Year: September 2026 Volume: 13 Issue: 3

E-ISSN: 2587-2478

The Journal of Pediatric Research

Review

Amino Acid Pattern in Pediatric Obesity
Soliman et al.

Original Articles

Genotype-phenotype Analysis in Congenital Adrenal Hyperplasia
Bilen et al.

Prospective Study of Blood Pressure Trends
Güran et al.

Immune Dysregulation in Children
Çelebi Çelik et al.

Pediatric Occipital Epilepsy Spectrum
Ayfer Arduç Akçay.

Heart Rate Variability Children Palliative Care
Beyazal and Ayar.

Cerebral Edema in Diabetes Ketoacidosis
Sarvari et al.

Virus-specific Early Hospitalization in Children
Bedir and Altuntaş.

CVC-related Complications in Paediatric Neurosurgery
Barros et al.

Pediatric Tracheostomy and Etiology
Durak et al.

Breastfeeding Self-efficacy and Infant Feeding Attitudes
Zengin et al.

LLM Concordance in Pediatric Penile Conditions
Türedi and Sezer.

Readability, Quality, and Consistency of a Large Language Model
Bitisik et al.

Case Report

Ochoa Syndrome: A Case Report
Ekberli et al.

galenos Publishing House

JPR
The Journal of Pediatric Research
Official Journal of Ege University Children's Hospital

EDITORIAL BOARD

The Journal of Pediatric Research is the official publication of Ege University Faculty of Medicine, Department of Pediatrics and Ege Children's Foundation.

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1
34093 İstanbul, Türkiye
Phone: +90 (530) 177 30 97 / +90 (539) 307 32 03
E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr
Publisher Certificate Number: 14521

Online Publication Date: September 2026
E-ISSN: 2587-2478
International scientific journal published quarterly.

JPR
The
Journal of Pediatric Research
Official Journal of Ege University Children's Hospital

EDITORIAL BOARD

JPR
The
Journal of Pediatric Research
Official Journal of Ege University Children's Hospital

EDITORIAL BOARD

JPR
The Journal of Pediatric Research
Official Journal of Ege University Children's Hospital

EDITORIAL BOARD

The Journal of Pediatric Research is the official publication of Ege University and Ege Children's Foundation

Please refer to the journal's webpage **(https://jpedres.org/)** for "Aims and Scope", "Instructions to Authors" and "Ethical Policy".

The editorial and publication processes of the Journal of Pediatric Research are shaped in accordance with the guidelines of ICMJE, WAME, CSE, COPE, EASE, and NISO. The journal is in conformity with the Principles of Transparency and Best Practice in Scholarly Publishing.

The Journal of Pediatric Research is indexed in **Emerging Sources Citation Index (ESCI), Embase, Directory of Open Access Journals (DOAJ), EBSCO, CINAHL Complete Database, ProQuest, CABI, Gale/Cengage Learning, Ulakbim TR Dizin, TurkMedline, J-GATE, IdealOnline, Hinari, GOALI, ARDI, OARE, AGORA** and **Türkiye Citation Index**.

The journal is published online.

Owner: Ege University and Ege Children's Foundation

Responsible Manager: Özgür ÇOĞULU

JPR
The
Journal of Pediatric Research
Official Journal of Ege University Children's Hospital

SCIENTIFIC ADVISORY BOARD

JPR
The
Journal of Pediatric Research
Official Journal of Ege University Children's Hospital

CONTENTS

Review

193 ▶ **Amino Acid Pattern Signatures in Pediatric Obesity, Dysglycemia, and Fatty Liver Disease: Links to Insulin Resistance and the IGF-1 Axis**
Ashraf T Soliman, Hendeh Zirak, Fawzia Alyafei, Nada Alaaraj, Shayma Ahmed, Noor Hamed, Sohair Abdeldaim Ali Elsiddig; Doha, Qatar

Original Articles

209 ▶ **Genotype-Phenotype Correlation Analysis and Identification of a Novel Variant in Children with Congenital Adrenal Hyperplasia due to 21-Hydroxylase Deficiency**
Nimet Melis Bilen, Gönül Çatlı, Cemil Koçyiğit, Bumin Nuri Dündar; İzmir, Türkiye

220 ▶ **Blood Pressure Trends and Influencing Factors in the First Week of Life: A Prospective Study in Healthy Term Turkish Newborns**
Ömer Güran, Sinan Uslu, Ali Bülbül, Evrim Kıray Baş, Muhittin Çelik, Umut Zübarioğlu; İstanbul, Türkiye

227 ▶ **Beyond Infections: Clinical and Genetic Spectrum of Pediatric Immune Dysregulation Disorders**
Figen Çelebi Çelik, Aymen Hişmioğulları, Soner Günder, Gülçin Kaymakoğlu, Nesrin Gülez, Ferah Genel; İzmir, Türkiye

238 ▶ **Clinical and Electroencephalographic Characteristics of Childhood Occipital Epilepsies: Comparison of Self-Limited and Symptomatic Forms**
Ayfer Arduç Akçay; İstanbul, Türkiye

247 ▶ **Heart Rate Variability in Children Receiving Palliative Care: An Exploratory Cross-Sectional Study**
Meryem Beyazal, Ganime Ayar; Ankara, Türkiye

254 ▶ **Prevalence and Risk Factors Related to Cerebral Edema Associated with Diabetes Ketoacidosis in Pediatrics with Type-1 Diabetes Mellitus**
Masoumeh Sarvari, Maedeh Sarvari, Reihaneh Mohsenipour, Mahsa Soti Khiabani, Mahya Sadat Mohammadi; Tehran, Qazvin, Iran

263 ▶ **Virus-Specific Patterns of Early Hospitalization after Pediatric Emergency Department Discharge: A Real-World Cohort Study**
Fırat Bedir, Mahmut Altuntaş; İstanbul, Türkiye

272 ▶ **Central Venous Catheter-Related Complications in a Paediatric Neurosurgical Population: A Retrospective Study**
Francisco Barros, Catarina Chaves, Patricia Santos, Helder Pereira, Joana Garcia Azevedo; Porto, Portugal

279 ▶ **Etiology-Based Clinical Outcomes after Pediatric Tracheostomy in Intensive Care: A Retrospective Observational Study**
Fatih Durak, Emine Pınar Küllüoğlu, Gökçen Özçifçi, Deniz Koçak Ter; İzmir, Türkiye

285 ▶ **Association between Breastfeeding Self-Efficacy and Infant Feeding Attitudes among Mothers of Infants Aged 0-6 Months**
Dilek Zengin, Feyza Koç, Selinnur Durmaz, Merve Tosyalı; İzmir, Türkiye

294 ▶ **Large Language Models in Pediatric Penile Conditions: Guideline Concordance, Readability, and Language-Based Performance**
Bilge Türedi, Ali Sezer; İstanbul, Kocaeli, Türkiye

303 ▶ **Can a Large Language Model Effectively Answer Parents' Questions About Children's Oral Health?**
Irem Nur Bitisik, Selcuk Savas, Ilgin Timarci, Mehmet Izgi, Ebru Kucukyilmaz Izgi; Izmir, Türkiye

Case Report

314 ▶ **Ochoa Syndrome Associated with a Novel HPSE2 Deletion: A Case Report**
Günay Ekberli, Sevgin Taner, Esra Işık, Türkan Turkut Tan; Ankara, İzmir, Türkiye

318 ▶ **Erratum**

JPR
The
Journal of Pediatric Research
Official Journal of Ege University Children's Hospital

EDITORIAL

Dear JPR Readers,

It is a privilege to introduce the September 2026 issue of The Journal of Pediatric Research. Across metabolism, genetics, neonatology, immunology, neurology, complex care, family support, and artificial intelligence, the articles ask a practical pediatric question: how can we recognize meaningful signals early, translate them into safer care, and equip families to act? In children, every signal is shaped by development, clinical context, and change over time.

The issue opens with a review of amino acid patterns in pediatric obesity, dysglycemia, and fatty liver disease. Elevations in branched-chain and aromatic amino acids and glutamate, together with reductions in glycine and serine, may provide earlier metabolic signals of insulin resistance and related risk, but their interpretation must remain sensitive to age, sex, puberty, and the IGF-1 axis. This leads naturally to the study of congenital adrenal hyperplasia, which found overall concordance between genotype and phenotype alongside discordant presentations and novel variants. The findings support integrating molecular information with hormonal findings, clinical expression, and longitudinal follow-up.

Recognition also requires understanding normal developmental variation. A prospective study provides descriptive blood-pressure values for healthy term Turkish newborns during the first postnatal week, demonstrating the expected increase over time. Although several maternal and neonatal factors were associated with first-day blood pressure, these relationships were weak and of limited clinical significance. The next study broadens recognition beyond familiar warning signs: children with immune dysregulation often presented with lymphoproliferation, cytopenias, autoimmunity, and dermatologic disease rather than infection alone. Its reported diagnostic delay reminds us that earlier diagnosis may depend on recognizing a wider phenotype.

From a pediatric neurologist's perspective, the value of following patterns over time is especially clear in childhood occipital epilepsies. Electroclinical and imaging findings support a dynamic spectrum rather than rigidly separated self-limited and symptomatic categories, reinforcing the importance of longitudinal EEG and neurocognitive follow-up. Heart-rate variability in children receiving palliative care may likewise offer a complementary window into autonomic regulation. Yet palliative care also reminds us that time can be measured in more than heartbeats. As clinicians attend to physiologic intervals, families may be trying to slow another kind of time. They may wish to be fully present, to hold a hand, and to create space for closeness with a precious child. Physiologic measures may help us understand the body, but they should also return us to the human purpose of palliative care: protecting comfort, presence, and meaningful time together.

Two acute-care studies continue this theme by examining clinical signals associated with cerebral edema in diabetic ketoacidosis and virus-specific patterns of early hospitalization after discharge from the pediatric emergency department. Both address which children may require closer observation, reassessment, or follow-up, without assuming that every child within a diagnostic category carries the same risk.

The issue then turns from recognition to preventing treatment-related harm. Research on central venous catheter complications in pediatric neurosurgical care emphasizes procedural risk and surveillance. After tracheostomy, outcomes differed according to etiology; chronic respiratory failure was associated with longer ventilation and

The
Journal of Pediatric Research

Official Journal of Ege University Children's Hospital

EDITORIAL

less successful weaning. Etiology-specific trajectories can guide individualized management and more realistic, compassionate conversations with families.

Equipping families to act is the next essential step. Higher breastfeeding self-efficacy was associated with more positive infant-feeding attitudes, making caregiver confidence a meaningful target for support. Two studies examine artificial intelligence as a source of pediatric information. Large language models provided generally guideline-concordant, readable answers about pediatric penile conditions in English and Turkish; Turkish responses about children's oral health showed high information quality but variable readability. As families increasingly turn to AI for health information, pediatric clinicians should serve as a trusted counterbalance, helping them weigh accuracy, relevance, and uncertainty. Our role is not simply to add another voice to an already crowded landscape. Journals such as JPR strengthen this work by building the evidence base from which trustworthy guidance can be translated. Looking ahead, teaching families how to ask informed questions, evaluate information, and advocate for their children will be an increasingly important part of pediatric care.

The issue concludes with a case of Ochoa syndrome associated with a novel homozygous HPSE2 deletion. The characteristic phenotype, recurrent urinary tract infections, and molecular finding form a relationship between genotype and phenotype with immediate clinical relevance. Early recognition and timely intervention may prevent progressive renal injury. This is a fitting conclusion to an issue centered on turning signals into thoughtful action.

We thank the authors, reviewers, editorial team, and Galenos Publishing House for their contributions. We hope this issue encourages readers to notice earlier signals, reduce avoidable harm, and help families translate an expanding body of information into informed advocacy and care.

