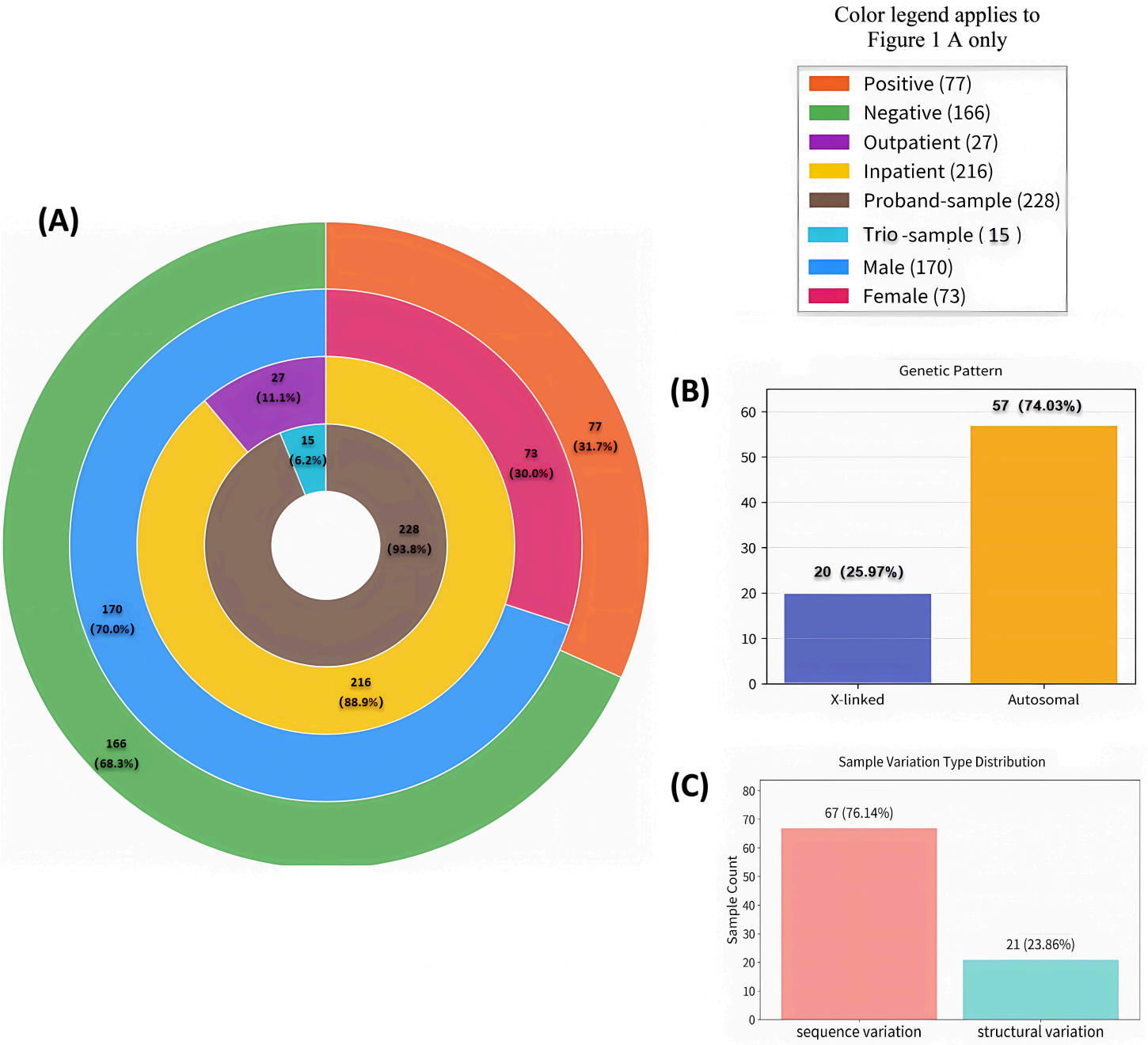


Figure 1

Overview of cohort characteristics and genetic architecture among 243 patients with suspected IEI.