Warm regards,

Tuba Rashid Khan, MD, MPH, EDM

J Pediatr Res 2026;13(3):193-208
DOI: 10.4274/jpr.galenos.2026.70094

Review

Amino Acid Pattern Signatures in Pediatric Obesity, Dysglycemia, and Fatty Liver Disease: Links to Insulin Resistance and the IGF-1 Axis

Ashraf T Soliman[1], Hendeh Zirak[2], Fawzia Alyafei[1], Nada Alaaraj[1], Shayma Ahmed[1], Noor Hamed[1], Sohair Abdeldaim Ali Elsiddig[1]

[1]Hamad Medical Corporation, Department of Pediatrics, Doha, Qatar
[2]Hamad Medical Corporation, Department of Nursing, Doha, Qatar

ABSTRACT

Amino acid (AA) pattern shifts are increasingly recognized as early biochemical fingerprints of pediatric obesity and its metabolic sequelae. In children and adolescents, elevations in branched-chain AAs (BCAAs: leucine, isoleucine, valine), aromatic AAs (AAAs: phenylalanine, tyrosine), and glutamate—often coupled with reductions in glycine and serine—have been repeatedly linked to insulin resistance (IR), dysglycemia, and metabolic dysfunction-associated fatty liver disease (MASLD/NAFLD). These signatures may reflect impaired BCAA catabolism, altered mitochondrial substrate handling, hepatic nitrogen flux, and adipose-liver crosstalk, and they may provide complementary risk stratification beyond standard clinical markers. (1) To summarize AA pattern changes characteristic of pediatric/adolescent obesity, prediabetes, diabetes, and fatty liver disorders and their relationship to insulin resistance. (2) To synthesize evidence linking AA signatures to IGF-1 biology and developmental covariates (age, sex, puberty), and to discuss management implications. A structured review (last 20 years) of PubMed/Scopus-indexed literature was performed, prioritizing pediatric/adolescent studies. We included targeted and untargeted metabolomics studies reporting fasting or standardized AA measures in: (i) obesity/obesity phenotypes, (ii) prediabetes/abnormal glucose tolerance, (iii) Type 1 or Type 2 diabetes, and (iv) NAFLD/MASLD, with outcomes involving IR (HOMA-IR, clamp-derived indices), hepatic steatosis (MRI/MRS, ultrasound, biopsy), or cardiometabolic risk clustering. We narratively synthesized AA directionality and consistency across conditions and platforms. Risk of bias was assessed using RoB 2 for randomized trials and ROBINS-I for non-randomized studies. Across multiple pediatric cohorts, obesity was most consistently associated with higher BCAAs, AAAs, glutamate, and several BCAA-derived intermediates, with lower glycine/serine in metabolically higher-risk phenotypes. Several studies linked AA signatures to future IR or hypertriglyceridemia and identified associations with metabolically unhealthy obesity. In youth with dysglycemia or Type 2 diabetes risk, AA profiles were heterogeneous across studies but increasingly support combined panels (BCAA/AAA elevation with glycine reduction and glutamine-glutamate imbalance) to capture β-cell stress and IR-related trajectories. In pediatric NAFLD, AA signatures (notably BCAA dysregulation and glutamate-related indices) were associated with steatosis burden and, in some studies, predicted liver fat independent of BMI and conventional risk factors. Developmental effects were substantial: age, sex, and pubertal stage modified AA distributions, and IGF-1-linked metabolomic relationships suggest GH/IGF axis-metabolism coupling. AA patterns provide biologically plausible, developmentally modulated candidate biomarkers which align with IR, dysglycemia, and pediatric fatty liver risk. Future pediatric studies should standardize sampling, puberty adjustment, and analytic pipelines, and should test whether AA-informed risk stratification improves prevention and treatment response monitoring.

Keywords: Amino acids, insulin resistance, pediatric obesity, NAFLD/MASLD, IGF-1

Corresponding Author
Prof. Ashraf T Soliman, Hamad Medical Corporation, Department of Pediatrics, Doha, Qatar
E-mail: atsoliman@yahoo.com **ORCID:** orcid.org/0000-0002-9585-9461
Received: 05.04.2026 **Accepted:** 31.07.2026 **Publication Date:** 18.09.2026

Cite this article as: Soliman AT, Zirak H, Alyafei F, et al. Amino acid pattern signatures in pediatric obesity, dysglycemia, and fatty liver disease: links to insulin resistance and the IGF-1 axis. J Pediatr Res. 2026;13(3):193-208

Introduction

Amino acids are not only building blocks for protein synthesis, but also key metabolic signals integrated with mitochondrial function, hepatic nitrogen handling, and nutrient-sensing pathways. Modern metabolomics has expanded the clinical view of AAs by revealing reproducible, disease-linked shifts in circulating AA patterns across metabolic disorders (1). Among these, elevations in BCAAs and AAAs have emerged as particularly robust correlates of insulin resistance and future diabetes risk in adult and pediatric populations, suggesting shared upstream mechanisms such as impaired BCAA catabolism, altered substrate competition, and dysregulated nutrient signaling (2-5).

Beyond BCAAs, several 'counter-pattern' AAs, especially glycine and serine, often show reductions in metabolically at-risk states. Glycine deficiency states may reflect increased one-carbon demand, oxidative stress buffering, and constrained conjugation capacity for acyl-group disposal, while the glutamine-glutamate axis can reflect altered transamination, hepatic metabolism, and mitochondrial flux (6,7). Evidence syntheses indicate that AA panels (BCAA/AAA elevation plus lower glycine/serine and glutamine-glutamate imbalance) are among the most consistent metabolite features linked to prediabetes/diabetes risk, supporting their role as integrated markers of early metabolic dysfunction (8).

In pediatrics, obesity is a heterogeneous condition with variable cardiometabolic risk clustering; therefore, metabolite-based fingerprints are appealing adjuncts to BMI-based classification. Systematic reviews of pediatric obesity metabolomics consistently highlight AA disturbances, particularly BCAAs, AAAs, and related intermediates, alongside lipid and acylcarnitine changes (9,10). Pediatric-focused reviews similarly emphasize that AA alterations can be detectable early in life, may precede overt dysglycemia, and plausibly connect adiposity to hepatic and muscle insulin resistance through impaired BCAA catabolism and nutrient signaling (11-13).

Multiple pediatric cohort and case-control studies have reported higher BCAAs and AAAs in children and adolescents with obesity and have linked these elevations to insulin resistance metrics and future metabolic risk (14-16). Metabolomics studies in obese prepubertal children indicate sex-dependent patterns and suggest that hyperinsulinemia/IR phenotypes are accompanied by distinct AA and related metabolite shifts (17,18). In addition, pediatric studies have shown that tyrosine and related signatures can track longitudinal IR trajectories, supporting AA measures as candidate early markers rather than mere epiphenomena (19,20).

Dysglycemia in youth spans impaired fasting glucose, impaired glucose tolerance, and youth-onset diabetes, with unique developmental considerations (puberty-related physiological IR, growth demands, and evolving β-cell compensation). Pediatric metabolomics studies evaluating AA profiles in dysglycemia and youth-onset diabetes have produced both convergent and divergent findings, likely reflecting differences in age, adiposity, diabetes type, pubertal stage, and analytic platform (21-25). Nevertheless, recent youth-focused work suggests that composite AA signatures may help identify youth at risk for β-cell failure and may distinguish diabetes types, which could support phenotype-aligned prevention and management strategies (23-25).

Pediatric fatty liver disease (NAFLD/MASLD) is closely linked to obesity and insulin resistance and is now a major concern in pediatric practice. As most cited pediatric studies predate the 2023 nomenclature revision, we retain the original study-reported terms (typically NAFLD) when summarizing individual investigations, while using the combined NAFLD/MASLD designation in synthesis statements which span the terminology transition; MASLD is used specifically when a cited study explicitly adopted the revised nomenclature. Clinical guidelines emphasize early identification and risk stratification, typically using clinical risk factors and imaging, yet blood-based metabolic fingerprints remain an active translational gap (26). Pediatric NAFLD metabolomics studies demonstrate AA pathway disruptions often involving BCAA dysregulation and glutamate-related patterns which can associate with hepatic fat independent of BMI and conventional risk markers (27-35). Importantly, normal developmental variation (age, sex, puberty) strongly influences AA distributions and intersects with growth biology and IGF-1 signaling (36,37). This creates a need for pediatric-specific interpretation frameworks and motivates this review's focus on AA pattern characteristics across obesity, dysglycemia, diabetes, and fatty liver disorders in relation to IR and the IGF-1 axis.

Objectives

1. To characterize AA pattern changes reported in pediatric/adolescent obesity, prediabetes, diabetes, and fatty liver disorders, and to summarize their associations with insulin resistance and metabolic risk phenotypes.

2. To integrate developmental and endocrine context, focusing on IGF-1 and puberty-related modifiers of AA signatures, and to discuss practical management implications and research priorities.

Materials and Methods

PRISMA 2020 principles were used to improve transparency of literature identification and study selection (14-53). As the included literature used heterogeneous analytic platforms (targeted versus untargeted metabolomics), varied sampling conditions (fasting versus standardized sampling windows), and reported outcomes which were not directly comparable across studies, we prioritized a structured narrative synthesis supported by tabulation rather than conducting a pooled meta-analysis (54-56).

We searched PubMed and Scopus for English-language articles published from January 1, 2006 through March 1, 2026. Search terms were built around four concept blocks: (i) amino acid exposure terms ('amino acid*,' 'branched chain amino acid*,' 'aromatic amino acid*,' 'glycine,' 'glutamate,' 'glutamine'), (ii) pediatric population terms ('child*,' 'adolescent*,' 'pediatric'), (iii) target conditions and outcomes (obesity, insulin resistance, prediabetes, impaired glucose tolerance, diabetes, fatty liver, NAFLD, MASLD), and (iv) metabolomic methods (metabolomics, mass spectrometry, LC-MS, GC-MS, NMR). In addition, reference lists of key reviews and eligible studies were hand-searched to identify relevant pediatric metabolomics papers that might not have been retrieved by the database terms alone. Titles and abstracts were screened for pediatric relevance and amino-acid content, followed by full-text assessment against the eligibility criteria described below.

Study selection and evidence identification. The database and reference-list searches identified literature addressing amino-acid and metabolomic patterns in pediatric obesity, dysglycemia/diabetes, fatty liver disease, and developmental or IGF-1-related metabolic phenotypes. Because the original record-level PubMed and Scopus search exports were not retained, exact database-specific identification, deduplication, and title/abstract screening counts could not be reliably reconstructed retrospectively and were therefore not estimated. A retrospective audit of the final evidence set identified 58 cited references. Three references were methodological/reporting sources relating to PRISMA 2020, RoB 2, and ROBINS-I and were not considered substantive evidence. Of the remaining 55 substantive references assessed, 23 consisted of systematic or narrative reviews, clinical practice guidelines, predominantly adult mechanistic/context studies, or other supporting sources that were not counted as primary pediatric studies. The final primary-study synthesis therefore comprised 32 unique pediatric or pediatric-relevant primary studies. Additional reviews, guidelines, and mechanistic studies were retained exclusively to provide biological, methodological, and clinical context and were clearly distinguished from the primary pediatric evidence (Figure 1).

Eligible studies primarily included children and adolescents (typically ≤18 years); however, we also included studies enrolling youth/young adults up to 24 years when the cohorts were clearly pediatric-relevant (for example, bariatric surgery cohorts initiated in adolescence). We included observational (cross-sectional, case-control, and cohort) and interventional designs which reported circulating amino acids measured under fasting or otherwise standardized sampling conditions. The included studies were required to report at least one relevant metabolic outcome, including insulin resistance (HOMA-IR, clamp-derived indices, or validated surrogates of insulin sensitivity), glycemic phenotypes (prediabetes, impaired glucose tolerance, youth-onset diabetes), and/or hepatic steatosis assessed by MRI/MRS, ultrasound, or histology. Studies using targeted amino-acid panels, untargeted metabolomics with explicit amino-acid outputs, or broader metabolite profiling which clearly reported amino-acid findings were all eligible. We excluded adult-only cohorts without pediatric stratification (unless used solely for mechanistic context in the IGF-1, NAFLD/MASLD, or intervention-focused sections, where they are explicitly labeled as adult/context-only and are not treated as pediatric evidence), non-human studies (not tabulated in this review), and reports which did not provide extractable directionality for amino-acid changes or did not include relevant metabolic outcomes.

For each eligible study, we extracted the study design, sample size, age range, phenotype definitions (obesity class, metabolic-risk clustering, dysglycemia definition), metabolomic platform, key amino-acid directionality (increase/decrease), and primary associations with insulin resistance and/or liver fat. As effect measures and analytic strategies varied widely, we synthesized findings using a structured qualitative framework. This included consistency mapping (whether a given amino-acid pattern was directionally replicated across multiple pediatric studies), phenotype anchoring (whether the pattern aligned more closely with metabolically unhealthy obesity, dysglycemia risk, or NAFLD/MASLD), and assessment of context

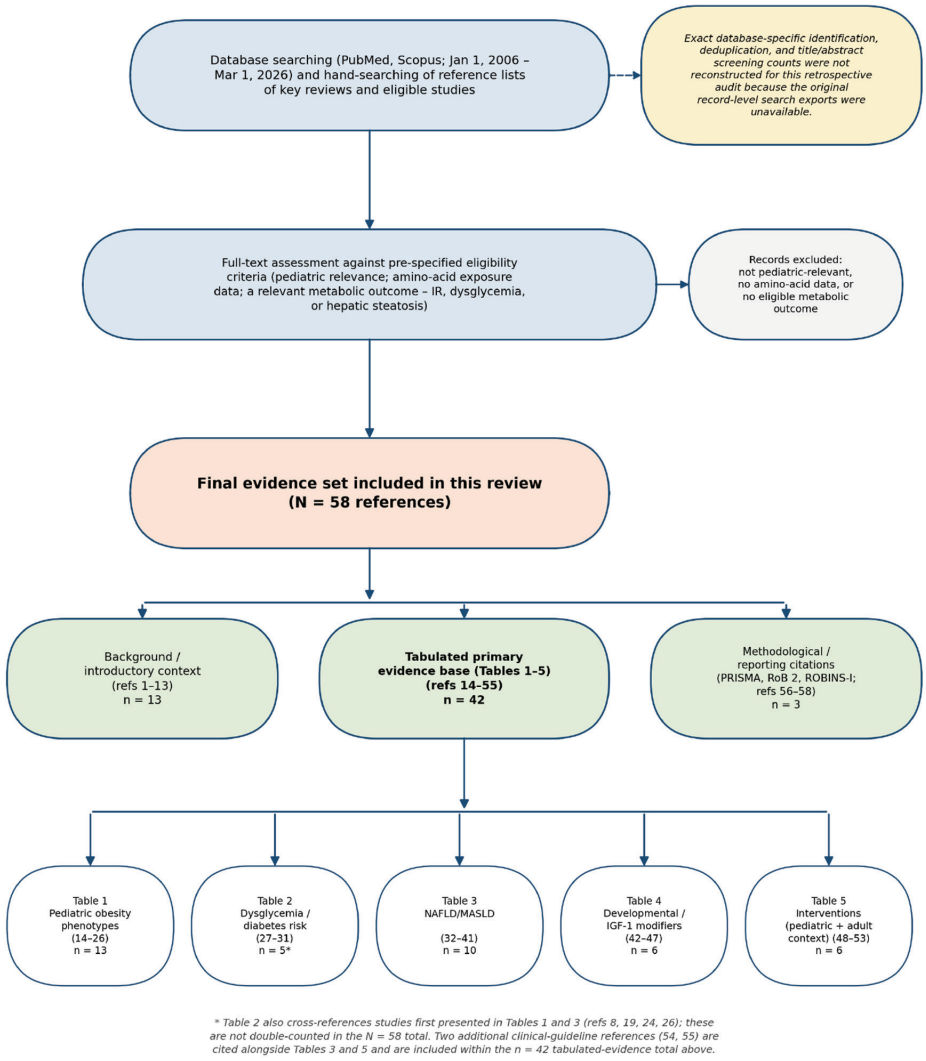

Figure 1. PRISMA 2020-style flow diagram of study selection (retrospective evidence-set audit). The diagram reports a retrospective audit of the final evidence set. Exact database-specific identification, deduplication, and title/abstract screening counts were not reconstructed because the original record-level search exports were unavailable

Algorithm (for figure box)

Step 1 - Identify the at-risk patient: Child/adolescent with overweight/obesity or metabolic risk (family history, rapid weight gain, acanthosis, elevated BP, dyslipidemia, or elevated ALT). Document age, sex, and pubertal stage.

Step 2 - Standard metabolic phenotyping: Fasting glucose and HbA1c (± OGTT if indicated), fasting insulin (HOMA-IR), lipids, ALT/AST (± GGT), and blood pressure; consider hs-CRP and uric acid when relevant.

Step 3 - Add endocrine context when indicated: If growth or pubertal concerns are present, measure IGF-1 (standard deviation score, SDS) and interpret metabolite findings with developmental context.

Step 4 - Trigger AA profiling (adjunct tool): Use fasting, standardized sampling for targeted AA panel (± acylcarnitines) when phenotype is metabolically unhealthy, IR is rising, dysglycemia risk is high, or NAFLD/MASLD is suspected.

Step 5 - Interpret the AA signature (puberty/sex-adjusted): Core IR-risk pattern: BCAAs (Val/Leu/Ile) increased, often AAAs (Phe/Tyr) increased, glycine decreased (± serine decreased), and glutamine-to-glutamate ratio decreased (and/or glutamate increased). Consider validated indices when available (e.g., GSG index, BCAA-based liver fat score).

Step 6 - Anchor interpretation to clinical phenotype: Normal glycemia: intensify lifestyle and monitor closely. Dysglycemia/prediabetes: follow pediatric dysglycemia pathway focusing on IR reversal. NAFLD/MASLD suspicion or ALT elevation: proceed to liver imaging (US or MRI-PDFF where available) and hepatology pathway when indicated.

Step 7 - Follow-up and response assessment: Reassess at 3-6 months (BMI-z/waist, glycemia, HOMA-IR, ALT, lipids). If AA profiling is used, repeat under identical conditions and interpret changes alongside clinical improvement.

modifiers such as pubertal stage, sex, ethnicity, fasting status, and sampling conditions.

No quantitative pooling of effect sizes was performed. Instead, we summarized whether each study reported an association, the direction of the effect, and whether the association persisted after adjustment for adiposity and other covariates when such models were reported. When studies developed composite scores or indices (for example, amino-acid-based prediction models for liver fat), we extracted the reported discrimination or predictive performance and described these findings narratively in order to maintain comparability across heterogeneous designs.

Risk of bias was assessed using RoB 2 for randomized trials (57). For non-randomized observational and interventional studies, we used ROBINS-I (58) domains covering confounding, participant selection, classification of exposures/interventions, deviations from intended interventions, missing data, outcome measurement, and selective reporting. Study-level judgments were summarized as low, moderate, or serious risk of bias and then condensed across evidence themes in order to contextualize confidence in the overall patterns described.

Results

Overview of Included Evidence

Across pediatric and youth studies, the most frequently observed AA pattern in metabolic risk states was BCAA/AAA elevation with lower glycine/serine and glutamine-glutamate imbalance, although the magnitude and specific AAs varied by phenotype definition, pubertal stage, and analytic platform (9-13). Evidence was strongest for obesity and NAFLD signatures (14-20,27-35) and increasingly supportive for dysglycemia risk signatures in youth (21-25).

Across diverse pediatric cohorts, obesity, especially metabolically higher-risk obesity, shows reproducible AA shifts centered on BCAA/AAA elevation and related intermediates, with additional signals involving glutamate, tyrosine, and BCAA-derived metabolites (e.g., 3-hydroxyisobutyrate). Longitudinal data suggest some AA changes (e.g., tyrosine) may appear early along the IR trajectory (14-26).

Pediatric dysglycemia evidence supports composite AA signatures rather than single AAs, but findings vary by developmental stage and diabetes subtype (27-31). Newer pediatric studies suggest AA panels may help capture β-cell stress risk and differentiate T1D vs T2D signatures (29,31), while systematic evidence synthesis provides interpretive scaffolding (8).

Pediatric NAFLD studies consistently implicate BCAA dysregulation and glutamate-related indices, with some approaches showing promise for non-invasive prediction of liver fat (34,36-38). Dietary AA composition may also play a role, but causal direction and pediatric intervention evidence are still limited (39-41).

AA signatures are developmentally dynamic and influenced by the GH/IGF axis. Pediatric interpretation of AA risk panels should therefore consider pubertal stage and IGF-1 status, especially when comparing across cohorts or tracking individuals longitudinally (42,43,51,52).

Pediatric intervention data suggest AA profiles, particularly BCAAs, can shift with substantial metabolic improvement (e.g., bariatric surgery), while lifestyle programs may yield broader metabolomic shifts reflecting improved cardiometabolic biology. Mechanistic adult data caution that tissue BCAA catabolism can change without large plasma BCAA shifts, and that glycine context matters for interpreting BCAA-IR relationships.

Evidence is moderate for obesity and NAFLD AA signatures because findings are consistent across cohorts and NAFLD studies often use imaging. Evidence is low-moderate for dysglycemia/diabetes and IGF-1 links due to heterogeneity (definitions, puberty effects, treatments) and fewer robust pediatric datasets. Intervention evidence is low-moderate because most studies are non-randomized, but AA signatures improve when metabolic state improves (especially after surgery). At the individual-study level, each reference in Tables I-VI inherits the risk-of-bias judgment of the evidence cluster to which it is assigned above (Ref numbers listed in the 'Studies included' column); studies which contribute only adult or mechanistic context (explicitly labeled as such in Tables I-VI). are excluded from the pediatric-evidence certainty ratings shown here.

Table I. Amino acid pattern changes in pediatric obesity and obesity phenotypes

Ref	First author (Year)	Design/population	Platform	Key amino-acid findings (direction)	Metabolic association (IR/risk phenotype)
(14)	McCormack (2013)	Cohort; children/adolescents (n=69, ages 8-18 y cross-sectional; longitudinal subset n=17, ages 8-13 y, followed 18 months)	Targeted AAs	↑BCAAs/AAAs in obesity	Baseline BCAA pattern associated with future insulin resistance
(15)	Butte (2015)	Cross-sectional; Hispanic children (n=803; mean age 11.1±3.9y; 56% with obesity)	Untargeted metabolomics	AA clusters (incl. BCAA/AAA) differed by obesity	AA PCs associated with hyperinsulinemia and TG risk
(16)	Mangge (2016)	Cross-sectional; lean/overweight/obese youth [cohort spanned adults and juveniles, mean age 25.3±12.8y; pediatric-relevant subset only]	Targeted BCAAs	↑BCAAs across adiposity strata	BCAAs associated with cardiometabolic risk profiles
(17)	Mastrangelo (2016)	Case-control; obese prepubertal	Targeted metabolomics	Sex-dependent AA shifts; ↑BCAA/AAA patterns in IR	Metabolomic profile correlated with IR severity
(18)	Martos-Moreno (2017)	Targeted metabolomics; obese prepubertal	Targeted metabolomics	Discriminative metabolome in hyperinsulinism	Metabolome separated hyperinsulinemic vs less-affected phenotypes
(19)	Suzuki (2019)	Cross-sectional; childhood obesity (n=26; ages 9-10y)	AA + lipid panels	AA profile linked to IGT; BCAA/AAA signals prominent	AA profile reflected early IGT and HOMA-IR variation
(20)	Hellmuth (2016)	Longitudinal; obese children	Longitudinal metabolomics	Tyrosine tracked IR trajectory; later BCAA shifts	Suggests early AA changes may precede broader BCAA disruption
(21)	Bugajska (2023)	Cross-sectional; overweight/obese prepubertal	AA profiling	Abnormal AA profile in OW/OB	Linked with early markers of cardiometabolic risk and MAFLD signals
(22)	Moran-Ramos (2017)	Cross-sectional + longitudinal school-age cohort	AA signature score	AA signature (incl. BCAA/AAA/proline/arginine)	Predicted 2-year hypertriglyceridemia risk
(23)	Perng (2020)	Cross-sectional adolescents (Project Viva)	Metabolomics	Metabolomic differences across OWOB-MetRisk phenotypes	MetRisk phenotype associated with AA differences incl. BCAA/AAA
(24)	Short (2019)	Cross-sectional + exercise intervention; American Indian adolescents	42 AA panel	Obesity: ↑BCAA/AAA/glutamate; ↓BAIBA; ↑2-AAA	AA levels correlated with adiposity and inversely with insulin sensitivity
(25)	McCann (2021)	Cohort biorepository (POMMS)	Metabolomics/microbiome	Adolescents with obesity show metabolomic patterns overlapping adult risk	Provides pediatric baseline metabolomic context for obesity interventions
(26)	Ng (2025)	Untargeted metabolomics; MHO vs MUO	Metabolomics	MUO linked to AA/FA differences; 3-HIB and BCAA signals	Differences primarily aligned with abnormal glucose homeostasis

AA: Amino acid(s), BCAA: Branched-chain amino acid(s), AAA: Aromatic amino acid(s), IR: Insulin resistance, TG: Triglycerides, IGT: Impaired glucose tolerance, HOMA-IR: Homeostasis model assessment of insulin resistance, MUO: Metabolically unhealthy obesity, MHO: Metabolically healthy obesity, 3-HIB: 3-hydroxyisobutyrate, BAIBA: β-aminoisobutyric acid, 2-AAA: 2-aminoadipic acid

Across these studies, BCAA/AAA elevation with reduced glycine is the most consistent amino-acid signature distinguishing metabolically unhealthy from metabolically healthy obesity.

This pattern is detectable in cross-sectional pediatric cohorts and, in longitudinal data, appears to precede measurable insulin resistance.

Table II. Amino acid patterns in pediatric prediabetes, dysglycemia risk, and diabetes phenotypes

Ref	First author (Year)	Phenotype/population	Platform	Key amino-acid findings	Clinical implication
(8)	Guasch-Ferré (2016)	Evidence synthesis (human studies)	Review/meta-analysis	AA panels (BCAA/AAA ↑; glycine ↓; glutamine-glutamate imbalance) linked to prediabetes/T2D risk	Framework for interpreting AA panels as risk markers
(27)	Michaliszyn (2012)	Obese adolescents ± dysglycemia	Targeted AAs	AA concentrations related to β-cell function relative to insulin sensitivity	Highlights developmental differences vs adults in AA-risk interpretation
(28)	Mihalik (2012)	NW vs obese vs T2D adolescents	Acylcarnitine + AA profiling	T2D youth showed distinct AA/acylcarnitine patterns vs obese/NW	Suggests youth T2D metabolome differs from adult paradigm
(19)	Suzuki (2019)	Childhood obesity with early metabolic abnormalities	AA + lipids	AA profile reflected impaired glucose tolerance early	Supports AA panels for early dysglycemia risk stratification in obesity
(24)	Short (2019)	Obese vs NW adolescents	Targeted AAs	↑BCAA/AAA/glutamate; ↓BAIBA; ↑2-AAA	Candidate AA markers for insulin sensitivity in adolescents
(26)	Ng (2025)	Obesity with abnormal glucose tolerance	Untargeted metabolomics	↑BCAA and 3-HIB; ↓1,5-Anhydroglucitol in abnormal glucose tolerance	Links AA changes to glycemic dysregulation phenotype
(29)	Bacha (2024)	Youth with or at risk for T2D	Targeted AAs	Higher BCAA/AAA and lysine; lower glycine/glutamine-glutamate signals	AA profile proposed to identify risk for β-cell failure
(30)	Perng (2022)	Two youth cohorts; fasting glucose outcomes	Metabolomics + RRR	Sex-specific metabolite predictors of future dysglycemia	Demonstrates predictive modeling using metabolite panels including AAs
(31)	Tosur (2023)	Pediatric T1D vs T2D	Targeted AA metabolomics	Distinct fasting AA signatures by diabetes type	Potential biomarker support for classification/phenotyping

AA: Amino acid(s), BCAA: Branched-chain amino acid(s), AAA: Aromatic amino acid(s), T1D: Type 1 diabetes, T2D: Type 2 diabetes, RRR: Reduced rank regression

Amino-acid signatures differ by diabetes type, with distinct fasting profiles distinguishing pediatric type 1 from type 2 diabetes.

Glycine reduction and BCAA/AAA elevation track with declining insulin sensitivity across the dysglycemia spectrum, supporting their use as early risk markers.