(A) Patient cohort stratification by sample type, clinical setting, diagnosis outcome, and sex.From innermost to outermost ring:Layer 1 (Sample Type): 228 proband-only samples (93.8%) and 15 trio samples (6.2%).Layer 2 (Clinical Setting): 216 inpatients (88.9%) and 27 outpatients (11.1%).Layer 3 (Diagnosis Outcome): 77 molecularly confirmed cases (31.7%) and 166 negative cases (68.3%).Layer 4 (Sex Distribution): 170 males (70.0%) and 73 females (30.0%).

(B) Genetic inheritance pattern among the 77 confirmed IEI cases.Autosomal disorders predominated (57 cases, 74.03%) over X-linked disorders (20 cases, 25.97%).

(C) Distribution of variant types among the 77 confirmed cases.Sequence variations (67 cases, 76.14%) were more frequent than structural variations (21 cases, 23.86%).

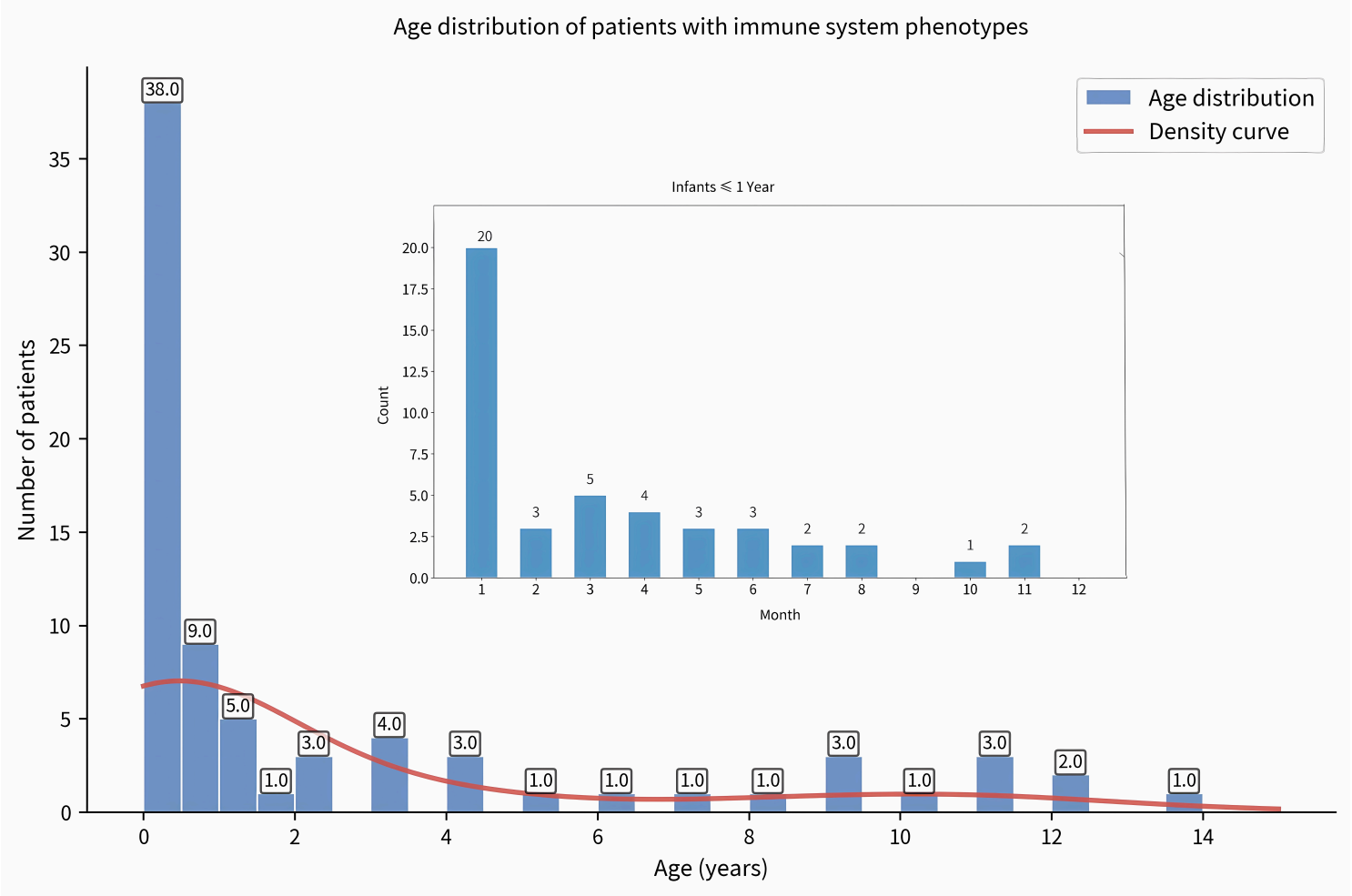


Figure 2

Age distribution of patients with immune system phenotypes. The main histogram displays the number of patients across different age groups (in years), with a density curve overlaid to show the distribution trend. A pronounced peak is observed in infants aged 0-1 year, particularly within the first 6 months (shown in the inset). The number of patients declines sharply after infancy, with smaller, scattered peaks in older age groups.