Table III. Amino acid signatures in pediatric NAFLD/MASLD and related fatty liver disorders

Ref	First author (Year)	Study/population	Steatosis assessment	Key amino-acid findings	Clinical signal
(32)	Vos (2017)	NASPGHAN guideline	Guideline	Recognizes NAFLD as obesity/IR-linked	Supports need for better non-invasive biomarkers
(33)	Jin (2016)	Obese adolescents; case-control	MRS	AA metabolism pathways disturbed in NAFLD	Early evidence that AA pathways integral to pediatric NAFLD
(34)	Goffredo (2017)	Obese adolescents with NAFLD (n=78; mean age 13.3±3.0y, range 8-18y)	Imaging + follow-up	BCAA-related metabolomic signature in NAFLD	Signature associated with NAFLD independent of obesity/IR
(35)	Tricò (2018)	Multiethnic obese adolescents	Imaging + metabolic phenotyping	NAFL associated with metabolic features including AA signals	Highlights ethnicity/race and IR interplay in NAFL risk
(36)	Lischka (2021)	Youth with severe obesity	Liver fat prediction	BCAA-based metabolic score predicted liver fat	Practical candidate-biomarker approach for risk stratification
(37)	Leonetti (2020)	Pediatric at-risk cohorts	NAFLD detection	'GSG index' (glutamate-serine-glycine) detected NAFLD	Index associated with NAFLD independent of traditional risks
(38)	Chae (2022)	Overweight pediatric population	Biopsy-proven NAFLD	NAFLD-specific metabolic features included valine/tyrosine/glutamate/glycine	Machine-learning models achieved high AUROC using metabolite features

Table III. Continued

Ref	First author (Year)	Study/population	Steatosis assessment	Key amino-acid findings	Clinical signal
(39)	Nikparast (2024)	Overweight/obese children	Ultrasound + clinician dx	Higher dietary BCAA (esp. leucine) associated with NAFLD odds	Suggests dietary AA composition may modulate NAFLD risk
(40)	Lake (2015)	Progressive human NAFLD [Adult; mechanistic context only] (context)	Histology spectrum	Hepatic BCAA metabolite changes with NAFLD→NASH progression	Mechanistic support for BCAA pathway involvement
(41)	Deng (2024)	Review (MAFLD mechanisms) [Not pediatric-specific; mechanistic context only]	Synthesis	Essential AA signaling intersects lipid handling, inflammation	Frames therapeutic interest but pediatric trials remain limited

AA: Amino acid(s), BCAA: Branched-chain amino acid(s), NAFLD: Non-alcoholic fatty liver disease, MASLD: Metabolic dysfunction-associated steatotic liver disease, NASH: Non-alcoholic steatohepatitis, IR: Insulin resistance, MRS: Magnetic resonance spectroscopy, AUROC: Area under the receiver operating characteristic curve, GSG index: Glutamate-serine-glycine index

BCAA-related amino-acid alterations associate with hepatic steatosis independently of BMI and conventional risk markers in several pediatric cohorts.

Amino-acid-based indices show early promise as noninvasive adjuncts to imaging for pediatric NAFLD/MASLD risk stratification, though external validation remains limited.

Table IV. Amino acids, IGF-1 axis, and developmental modifiers relevant to pediatric interpretation

Ref	First author (Year)	Population/context	Key observation	Importance for this review
(42)	Jensch (2024)	Healthy children/adolescents	Whole-blood AA/acylcarnitines associated with anthropometry and IGF-I	Supports developmental and endocrine coupling of AA patterns with IGF-I
(43)	Hirschel (2020)	Children/adolescents	AA/acylcarnitine levels vary with age, sex, BMI, puberty and correlate with metabolic markers	Puberty and sex must be adjusted when interpreting AA signatures
(44)	Knacke (2016)	Metabolomics + IGF-1 (multifluid; context) [Adult/general population; mechanistic context only]	IGF-1 associated with metabolites linked to AA metabolism (e.g., pipecolate)	Provides mechanistic plausibility for IGF-linked AA fingerprints
(45)	Barazani (2020)	IGF-I deficiency (Laron syndrome) ± IGF-I therapy	IGF-I deficiency associated with abnormal plasma AA metabolism partially reversed by IGF-I	Direct evidence that IGF-I status can shape AA profiles
(46)	Aggelaki (2025)	Children with GHD before/after GH therapy	Metabolomic fingerprint differs in GHD; correlations explored with IGF-1	Reinforces endocrine context for metabolomic interpretation
(47)	Inoue (2025)	SGA infants (cord metabolomics)	Cord AA alterations associated with impaired linear growth	Demonstrates early-life AA-growth links relevant to IGF biology

AA: Amino acid(s), IGF-I/IGF-1: Insulin-like growth factor-1, GH: Growth hormone, GHD: Growth hormone deficiency, SGA: Small for gestational age

Pubertal stage, sex, and IGF-1 status are important modifiers of amino-acid profiles and must be accounted for when interpreting AA signatures in children.

Conditions altering the IGF-1 axis (growth hormone deficiency, IGF-1 deficiency, small-for-gestational-age status) are associated with distinct amino-acid fingerprints, supporting a mechanistic link between growth endocrinology and AA metabolism.

Table V. Management and intervention studies: how amino-acid patterns shift with treatment and their importance for this review

Ref	First author (Year)	Intervention/design	Population	Key AA-related findings	Clinical takeaway
(48)	Hampl (2023)	Pediatric obesity guideline	Children/adolescents	Lifestyle-first; pharmacotherapy/ surgery for selected cases	AA biomarkers not routine yet; focus remains on weight/IR management
(49)	Rigamonti (2023)	Short-term body weight reduction program (prospective)	Adolescents with obesity	Program altered metabolomic profile toward improved cardiometabolic pathways	Supports metabolomics as response-tracking tool in structured programs
(50)	Becetti (2023)	Sleeve gastrectomy vs non-surgical (12 months)	Youth 13-24 yrs	Post-surgery BCAA reductions associated with improved IR	Bariatric surgery may normalize BCAA burden alongside IR improvement
(20)	Hellmuth (2016)	Longitudinal observational (weight/IR trajectory)	Obese children	Tyrosine alterations associated with IR preceded broader BCAA changes	Early AA markers may help identify high-risk trajectories
(51)	Glynn (2015)	Aerobic + resistance training (exercise intervention; context)	Overweight adults *[Adult; mechanistic context only]*	Exercise improved IS; glycine conjugation pathways implicated	Mechanistic insight: AA handling changes may mediate IS improvements
(52)	Lee (2021)	Exercise training (context)	Dysglycemic men *[Adult; mechanistic context only]*	IS improved with enhanced tissue BCAA catabolism despite stable plasma BCAA	Plasma BCAA may not fully reflect tissue catabolism changes
(53)	Lim (2024)	Weight-loss RCT secondary analysis (context)	Prediabetes adults *[Adult; mechanistic context only]*	Low glycine strengthened BCAA-IR association; concurrent assessment improved interpretation	Supports combined glycine + BCAA measurement for IR phenotyping

AA: Amino acid(s), BCAA: Branched-chain amino acid(s), IR: Insulin resistance, IS: Insulin sensitivity, RCT: Randomized controlled trial

Pediatric weight-loss interventions, particularly bariatric surgery, are associated with measurable reductions in BCAA levels alongside improved insulin resistance.

Adult mechanistic studies suggest amino-acid catabolism can shift with exercise training before changes are detectable in plasma BCAA, underscoring the need for pediatric-specific intervention trials.

Table VI. Quality assessment summary table (Cochrane-aligned)

Evidence cluster	Studies included from Tables I-V (Ref no.)	Typical study designs	Key bias concerns (ROBINS-I/RoB 2 domains)	Overall certainty (narrative)
Pediatric obesity AA signatures (Table I)	(14-26)	Mostly cross-sectional and cohort; some longitudinal and phenotype-stratified cohorts	Confounding (puberty stage, diet/protein source, ethnicity, adiposity distribution); selection bias (clinic-based cohorts); measurement variability (fasting status, assay/platform/batch effects); multiple testing/selective reporting in high-dimensional metabolomics	Moderate (directional consistency across cohorts; heterogeneity and residual confounding remain)
Pediatric dysglycemia/ diabetes AA signatures (Table II)	(8,19,24,26-31)	Case-control and cohort (often small); some studies with gold-standard physiology/phenotyping; diabetes subtype comparisons	Phenotype misclassification (definitions of dysglycemia; T1D vs T2D overlap); treatment effects and disease duration; puberty confounding; small samples/selection bias; reverse causality (cross-sectional designs)	Low-moderate (signals emerging; strongest where detailed phenotyping is present; heterogeneity remains)
Pediatric NAFLD/ MASLD AA signatures (Table III)	(32-41)	Case-control/cohort with imaging (MRI/MRS/US) ± metabolic phenotyping; prediction models/indices	Outcome ascertainment heterogeneity (US vs MRI vs MRS vs biopsy); confounding by obesity/IR/lifestyle; selection toward tertiary/ severe obesity cohorts; model overfitting and limited external validation for indices/ML models	Moderate (repeatability of BCAA/glutamate-related patterns; broader validation still needed)

Table VI. Continued

Evidence cluster	Studies included from Tables I-V (Ref no.)	Typical study designs	Key bias concerns (ROBINS-I/RoB 2 domains)	Overall certainty (narrative)
IGF-1/ development modifiers (Table IV)	(42-47)	Large population cohorts plus special endocrine/ perinatal states	Developmental confounding (age, sex, puberty staging); generalizability limits for rare endocrine states; assay/platform differences; within-person variability across growth	Low-moderate (strong evidence that puberty/ sex/adiposity modify AA distributions; fewer disease-state datasets directly linking AA patterns to IGF-1 outcomes)
Intervention effects (Table V)	Pediatric interventions/ guideline: (20,48-50); adult mechanistic context: (51-53)	Prospective lifestyle programs and surgery cohorts; guideline framing; adult mechanistic trials for interpretation	Non-randomized designs (selection bias), adherence/co-interventions; short follow-up for lifestyle; small samples; outcome changes may lag metabolite shifts	Low-moderate (clear biologic shifts when metabolic state improves substantially—especially surgery; limited pediatric randomized metabolomics evidence)

RoB 2: Cochrane risk of bias tool version 2, ROBINS-I: Risk of bias in non-randomized studies of interventions, AA: Amino acid(s), IR: Insulin resistance, NAFLD: Non-alcoholic fatty liver disease, MASLD: Metabolic dysfunction-associated steatotic liver disease, IGF-1: Insulin-like growth factor-1, MRI: Magnetic resonance imaging, MRS: Magnetic resonance spectroscopy, US: Ultrasound

Most included pediatric evidence derives from cross-sectional or case-control designs, supporting moderate (and in some clusters low-to-moderate) certainty for the amino-acid signatures described.

Longitudinal and interventional pediatric studies remain comparatively scarce, representing the key evidence gap for future confirmatory research.

Algorithm Text (Figure 2)

The algorithm summarizes a stepwise clinical-research pathway for children and adolescents with overweight/ obesity or metabolic risk. Baseline phenotyping includes anthropometrics and pubertal staging, followed by standard metabolic evaluation (fasting glucose/HbA1c ± OGTT when indicated, fasting insulin/HOMA-IR, lipids, and liver enzymes). IGF-1 (standard deviation score, SDS) is incorporated when growth or pubertal concerns are present to contextualize developmental endocrine influences on metabolite patterns. Amino-acid (AA) profiling (fasting, standardized sampling; targeted AA panel ± acylcarnitines) is positioned as an adjunct tool for advanced risk stratification in metabolically unhealthy obesity, rising insulin resistance, dysglycemia risk, or suspected NAFLD/MASLD. Interpretation is anchored to phenotype (normal glycemia vs dysglycemia vs NAFLD suspicion) to guide escalation of guideline-based management and liver imaging when indicated. Follow-up at 3-6 months prioritizes clinical outcomes (BMI-z/ waist, glycemia, HOMA-IR/insulin sensitivity markers, liver enzymes, lipids); repeat AA profiling, when used, should be performed under identical conditions and interpreted alongside metabolic improvement. Throughout this pathway, AA profiling is intended strictly as an adjunct to standard clinical, biochemical, and imaging assessment rather than as a replacement for it; the algorithm reflects a proposed research-informed framework, and prospective validation in diverse pediatric populations is required before broader clinical implementation.

Graphic Abstract (Figure 3)

Provides graphical synthesis of this composite amino-acid signature—BCAA/AAA elevation with glycine reduction and glutamine–glutamate imbalance—and summarizes how these candidate biomarker signatures intersect with standard clinical phenotyping, puberty/IGF-1 context, and downstream risk for prediabetes/T2D and NAFLD/MASLD across the pediatric evidence base described in Tables I-V.