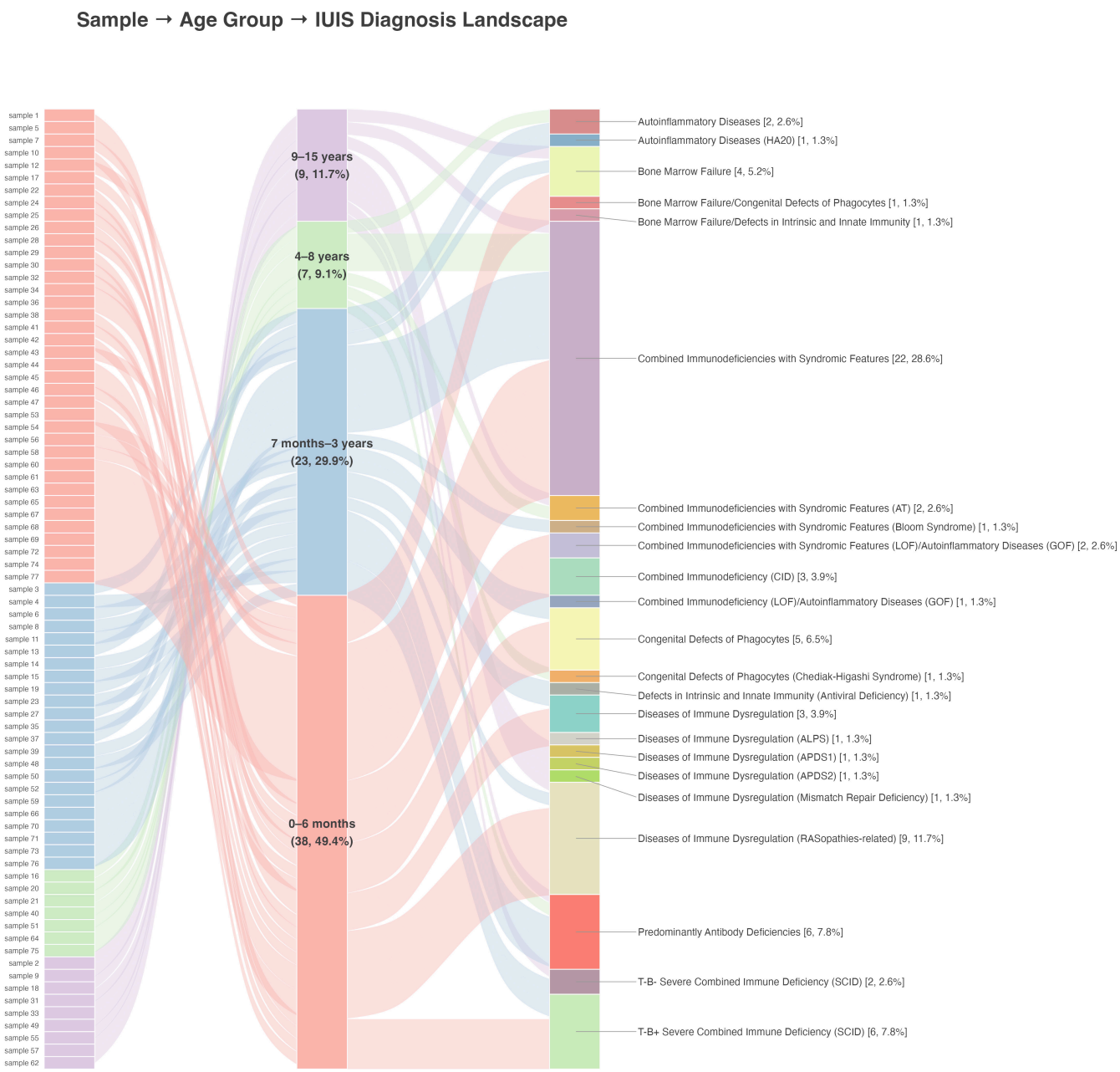


Figure 3

Sankey diagram illustrating the age-stratified evolution of IEI phenotypes. This diagram visualizes the flow of individual patient cases (labeled as Sample 1 to Sample 77) from their age at diagnosis (left) to their specific IUIS-classified disease category (right). It clearly demonstrates that infants (0–6 months) are predominantly diagnosed with severe combined immunodeficiencies (SCID) and congenital defects of phagocytes. As patients age, the spectrum shifts towards antibody deficiencies, autoinflammatory diseases, and complex syndromic IEI in the 9–15-year group.

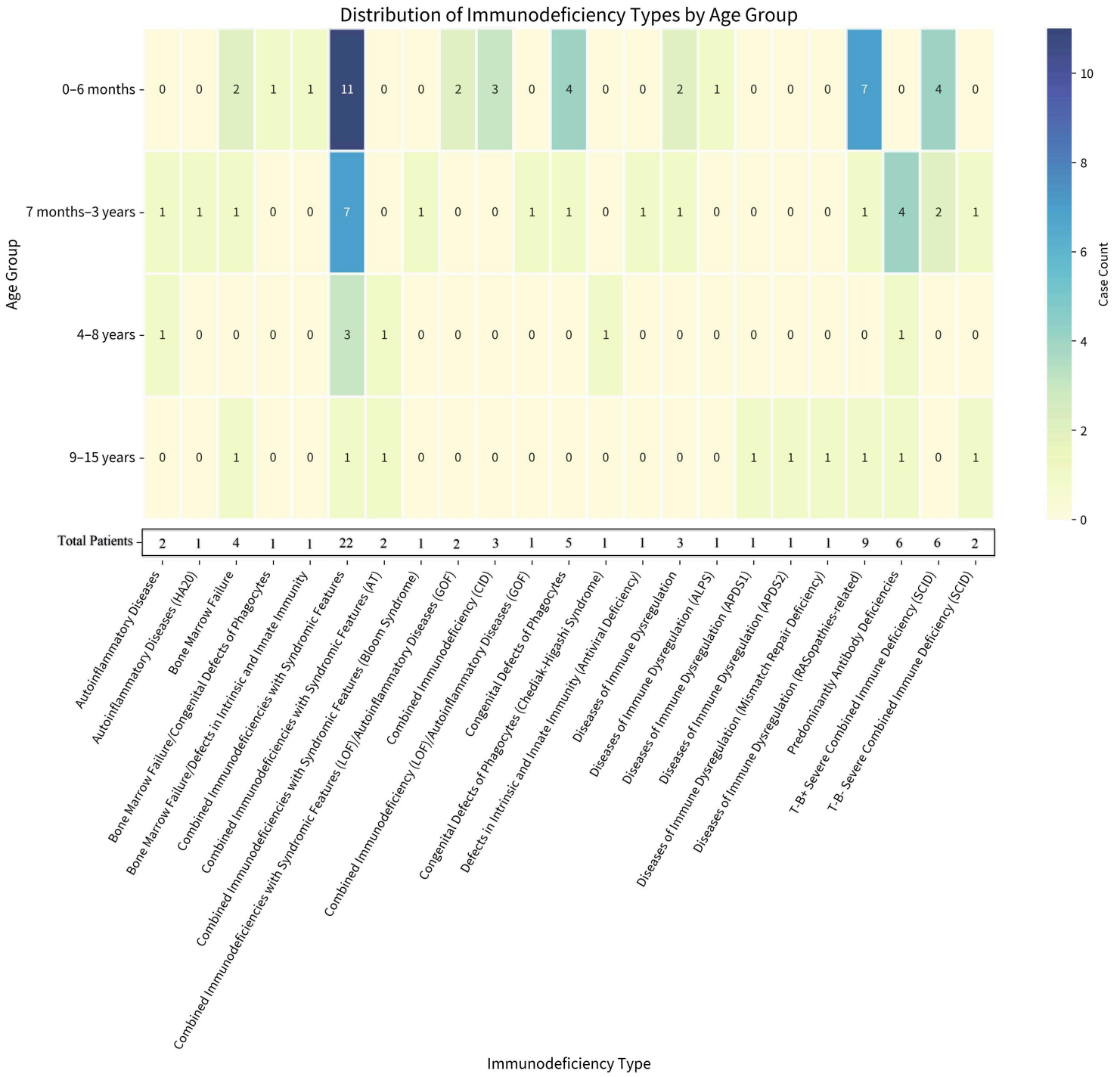


Figure 4

Heatmap of immunodeficiency type distribution across age groups.

This heatmap displays the case count for each IUIS-defined immunodeficiency type (columns) stratified by four age groups (rows: 0–6 months, 7 months–3 years, 4–8 years, 9–15 years). The color intensity reflects the number of cases, with darker shades indicating higher counts. The data reveal a clear shift: the 0–6 month group is dominated by T-B+ SCID (11 cases) and DiGeorge syndrome (7 cases), while the 9–15 year group exhibits greater diversity, including multiple cases of immune dysregulation syndromes (e.g., APDS1, APDS2) and cancer-predisposing conditions.