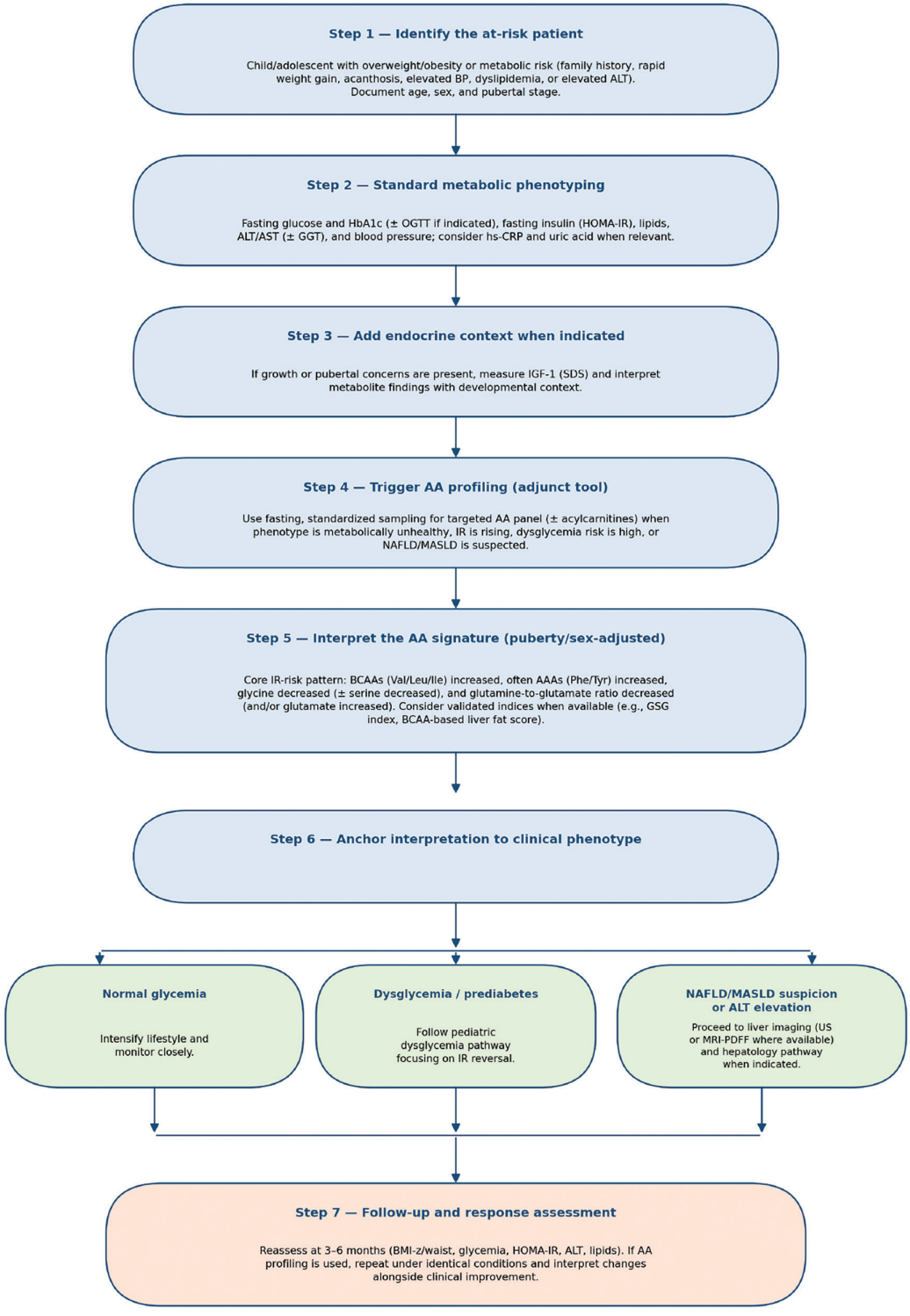

Figure 2. Algorithmic framework for interpreting amino-acid pattern signatures across pediatric obesity, dysglycemia, and fatty liver disease, incorporating insulin resistance and IGF-1 context. This algorithm is intended as an adjunct to, not a replacement for, standard clinical assessment; amino-acid profiling requires further prospective validation before broader clinical implementation.

Figure 2. Algorithmic framework for interpreting amino-acid pattern signatures across pediatric obesity, dysglycemia, and fatty liver disease, incorporating insulin resistance and IGF-1 context. This algorithm is intended as an adjunct to, not a replacement for, standard clinical assessment; amino-acid profiling requires further prospective validation before broader clinical implementation

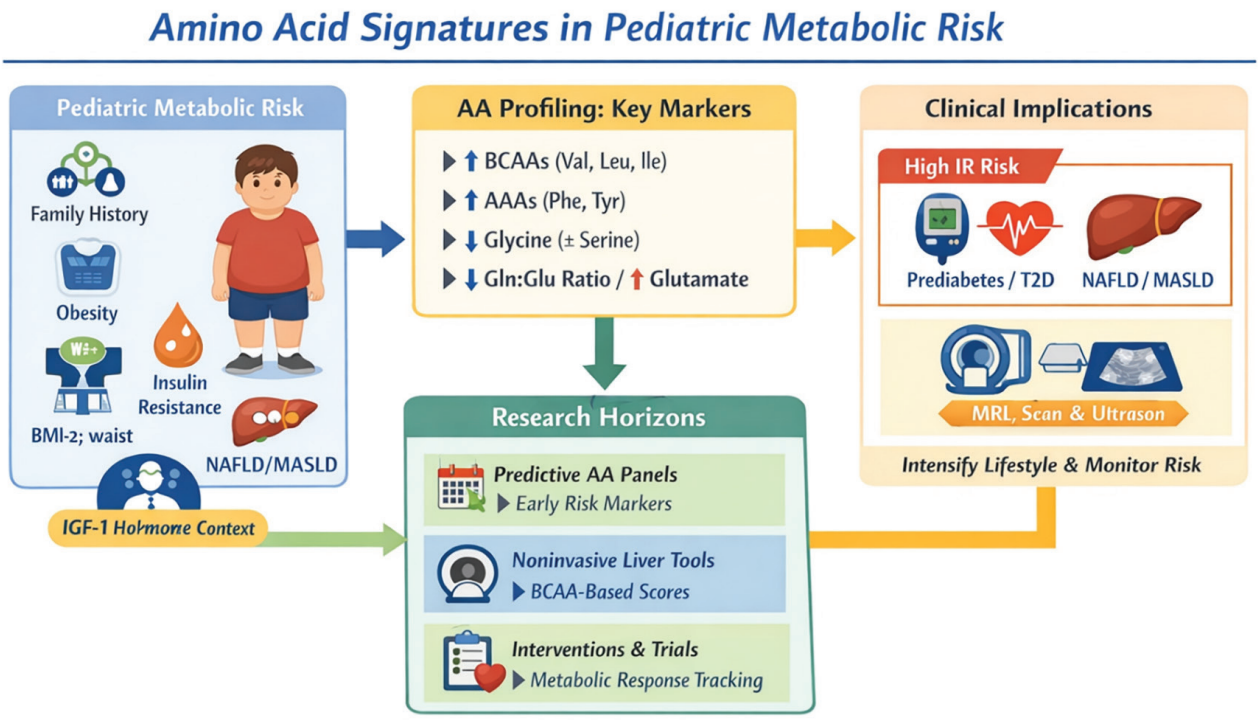

Figure 3. Amino acid signatures in pediatric metabolic risk: linking obesity to insulin resistance, dysglycemia, and fatty liver disease. This graphical abstract summarizes the key amino-acid pattern associated with pediatric metabolic risk (candidate biomarker signatures: BCAA/AAA elevation with glycine reduction and glutamine-glutamate imbalance) and shows how these signatures integrate with standard clinical phenotyping in order to identify children at higher risk of prediabetes/T2D and NAFLD/MASLD, while emphasizing puberty/IGF-1 context and the need for prospective validation before these candidate signatures can inform intervention trials

Discussion

Across pediatric studies, the most reproducible amino-acid (AA) fingerprint linked to metabolic risk is an increase in branched-chain amino acids (BCAAs) and aromatic amino acids (AAAs), frequently accompanied by higher glutamate and lower glycine/serine, with substantial modulation by age, pubertal stage, sex, and adiposity distribution (9-13,42,43). These patterns are clinically meaningful rather than purely descriptive as multiple pediatric cohorts show associations with insulin resistance (IR) and cardiometabolic risk clustering, indicating that AA profiles can serve as integrated readouts of disrupted nutrient handling in pediatric obesity (14-16,19,24).

Mechanistic interpretations converge on altered BCAA catabolism and downstream metabolic 'spillover.' Elevated circulating BCAAs may reflect reduced activity or regulation of catabolic enzymes and altered tissue partitioning, while BCAA-derived intermediates can mirror mitochondrial substrate overload and incomplete oxidation (2-5,13). In parallel, recurring reductions in glycine/serine may reflect constrained one-carbon and antioxidant capacity and limited acyl-group buffering, while glutamine-glutamate imbalance may indicate altered transamination and hepatic nitrogen flux (6,7). Evidence synthesis supports the concept that composite AA panels (BCAA/AAA elevation with glycine reduction and glutamine-glutamate imbalance) are among the most consistent metabolite features associated with prediabetes/diabetes risk, providing an interpretive scaffold relevant to youth as pediatric data mature (8).

A key pediatric nuance is that metabolic risk is not uniform within obesity, and AA signatures appear most informative when aligned to phenotype rather than BMI alone. Metabolomic comparisons across adolescent overweight/obesity phenotypes indicate that AA differences cluster with 'metabolically at-risk' profiles (e.g., adverse glycemia and lipid patterns) rather than with adiposity per se (23). Similarly, metabolomic differences between metabolically healthy obesity and metabolically unhealthy obesity are driven primarily by abnormal glucose homeostasis and related metabolic disturbances, supporting AA profiling as a potential adjunct to phenotype anchoring (26).

Longitudinal pediatric findings add strength by moving beyond cross-sectional discrimination toward prediction. Baseline BCAA elevations have been associated with future

insulin resistance in children and adolescents, supporting their potential as early-risk markers (14). Longitudinal metabolomic profiling also suggests that certain signals, particularly tyrosine, may track IR trajectories and could change earlier than broader BCAA disturbances in some pediatric obesity cohorts (20). In addition, AA-based signatures in school-age children have predicted future hypertriglyceridemia, reinforcing the concept that AA patterns can act as early 'stress integrators' across metabolic pathways (22). Collectively, these studies argue that the most clinically relevant next step is prospective validation of AA signatures for dysglycemia and progression to fatty liver disease.

For dysglycemia and pediatric diabetes phenotypes, the evidence is evolving and heterogeneous, reflecting developmental physiology, pubertal insulin resistance, differing diabetes subtypes, and treatment effects. Clamp-based pediatric studies indicate that AA-β-cell relationships can differ from adult paradigms and may reflect compensatory physiology early in the disease trajectory (27). Youth-onset Type 2 diabetes has also shown distinct AA/acylcarnitine patterns compared with obesity alone, suggesting that pediatric T2D metabolomic features cannot be assumed to mirror adult profiles (28). More recent youth-focused targeted profiling supports a 'risk-like' AA pattern (higher BCAA/AAA and lysine with lower glycine and glutamine-glutamate signals) in youth with or at risk of Type 2 diabetes, consistent with impaired insulin sensitivity and potential β-cell stress (29). Prospective modeling in youth cohorts further indicates that metabolite panels including AAs can predict dysglycemia, with sex-specific signals which underscore the need for developmentally stratified interpretation (30). Finally, targeted AA signatures distinguishing pediatric diabetes types suggest potential future utility for phenotype refinement in research and specialized clinical settings (31).

Pediatric NAFLD/MASLD shows particularly strong convergence of AA disturbances across untargeted pathway analyses, targeted metabolite panels, and prediction tools. Studies in obese adolescents demonstrate AA pathway disruption in NAFLD, including BCAA-related signatures and broader perturbations which align with hepatic fat and insulin resistance phenotypes (33-35). In youth with severe obesity, a BCAA-based metabolic score has been proposed to predict liver fat, offering a practical approach to non-invasive risk stratification which may be especially valuable where imaging access is constrained (36). Complementary approaches such as the glutamate-serine-glycine (GSG) index further highlight the recurring role of glutamate and glycine/serine biology in pediatric fatty liver phenotypes (37). In biopsy-proven pediatric NAFLD, models incorporating valine, tyrosine, glutamate, and glycine have achieved strong discrimination, but require external validation and pragmatic evaluation before clinical translation (38). Current pediatric NAFLD guidelines emphasize early identification and risk stratification but acknowledge the need for better non-invasive biomarkers, providing a clear translational rationale for continued AA biomarker development (32).

Dietary AA composition is a plausible contributor to circulating AA signatures and liver-fat risk, but causal evidence in children remains limited and confounded by total energy intake, protein source, and dietary pattern quality. Observational pediatric work suggests that higher dietary BCAA intake, particularly leucine, is associated with NAFLD odds in overweight/obese youth, but these findings require cautious interpretation and prospective validation (39). Mechanistic context from progressive NAFLD studies supports BCAA pathway remodeling across disease severity and highlights intersections with hepatic inflammation and lipid handling (40). Recent syntheses also emphasize that essential AAs can influence metabolic liver disease biology through nutrient signaling pathways, but pediatric interventional trials specifically targeting AA composition while preserving growth needs remain a priority (41).

The IGF-1 axis adds an endocrine-growth lens to AA interpretation in pediatrics. Large healthy pediatric cohorts demonstrate that whole-blood AAs and acylcarnitines vary systematically with age, sex, BMI, and pubertal stage, and can associate with IGF-1 levels and growth-related phenotypes, reinforcing the need for puberty- and sex-aware reference frameworks (42,43). Mechanistic possibility is supported by metabolomic analyses linking circulating IGF-1 and IGF-1/IGFBP-3 ratios to AA-related metabolites (44). Moreover, endocrine states characterized by altered IGF-1 or GH biology provide direct evidence that hormonal status can shape AA metabolism: long-term IGF-1 deficiency shows abnormal plasma AA metabolism which is partially reversible with IGF-1 therapy, and GH deficiency cohorts demonstrate distinct metabolomic fingerprints with IGF-1-related correlations (45,46). Early-life metabolomic data in growth-restricted infants further support the concept that AA availability and handling relate to linear growth trajectories, strengthening the rationale to integrate IGF-1 context when interpreting AA risk panels across pediatric ages (47).

From a clinical perspective, current pediatric guidelines for obesity, NAFLD, and youth-onset Type 2 diabetes remain anchored in lifestyle intervention, weight management, and cardiometabolic risk control, with imaging/laboratory assessment for NAFLD rather than AA biomarkers (32,48,54,55). This is appropriate given the present heterogeneity in platforms, the lack of standardized pediatric cutoffs, and limited evidence that AA-informed stratification improves outcomes. Nevertheless, AA profiling is positioned for near-term research-to-practice transition in three areas: identifying metabolically higher-risk obesity phenotypes, enriching screening for NAFLD risk when imaging is limited, and monitoring metabolic response in intensive programs or post-surgical settings.

Intervention evidence supports partial 'reversibility' of AA disturbances when metabolic states improve meaningfully. In adolescents with obesity, short-term structured weight-reduction programs can shift metabolomic profiles toward more favorable cardiometabolic patterns, supporting metabolomics as a potential response-tracking tool (49). In youth undergoing sleeve gastrectomy, reductions in BCAAs have been reported in parallel with improvements in insulin resistance, suggesting that major weight loss and metabolic remodeling can normalize at least part of the BCAA-related signature (50). Longitudinal observational data also indicate that specific AAs such as tyrosine track insulin resistance trajectories in children, highlighting candidate markers for monitoring high-risk courses (20). Mechanistic adult intervention studies provide context that improved insulin sensitivity may be mediated by enhanced handling of BCAA-derived acyl groups and glycine conjugation and that tissue BCAA catabolism can improve even when plasma BCAA changes are modest, concepts which may guide pediatric trial design without being assumed as directly generalizable to children (51-53).

Key research gaps now center on standardization and validation. Pediatric studies need harmonized sampling (fasting status, timing relative to puberty), consistent reporting of pubertal stages and body composition, and cross-platform analytic comparability, as highlighted by pediatric metabolomics syntheses (9,10). Longitudinal cohorts are needed in order to test whether AA signatures add predictive value beyond established clinical risk factors for incident dysglycemia and NAFLD progression, and to define whether AA-informed phenotyping can guide interventions without compromising growth and nutritional adequacy. Finally, pragmatic studies should evaluate whether incorporating AA panels into clinical pathways improves risk stratification, patient engagement, or response monitoring compared with guideline-based care alone (32,48,54,55).

Overall, pediatric literature supports AA signatures as biologically coherent, developmentally modulated candidate biomarkers aligned with insulin resistance, dysglycemia risk, and fatty liver burden. This field is now positioned to move from association to action by establishing pediatric reference frameworks, validating prediction tools in diverse populations, and embedding AA profiling into interventional trials designed around clinically meaningful endpoints and growth-safe nutrition.

Conclusion

Pediatric and adolescent metabolic disorders, namely obesity, dysglycemia/diabetes risk, and NAFLD/MASLD, share partially overlapping amino-acid fingerprints, most commonly involving elevated BCAAs and AAAs, glutamate-related shifts, and reduced glycine/serine in higher-risk phenotypes. These patterns align with insulin resistance biology and can associate with liver fat burden and adverse cardiometabolic trajectories. However, AA signatures are strongly influenced by age, sex, puberty, and IGF-1/GH axis status, mandating pediatric-specific interpretive frameworks. While the current clinical management remains guideline-based, AA profiling is poised to become a valuable adjunct for research-grade risk stratification and response monitoring once standardized, validated, and tested for incremental clinical utility.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.S., H.Z., F.A., N.A., S.A., N.H., S.E., Concept: A.S., H.Z., F.A., N.A., Design: A.S., H.Z., Data Collection or Processing: A.S., H.Z., N.A., S.A., N.H., S.E., Analysis or Interpretation: A.S., F.A., Literature Search: A.S., F.A., N.A., S.A., Writing: A.S., H.Z., N.H., S.E.

Conflict of Interest: The authors declare no conflicts of interest.

Financial Disclosure: The authors received no financial support for the conduct, authorship, or publication of this study.

References

1. Ling ZN, Jiang YF, Ru JN, Lu JH, Ding B, Wu J. Amino acid metabolism in health and disease. Sig Transduct Target Ther. 2023; 8:345.
2. White PJ, McGarrah RW, Herman MA, Bain JR, Shah SH, Newgard CB. Insulin action, type 2 diabetes, and branched-chain amino acids: A two-way street. Mol Metab. 2021; 52:101261.

3. Lynch CJ, Adams SH. Branched-chain amino acids in metabolic signalling and insulin resistance. Nat Rev Endocrinol. 2014; 10:723-36.

4. De Bandt JP, Coumoul X, Barouki R. Branched-chain amino acids and insulin resistance, from protein supply to diet-induced obesity. Nutrients. 2022; 15:68.

5. Yoon MS. The emerging role of branched-chain amino acids in insulin resistance and metabolism. Nutrients. 2016; 8:405.

6. Adeva-Andany M, Souto-Adeva G, Ameneiros-Rodríguez E, Fernández-Fernández C, Donapetry-García C, Domínguez-Montero A. Insulin resistance and glycine metabolism in humans. Amino Acids. 2018; 50:11-27.

7. Alves A, Bassot A, Bulteau AL, Pirola L, Morio B. Glycine metabolism and its alterations in obesity and metabolic diseases. Nutrients. 2019; 11:1356.

8. Guasch-Ferré M, Hruby A, Toledo E, et al. Metabolomics in prediabetes and diabetes: a systematic review and meta-analysis. Diabetes Care. 2016; 39:833-46.

9. Handakas E, Lau CH, Alfano R, et al. A systematic review of metabolomic studies of childhood obesity: state of the evidence for metabolic determinants and consequences. Obes Rev. 2022; 23(Suppl 1):e13384.

10. De Spiegeleer M, De Paepe E, Van Meulebroek L, Gies I, De Schepper J, Vanhaecke L. Paediatric obesity: a systematic review and pathway mapping of metabolic alterations underlying early disease processes. Mol Med. 2021; 27:145.

11. Polidori N, Grasso EA, Chiarelli F, Giannini C. Amino acid-related metabolic signature in obese children and adolescents. Nutrients. 2022; 14:1454.

12. Zamosteanu D, Filip N, Trandafir LM, et al. Current data on the role of amino acids in the management of obesity in children and adolescents. Int J Mol Sci. 2025; 26:7129.

13. Dimou A, Tsimihodimos V, Bairaktari E. The critical role of the branched chain amino acids (BCAAs) catabolism-regulating enzymes, branched-chain aminotransferase (BCAT) and branched-chain α-keto acid dehydrogenase (BCKD), in human pathophysiology. Int J Mol Sci. 2022; 23:4022.

14. McCormack SE, Shaham O, McCarthy MA, et al. Circulating branched-chain amino acid concentrations are associated with obesity and future insulin resistance in children and adolescents. Pediatr Obes. 2013; 8:52-61.

15. Butte NF, Liu Y, Zakeri IF, et al. Global metabolomic profiling targeting childhood obesity in the Hispanic population. Am J Clin Nutr. 2015; 102:256-67.

16. Mangge H, Zelzer S, Prüller F, et al. Branched-chain amino acids are associated with cardiometabolic risk profiles found already in lean, overweight and obese young. J Nutr Biochem. 2016; 32:123-7.

17. Mastrangelo A, Martos-Moreno GÁ, García A, et al. Insulin resistance in prepubertal obese children correlates with sex-dependent early onset metabolomic alterations. Int J Obes (Lond). 2016; 40:1494-502.

18. Martos-Moreno GÁ, Mastrangelo A, Barrios V, et al. Metabolomics allows the discrimination of the pathophysiological relevance of hyperinsulinism in obese prepubertal children. Int J Obes (Lond). 2017; 41:1473-80.

19. Suzuki Y, Kido J, Matsumoto S, Shimizu K, Nakamura K. Associations among amino acid, lipid, and glucose metabolic profiles in childhood obesity. BMC Pediatr. 2019; 19:273.

20. Hellmuth C, Kirchberg FF, Lass N, et al. Tyrosine is associated with insulin resistance in longitudinal metabolomic profiling of obese children. J Diabetes Res. 2016; 2016:2108909.

21. Bugajska J, Berska J, Wójcik M, Sztefko K. Amino acid profile in overweight and obese prepubertal children - can simple biochemical tests help in the early prevention of associated comorbidities?. Front Endocrinol (Lausanne). 2023; 14:1274011.

22. Moran-Ramos S, Ocampo-Medina E, Gutierrez-Aguilar R, et al. An amino acid signature associated with obesity predicts 2-year risk of hypertriglyceridemia in school-age children. Sci Rep. 2017; 7:5607.

23. Perng W, Rifas-Shiman SL, Sordillo J, Hivert MF, Oken E. Metabolomic profiles of overweight/obesity phenotypes during adolescence: a cross-sectional study in project viva. Obesity (Silver Spring). 2020; 28:379-87.

24. Short KR, Chadwick JQ, Teague AM, et al. Effect of obesity and exercise training on plasma amino acids and amino metabolites in American Indian adolescents. J Clin Endocrinol Metab. 2019; 104:3249-61.

25. McCann JR, Bihlmeyer NA, Roche K, et al. The pediatric obesity microbiome and metabolism study (POMMS): methods, baseline data, and early insights. Obesity (Silver Spring). 2021; 29:569-78.

26. Ng DZW, Shie YH, Piyasanka SR, et al. Metabolome alterations in pediatric metabolically unhealthy obesity are primarily linked to abnormal glucose homeostasis. Sci Rep. 2025; 15:23934.

27. Michaliszyn SF, Sjaarda LA, Mihalik SJ, et al. Metabolomic profiling of amino acids and β-cell function relative to insulin sensitivity in youth. J Clin Endocrinol Metab. 2012; 97:E2119-2124.

28. Mihalik SJ, Michaliszyn SF, de las Heras J, et al. Metabolomic profiling of fatty acid and amino acid metabolism in youth with obesity and type 2 diabetes: evidence for enhanced mitochondrial oxidation. Diabetes Care. 2012; 35:605-11.

29. Bacha F, El-Ayash H, Mohamad M, et al. Distinct amino acid profile characterizes youth with or at risk for type 2 diabetes. Diabetes. 2024; 73:628-36.

30. Perng W, Hivert MF, Michelotti G, Oken E, Dabelea D. Metabolomic predictors of dysglycemia in two U.S. youth cohorts. Metabolites. 2022; 12:404.

31. Tosur M, Hsu JW, Deen S, et al. Plasma amino acid signatures define types of pediatric diabetes. Clin Nutr ESPEN. 2023; 57:21-8.

32. Vos MB, Abrams SH, Barlow SE, et al. NASPGHAN Clinical Practice Guideline for the Diagnosis and Treatment of Nonalcoholic Fatty Liver Disease in Children: Recommendations from the Expert Committee on NAFLD (ECON) and the North American Society of Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN). J Pediatr Gastroenterol Nutr. 2017; 64:319-34.

33. Jin R, Banton S, Tran VT, et al. Amino acid metabolism is altered in adolescents with nonalcoholic fatty liver disease-an untargeted, high resolution metabolomics study. J Pediatr. 2016; 172:14-9.e5.

34. Goffredo M, Santoro N, Tricò D, et al. A branched-chain amino acid-related metabolic signature characterizes obese adolescents with non-alcoholic fatty liver disease. Nutrients. 2017; 9:642.

35. Tricò D, Caprio S, Rosaria Umano G, et al. Metabolic features of nonalcoholic fatty liver (NAFL) in obese adolescents: findings from a multiethnic cohort. Hepatology. 2018; 68:1376-90.

36. Lischka J, Schanzer A, Hojreh A, et al. A branched-chain amino acid-based metabolic score can predict liver fat in children and adolescents with severe obesity. Pediatr Obes. 2021; 16:e12739.

37. Leonetti S, Herzog RI, Caprio S, Santoro N, Tricò D. Glutamate-serine-glycine index: a novel potential biomarker in pediatric non-alcoholic fatty liver disease. Children (Basel). 2020; 7:270.

38. Chae W, Lee KJ, Huh KY, Moon JS, Ko JS, Cho JY. Association of metabolic signatures with nonalcoholic fatty liver disease in pediatric population. Metabolites. 2022; 12:881.

39. Nikparast A, Razavi M, Sohouli MH, et al. The association between dietary intake of branched-chain amino acids and the odds of nonalcoholic fatty liver disease among overweight and obese children and adolescents. J Diabetes Metab Disord. 2024; 24:19.

40. Lake AD, Novak P, Shipkova P, et al. Branched chain amino acid metabolism profiles in progressive human nonalcoholic fatty liver disease. Amino Acids. 2015; 47:603-15.

41. Deng Y, Hu M, Huang S, Fu N. Molecular mechanism and therapeutic significance of essential amino acids in metabolically associated fatty liver disease. J Nutr Biochem. 2024; 126:109581.

42. Jensch R, Baber R, Körner A, et al. Association of whole blood amino acid and acylcarnitine metabolome with anthropometry and IGF-I serum levels in healthy children and adolescents in Germany. Metabolites. 2024; 14:489.