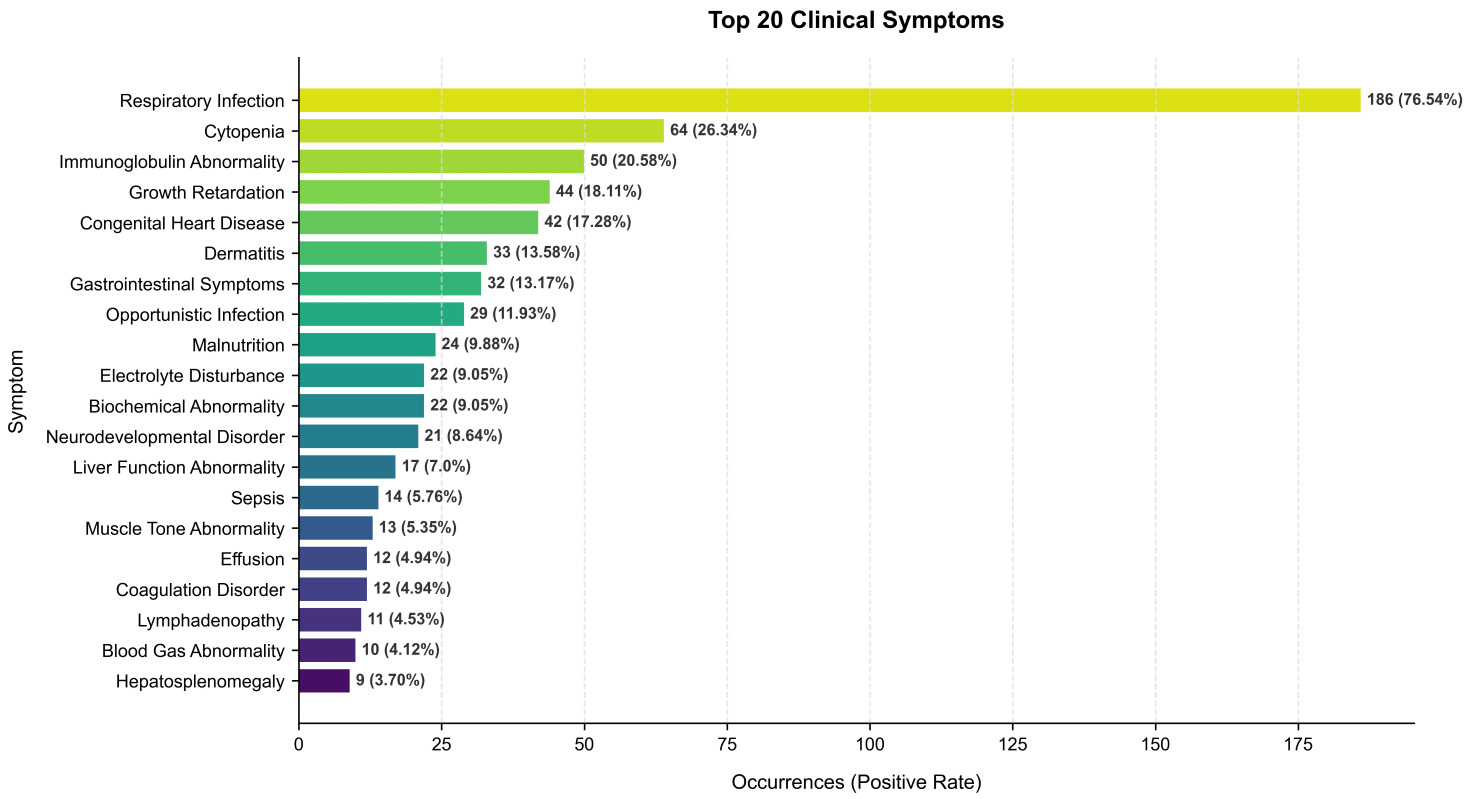


Figure 5

Top 20 clinical symptoms among the 243 pediatric IEI suspects. A horizontal bar chart ranks the top 20 clinical manifestations by their total occurrences across all patients. Respiratory infection was the most prevalent symptom (186 occurrences), followed by cytopenia (64) and immunoglobulin abnormality (50). Notably, non-infectious features such as growth retardation (44), congenital heart disease (42), and neurodevelopmental disorder (21) also ranked highly, underscoring the systemic nature of IEI.