43. Hirschel J, Vogel M, Baber R, et al. Relation of whole blood amino acid and acylcarnitine metabolome to age, sex, BMI, puberty, and metabolic markers in children and adolescents. Metabolites. 2020; 10:149.

44. Knacke H, Pietzner M, Do KT, et al. Metabolic fingerprints of circulating IGF-1 and the IGF-1/IGFBP-3 ratio: a multifluid metabolomics study. J Clin Endocrinol Metab. 2016; 101:4730-42.

45. Barazani C, Werner H, Laron Z. Changes in plasma amino acids metabolites, caused by long-term IGF-I deficiency, are reversed by IGF-I treatment - a pilot study. Growth Horm IGF Res. 2020; 52:101312.

46. Aggelaki EA, Giannakopoulos A, Georgiopoulou PD, et al. Unveiling the metabolomic profile of growth hormone deficiency children using NMR spectroscopy. Metabolomics. 2025; 21:25.

47. Inoue M, Sekiguchi K, Kishimoto S, Maeda T, Ihara K. Umbilical cord serum metabolomics identifies amino acid alterations associated with impaired linear growth in small for gestational age infants. Sci Rep. 2025; 15:27378.

48. Hampl SE, Hassink SG, Skinner AC, et al. Clinical practice guideline for the evaluation and treatment of children and adolescents with obesity. Pediatrics. 2023; 151:e2022060640.

49. Rigamonti AE, Frigerio G, Caroli D, et al. A metabolomics-based investigation of the effects of a short-term body weight reduction program in a cohort of adolescents with obesity: a prospective interventional clinical study. Nutrients. 2023; 15:529.

50. Becetti I, Lauze M, Lee H, Bredella MA, Misra M, Singhal V. Changes in branched-chain amino acids one year after sleeve gastrectomy in youth with obesity and their association with changes in insulin resistance. Nutrients. 2023; 15:3801.

51. Glynn EL, Piner LW, Huffman KM, et al. Impact of combined resistance and aerobic exercise training on branched-chain amino acid turnover, glycine metabolism and insulin sensitivity in overweight humans. Diabetologia. 2015; 58:2324-35.

52. Lee S, Gulseth HL, Langleite TM, et al. Branched-chain amino acid metabolism, insulin sensitivity and liver fat response to exercise training in sedentary dysglycaemic and normoglycaemic men. Diabetologia. 2021; 64:410-23.

53. Lim JJ, Prodhan UK, Silvestre MP, et al. Low serum glycine strengthens the association between branched-chain amino acids and impaired insulin sensitivity assessed before and after weight loss in a population with pre-diabetes: the PREVIEW_NZ cohort. Clin Nutr. 2024; 43:17-25.

54. Shah AS, Zeitler PS, Wong J, et al. ISPAD Clinical Practice Consensus Guidelines 2022: Type 2 diabetes in children and adolescents. Pediatr Diabetes. 2022; 23:872-902.

55. American Diabetes Association Professional Practice Committee. 14. Children and Adolescents: Standards of Care in Diabetes-2025. Diabetes Care. 2025; 48(1 Suppl 1):S283-305.

56. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021; 372:n71.

57. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019; 366:l4898.

58. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016; 355:i4919.

J Pediatr Res 2026;13(3):209-19
DOI: 10.4274/jpr.galenos.2026.07448

Original Article

Genotype-Phenotype Correlation Analysis and Identification of a Novel Variant in Children with Congenital Adrenal Hyperplasia due to 21-Hydroxylase Deficiency

Nimet Melis Bilen[1], Gönül Çatlı[2], Cemil Koçyiğit[3], Bumin Nuri Dündar[2]

[1]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Clinic of Pediatrics, İzmir, Türkiye
[2]İzmir Katip Çelebi University Faculty of Medicine, Department of Pediatric Endocrinology, İzmir, Türkiye
[3]University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Clinic of Pediatrics, Division of Pediatric Endocrinology, İzmir, Türkiye

ABSTRACT

Aim: Congenital adrenal hyperplasia (CAH) due to 21-hydroxylase deficiency is an autosomal recessive disorder caused by pathogenic variants in the *CYP21A2* gene. Although genotype is generally predictive of clinical phenotype, genotype-phenotype discordance may occur. This study aimed to evaluate genotype-phenotype correlations and to characterize the molecular spectrum of *CYP21A2* variants in children with CAH.

Materials and Methods: In this retrospective single-center study, 22 patients with CAH due to 21-hydroxylase deficiency were included. The patients were classified as salt-wasting (SW), simple-virilizing (SV), or non-classical (NC) according to clinical and hormonal findings. The detected variants were grouped according to their predicted residual enzyme activity, and genotype-phenotype concordance was evaluated.

Results: Eight patients (36.4%) had the SW phenotype, six (27.3%) had the SV phenotype, and eight (36.4%) had the NC phenotype. The most frequent variant was I2 splice (c.293-13C>G) (25%), followed by I172N (c.515T>A; p.Ile172Asn) (14.2%). Overall genotype-phenotype concordance was observed in 20 of the 22 patients (90%), and genotype severity groups were significantly associated with the clinical phenotype (p=0.001). One homozygous I172N patient exhibited an unexpectedly mild NC phenotype despite an anticipated SV phenotype. A previously unreported compound heterozygous variant, P30L(c.92C>T)/p.E321Afs*60 (c.961_962delGA), was identified in one patient with the NC phenotype.

Conclusion: Our findings demonstrate a strong overall genotype-phenotype correlation in CAH, while also highlighting that patients carrying moderate-severity or novel variants may present with unexpectedly mild or variable clinical phenotypes. Integration of molecular findings with detailed clinical, hormonal, and longitudinal follow-up data remains essential for accurate phenotypic prediction.

Keywords: Congenital adrenal hyperplasia, *CYP21A2*, 21-hydroxylase deficiency, genotype-phenotype correlation, novel variant

Introduction

Congenital adrenal hyperplasia (CAH) is a group of autosomal recessive disorders resulting from enzyme deficiencies in the adrenal steroid biosynthesis pathway (1). These enzyme defects result in decreased cortisol synthesis, which leads to adrenocorticotropic hormone (ACTH) secretion. This causes hyperplasia and the accumulation of steroid precursors in the adrenal cortex (2,3).

Corresponding Author
Nimet Melis Bilen, MD, University of Health Sciences Türkiye, İzmir Tepecik Education and Research Hospital, Clinic of Pediatrics, İzmir, Türkiye
E-mail: nmelisaydin@hotmail.com **ORCID:** 0000-0002-4762-0747
Received: 18.05.2026 **Accepted:** 01.07.2026 **Epub:** 21.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Bilen NM, Çatlı G, Koçyiğit C, Dündar BN. Genotype-phenotype correlation analysis and identification of a novel variant in children with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. J Pediatr Res. 2026;13(3):209-19

The most common cause of CAH is 21-hydroxylase deficiency, accounting for 90-95% of cases (4,5). 11-beta-hydroxylase deficiency accounts for approximately another 5% of cases.

Although the incidence of CAH varies across populations, studies based on data from approximately 6.5 million newborns screened worldwide indicate that the overall incidence of classic CAH is 1 in 13,000 to 1 in 15,000 live births (2,6). These differences may be related to ethnicity, consanguineous marriage rates, and the effectiveness of newborn screening programs (2,7).

21-hydroxylase deficiency results from pathogenic variants in the *CYP21A2* gene, located on the short arm of chromosome 6 (6p21.3) (8). This disease presents in three main clinical phenotypes: the classic salt-wasting (SW) form, the classic simple-virilizing (SV) form, and the non-classic (NC) (late-onset) form (9). The clinical spectrum ranges from severe neonatal salt-losing crises to mild forms presenting with hyperandrogenism in adolescence or adulthood (3,10).

Numerous variants leading to CAH have been identified, and many have been reported to correlate strongly with the clinical phenotype (2,4,11). Therefore, genotype analysis provides several major contributions, especially in prenatal diagnosis, disease prognosis prediction, and cases which are difficult to classify clinically (2,4,12). In addition, recent advances in molecular genetic techniques have allowed for a better understanding of the genotype-phenotype relationship in the diagnosis and follow-up of CAH (9). Genotype-phenotype mismatches in CAH may be influenced by residual enzyme activity, modifier genes, androgen receptor sensitivity, and epigenetic factors, resulting in variable clinical severity despite identical genotypes (2,12).

This study aimed to evaluate the genotype-phenotype correlation in those patients with CAH due to 21-hydroxylase deficiency and to provide an individualized assessment of rare variants and clinically discordant cases.

Materials and Methods

This retrospective single-center study included 22 patients diagnosed with CAH due to 21-hydroxylase deficiency and a molecularly confirmed *CYP21A2* gene mutation. We retrospectively reviewed patient data via their pediatric endocrinology outpatient clinic records and the hospital information management system. We included in this study those patients who had complete clinical, hormonal, and genetic data. We excluded those cases with other adrenal steroidogenesis disorders and those with incomplete follow-up data from this study.

We obtained demographic characteristics, age at presentation, age at diagnosis, gender, presenting complaints, physical examination findings, height, weight, body mass index (BMI), bone age, stage of puberty, Prader stage in female patients, and the type and dosage of treatment from the patient file records.

Height, weight, and BMI standard deviation scores (SDS) were calculated according to age and sex. Bone age was assessed using the Greulich and Pyle atlases, and the bone age/chronological age ratio was calculated (13). In the laboratory evaluation, baseline levels of 17-hydroxyprogesterone (17-OHP), ACTH, cortisol, and total testosterone were recorded at the time of diagnosis and at the last follow-up.

The patients were divided into three phenotypic groups based on their clinical and hormonal characteristics: SW type, SV type, and NC type. The SW phenotype was defined as those cases meeting the criteria for hyponatremia, hyperkalemia, dehydration, and elevated plasma renin activity. The SV phenotype was defined as those patients with prenatal or postnatal androgen-excess findings, without mineralocorticoid deficiency. The NC phenotype included patients presenting with late-onset hyperandrogenism, premature pubarche, or hirsutism (14).

CYP21A2 gene variant analyses of the patients were performed using reverse dot blot analysis, Sanger sequencing, and multiplex ligation-dependent probe amplification (MLPA). Variant nomenclature was reported according to the Human Genome Variation Society recommendations. Novel variants were interpreted according to American College of Medical Genetics and Genomics (ACMG) variant classification criteria.

The detected variants were grouped into four categories based on the literature classification of residual 21-hydroxylase enzyme activity:

Group 1: Null mutations (gene deletion, conversion, 8-bp deletion, E6 cluster, Q318X (c.955C>T;p.Gln319Ter),R356W(c.1066C>T;p.Arg356Trp),

Group 2: I2 splice (c.293-13C>G) homozygous or compound heterozygous variants with severe mutations,

Group 3: I172N (c.515T>A; p.Ile172Asn) homozygous or compound heterozygous mutations with severe variants,

Group 4: mild mutations V281L (c.844G>T; p.Val282Leu), P30L (c.92C>T; p.Pro31Leu), P453S (c.1358C>T; p.Pro453Ser).

According to this grouping, the expected phenotypes were SW, SW/SV, SV, and NC, respectively. Genotype-phenotype concordance was evaluated by comparing the expected and observed phenotypes.

Ethical Consideration

Ethical approval for this study was obtained from the Non-Interventional Research Ethics Committee of İzmir Katip Çelebi University Faculty of Medicine on September 29, 2016, with decision number 220.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 20.0 (IBM Corp., Armonk, NY, USA). Due to the small sample size, the distribution of continuous variables was evaluated using non-parametric methods, and data are presented as median (minimum-maximum) or mean ± standard deviation, as appropriate.

The Kruskal-Wallis test was used to compare continuous variables across the three phenotype groups, and the Wilcoxon signed-rank test was used for intragroup comparisons between the diagnostic and last follow-up measurements. Categorical variables are expressed as numbers and percentages. The association between genotype severity groups and observed phenotypes was evaluated using Fisher's exact test. A p value <0.05 was considered statistically significant.

Results

A total of 22 patients with genetically confirmed CAH due to 21-hydroxylase deficiency were included in this study. Eight patients (36.4%) had the SW phenotype, six (27.3%) had the SV phenotype, and eight (36.4%) had the NC phenotype. The SW group consisted of 5 boys and 3 girls, the SV group consisted entirely of girls, and the NC group consisted of 5 girls and 3 boys.

Age at diagnosis differed significantly between the phenotype groups (p=0.006). The median age at diagnosis was 1.6 months (0.5-4.0) in the SW group, 5.6 months (0.3-72.0) in the SV group, and 76.8 months (1.0-198.0) in the NC group. Those patients with the SW and SV phenotypes were diagnosed at significantly earlier ages than those in the NC group.

BMI SDS showed an increasing trend in all phenotype groups during the follow-up period. This increase was more pronounced in the classic phenotype patients, and at the last follow-up, BMI SDS values had moved into the positive range in all groups. The individual clinical presentations, phenotypes, and genotypes of the patients are presented in Table I.

No significant difference was found between the phenotype groups in terms of their baseline 17-OHP and ACTH levels (p>0.05). However, cortisol levels were higher in the SW and NC groups when compared to the SV group (p=0.028). Total testosterone levels were significantly higher in the SV group in comparison to the other two groups (p=0.026). There were no significant differences among the phenotype groups in terms of their follow-up hormone values.