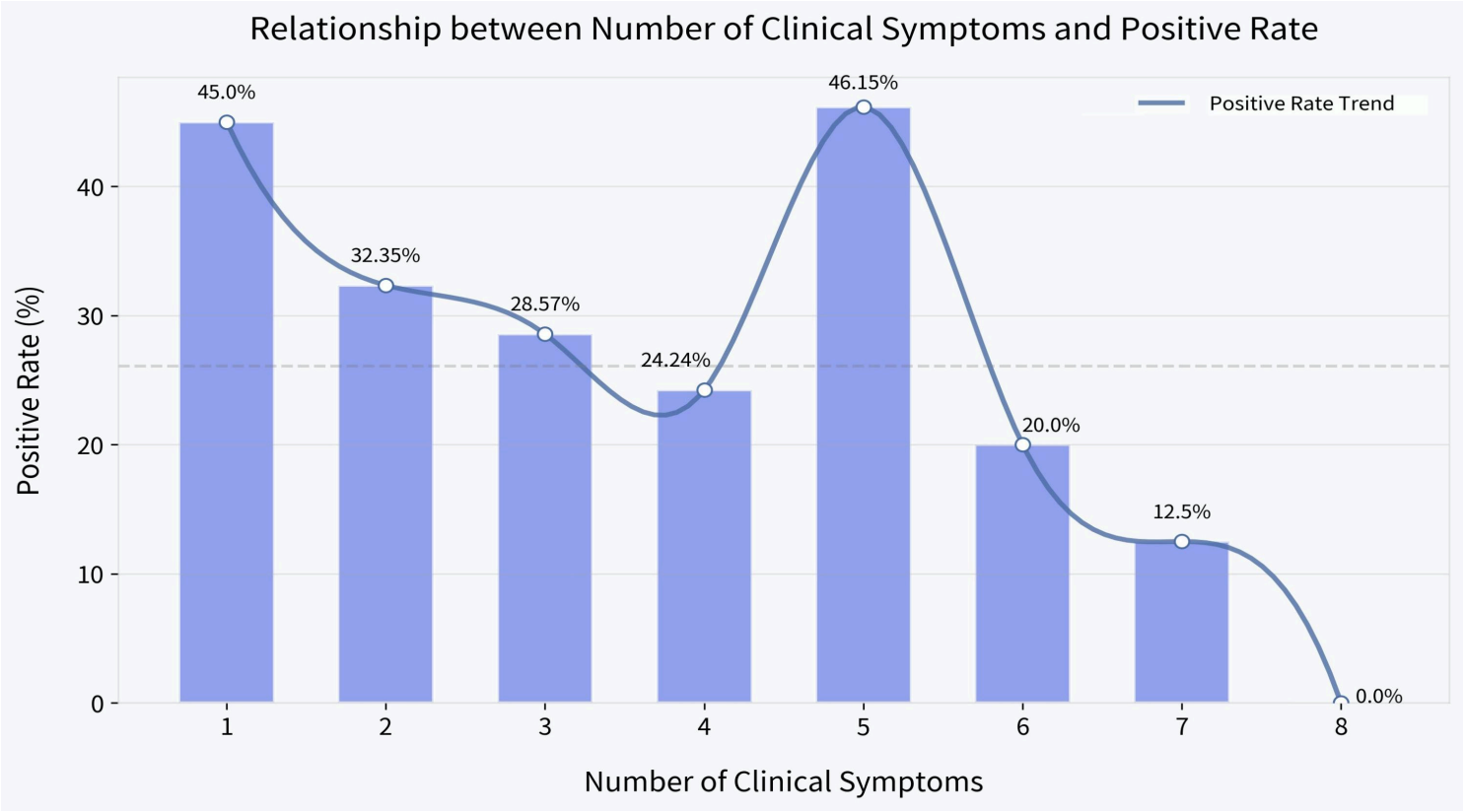


Figure 6

Relationship between the number of clinical symptoms and diagnostic positivity rate.The data reveal an approximately inverted U-shaped curve, with the highest positivity rate (46.15%) observed in patients presenting with five symptoms. In contrast, the diagnostic positivity rate drops markedly among patients with more complex, multi-system presentations (≥6 symptoms). Notably, even patients exhibiting only a single, highly suggestive symptom achieved a substantial positivity rate of 45.0%.

Data for ≥9 symptoms are based on a very small sample size (n ≤ 4), resulting in insufficient statistical power and high susceptibility to random variation; therefore, they were excluded from the formal statistical analysis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx
- Table2.docx
- SupplementaryFigure1.pdf
- SupplementaryTableS1.pdf