Table I. Patients' presenting symptoms, phenotypes, and detected genetic mutations

Patient	Gender	Presenting symptoms	Phenotype	Allele 1	Allele 2	Allelic interpretation
1	Male	Vomiting	SW	R356W (c.1066C>T; p.Arg356Trp)	R356W (c.1066C>T; p.Arg356Trp)	Homozygous
2	Female	Vomiting, ambiguous genitalia	SW	I2splice (c.293-13C>G) + Q318X (c.955C>T; p.Gln319Ter)	Q318X (c.955C>T; p.Gln319Ter)+ E6 cluster + R356W (c.1066C>T; p.Arg356Trp)	Complex alleles
3	Female	Ambiguous genitalia, hyponatremia	SW	I2 splice (c.293-13C>G)+ Q318X (c.955C>T; p.Gln319Ter)+ E6 cluster + R356W (c.1066C>T; p.Arg356Trp)	R356W (c.1066C>T; p.Arg356Trp)	Complex allele
4	Male	Vomiting	SW	I2 splice (c.293-13C>G)	I2 splice (c.293-13C>G)	Homozygous
5	Male	Vomiting, diarrhea	SW	Exon 1-7 conversion	Exon 1-3 conversion	Compound heterozygous conversion
6	Female	Ambiguous genitalia	SW	Q318X (c.955C>T; p.Gln319Ter)	Q318X (c.955C>T; p.Gln319Ter)+ E6 cluster + R356W (c.1066C>T; p.Arg356Trp)	Complex allele
7	Male	Vomiting	SW	I2 splice (c.293-13C>G)	I2 splice (c.293-13C>G)	Homozygous

Table I. Continued

Patient	Gender	Presenting symptoms	Phenotype	Allele 1	Allele 2	Allelic interpretation
8	Male	Vomiting	SW	No mutation detected	No mutation detected	Mutation-negative clinically confirmed case
9	Female	Clitoromegaly	SV	I172N (c.515T>A; p.Ile172Asn)	I2 splice (c.293-13C>G)	Compound heterozygous
10	Female	Ambiguous genitalia	SV	I2 splice (c.293-13C>G)	I2 splice (c.293-13C>G)	Homozygous
11	Female	Clitoromegaly	SV	I172N (c.515T>A; p.Ile172Asn)	I172N (c.515T>A; p.Ile172Asn)	Homozygous
12	Female	Ambiguous genitalia	SV	I172N (c.515T>A; p.Ile172Asn)	I2 splice (c.293-13C>G)	Compound heterozygous
13	Female	Ambiguous genitalia	SV	I2 splice (c.293-13C>G)	I2 splice (c.293-13C>G)	Homozygous
14	Female	Ambiguous genitalia	SV	P30L (c.92C>T; p.Pro31Leu) + I2 splice (c.293-13C>G) + 8-bp deletion	P30L (c.92C>T; p.Pro31Leu)+I2 splice c.293-13C>G	Complex alleles
15	Male	Premature pubarche	NC	I172N (c.515T>A; p.Ile172Asn)	I172N (c.515T>A; p.Ile172Asn)	Homozygous
16	Female	Premature pubarche	NC	Q318X (c.955C>T; p.Gln319Ter) + V281L (c.844G>T; p.Val282Leu)	V281L (c.844G>T; p.Val282Leu) + I172N (c.515T>A; p.Ile172Asn)	Complex alleles
17	Female	Premature pubarche	NC	V281L (c.844G>T; p.Val282Leu)	P30L (c.92C>T; p.Pro31Leu)	Compound heterozygous
18	Male	Premature pubarche	NC	P30L (c.92C>T; p.Pro31Leu)	c.961_962delGA (p.Glu321Alafs*60)	Compound heterozygous novel variant
19	Female	No symptoms	NC	P453S (c.1358C>T; p.Pro453Ser)	P453S (c.1358C>T; p.Pro453Ser)	Homozygous
20	Male	Premature pubarche	NC	p.R340H c.1019G>A (p.Arg340His) + P453S (c.1358C>T; p.Pro453Ser)	P453S (c.1358C>T; p.Pro453Ser)	Complex allele/ compound heterozygous
21	Female	Premature pubarche	NC	P453S (c.1358C>T; p.Pro453Ser)	P453S (c.1358C>T; p.Pro453Ser)	Homozygous
22	Female	Virilization	NC	V281L (c.844G>T; p.Val282Leu)	I172N (c.515T>A; p.Ile172Asn)	Compound heterozygous

*Traditional *CYP21A2* variant nomenclature is presented together with HGVS nomenclature. Variant nomenclature has been reported according to HGVS recommendations using the *CYP21A2* reference transcript NM_000500.9
SW: Salt-wasting, SV: Simple-virilizing, NC: Non-classical, HGVS: Human Genome Variation Society

The detailed clinical, auxological, hormonal, and longitudinal follow-up characteristics of the phenotype groups are summarized in Table II. The individual clinical presentations, Prader stages, pubertal characteristics, surgical interventions, and genotypes of the patients are presented in Table III.

Molecular analysis identified a total of 56 variant events representing 17 distinct variants across 44 alleles from 22 patients. As some alleles carried multiple pathogenic changes, the total number of detected variant events exceeded the number of alleles analyzed. One patient exhibited a clinically significant SW phenotype, but no mutation could be demonstrated. The most frequent variant was I2 splice (c.293-13C>G) (25%), followed by I172N (c.515T>A; p.Ile172Asn) (14.2%). These were followed by R356W (c.1066C>T; p.Arg356Trp) (10.7%), Q318X (c.955C>T; p.Gln319Ter) (10.7%), and P453S (c.1358C>T; p.Pro453Ser) (10.7%). The distribution and frequencies of the detected variants according to phenotype groups are summarized in Table IV.

Table II. Clinical, auxological, and hormonal characteristics according to phenotype groups

	Salt-wasting (n=8)	Simple-virilizing (n=6)	Non-classical (n=8)	p value
Age at diagnosis (months)	1.6±1.3	5.6±4.2	76.8±63	0.006
Target height SDS at diagnosis	-0.87±1.4	-0.45±1.17	0.46±1.87	0.2
Target height SDS at last follow-up	0.08±1.15	0.92±0.7	0.30±1.49	0.4
Height SDS at diagnosis	0.17±2	-0.02±1.3	1.4±1.6	0.161
Height SDS at last follow-up	-1.64±0.82	-0.36±1.5	-0.55±2.03	0.2
BMI SDS at diagnosis	-1.67±1.09	-0.82±1.73	-0.45±2.2	0.2
BMI SDS at last follow-up	1.2±1.2	1.39±0.49	1.01±1.3	0.9
Basal 17-OHP at diagnosis (ng/mL)	58.6 (4.4-207)	36.5 (1.9-110)	23.4 (3.2-41.8)	0.5
Basal 17-OHP at last follow-up (ng/mL)	27.83 (0.08-128)	20.9 (1.08-27)	9.7 (0.40-32.1)	0.2
ACTH at diagnosis (pg/mL)	42.8 (18-70)	153.5 (13-338)	48.3 (29-100)	0.5
ACTH at last follow-up (pg/mL)	19.6 (2-55)	152.3 (3-347)	30.4 (2-138)	0.2
Basal cortisol at diagnosis (µg/dL)	17.07 (1-50)	10.7 (4.9-18)	21.2 (14.2-27)	0.028
Basal cortisol at last follow-up (µg/dL)	8.63 (0.5-34)	9.07 (3-14.8)	10.1 (1-33.8)	0.7
Testosterone at diagnosis (ng/dL)	136.9 (0.4-1014)	494.7 (23.4-1600)	40.3 (0.17-107)	0.026
Testosterone at last follow-up (ng/dL)	32.2 (1.2-163.10)	58.7 (15-161)	34.1 (0.12-115)	0.23

Data are presented as mean±standard deviation or median (minimum-maximum), as appropriate.
SDS: Standard deviation score, BMI: Body mass index, 17-OHP: 17-hydroxyprogesterone, ACTH: Adrenocorticotropic hormone

Table III. Individual clinical characteristics including prader staging, pubertal development, and surgical outcomes

Patient ID	Phenotype	Prader stage at diagnosis	Pubertal stage at diagnosis	Surgical procedure	Age at surgery (years)
2	SW	3	1	Vaginal reconstruction + clitoroplasty	9
3	SW	3	1	None	-
6	SW	4	1	None	-
9	SV	3	1	Clitoroplasty	4
10	SV	4	1	Clitoroplasty	3
11	SV	4	1	Clitoroplasty	3
12	SV	4	1	Vaginal reconstruction + Clitoroplasty	12
13	SV	4	1	None	-
14	SV	3	1	Clitoroplasty	5
15	NC	2	2	None	-
16	NC	1	4	None	-
18	NC	1	5	None	-
20	NC	2	1	None	-
22	NC	1	5	None	-

SW: Salt-wasting, SV: Simple-virilizing, NC: Non-classical

When the variant distribution according to phenotype groups was examined, the most frequent variants were I2 splice, R356W, and Q318X in the SW phenotype; I2 splice and I172N in the SV phenotype; and P453S, V281L, and I172N in the NC phenotype (Table IV). A previously unreported P30L (c.92C>T)/p. E321Afs*60 (c.961_962delGA) variant was detected in one patient. According to ACMG variant classification criteria, this novel variant was considered *likely pathogenic* based on the presence of Pathogenic Very Strong 1 (PVS1) and Pathogenic Moderate 2 (PM2) evidence (15). The distribution of detected *CYP21A2* variants across the gene is summarized in a lollipop plot in Figure 1.

For genotype-phenotype correlation analysis, the patients were divided into four groups based on their residual enzyme activity. In Groups 1 and 2, the expected and observed phenotypes were completely concordant in all patients. In Group 3, one homozygous I172N patient exhibited the NC phenotype instead of the expected SV phenotype. In Group 4, the expected and observed NC phenotypes were completely concordant in all patients. The results of the genotype-phenotype concordance are presented in Table V. The genotype severity groups were significantly associated with the observed clinical phenotypes (p=0.001). Discordance was detected in only one case of homozygous I172N.

Table IV. Distribution of detected ***CYP21A2*** variant events and estimated residual enzyme activities

Variant	Protein change	Estimated residual enzyme activity	Salt-wasting (n=22)	Simple-virilizing (n=15)	Non-classical (n=19)	Total (n=56)
R356W (c.1066C>T)	p.Arg356Trp	0% (null)	6 (27.2%)	-	-	6 (10.7%)
Q318X (c.955C>T)	p.Gln319Ter	0% (null)	5 (22.7%)	-	1 (5.2%)	6 (10.7%)
E6 cluster	-	0% (null)	3 (13.6%)	-	-	3 (5.3%)
I2 splice (c.293-13C>G)	Splice-site variant	<1%	6 (27.2%)	8 (53.3%)	-	14 (25%)
Conversion	-	0% (null)	2 (9%)	-	-	2 (3.5%)
I172N (c.515T>A)	p.Ile172Asn	~1-2%	-	4 (26.6%)	4 (21%)	8 (14.2%)
P30L (c.92C>T)	p.Pro31Leu	~30-60%	-	2 (13.3%)	2 (10.5%)	4 (7.1%)
8-bp deletion (c.332_339del)	Frameshift variant	0% (null)	-	1 (6.6%)	-	1 (1.7%)
V281L (c.844G>T)	p.Val282Leu	~20-50%	-	-	4 (21%)	4 (7.1%)
P453S (c.1358C>T)	p.Pro453Ser	~20-50%	-	-	6 (31.5%)	6 (10.7%)
p.R340H c.1019G>A	p.Arg340His	Unknown/likely mild-modifying	-	-	1 (5.2%)	1 (1.7%)
p.E321Afs*60 (c.961_962delGA)	p.Glu321Alafs*60	Unknown; frameshift component likely severe	-	-	1 (5.2%)	1 (1.7%)

Variant nomenclature follows Human Genome Variation Society (HGVS) recommendations using the ***CYP21A2*** reference transcript NM_000500.9.
Estimated residual enzyme activities and mutation severity classifications were derived from previously published functional and genotype-phenotype studies (1,2,12,15,20,25).
Null variants include non-sense, frameshift, splice-site, large conversion, and severe cluster variants associated with absent or markedly impaired 21-hydroxylase activity

Table V. Genotype severity groups, hormonal characteristics, and genotype-phenotype concordance in patients with 21-hydroxylase deficiency

Group	Detailed genotype classification	n	SW	SV	NC	Median basal 17-OHP (ng/mL)	Median basal cortisol (µg/dL)	Expected phenotype	Genotype-phenotype concordance	Comments	p value
Group 1	**Null/severe variants**	5	5	-	-	20.0	15.0	SW	100%		
	R356W (c.1066C>T; p.Arg356Trp) / R356W (c.1066C>T; p.Arg356Trp)	1	1	-	-			SW	100%		
	Q318X (c.955C>T; p.Gln319Ter) / E6 cluster / R356W (c.1066C>T; p.Arg356Trp) / I2 splice (c.293-13C>G)	3	3	-	-			SW	100%		
	Large conversion variants	1	1	-	-			SW	100%		
Group 2	**Severe splice variants**	5	2	3	-	108.0	8.2	SW-SV	100%		
	I2 splice (c.293-13C>G) / I2 splice (c.293-13C>G)	4	2	2	-			SW-SV	100%		
	I2 splice (c.293-13C>G) / I2 splice (c.293-13C>G) / 8-bp deletion / P30L (c.92C>T; p.Pro31Leu) / P30L (c.92C>T; p.Pro31Leu)	1	-	1	-			SW-SV	100%		
Group 3	**Moderate residual activity variants**	4	-	3	1	14.0	14.0	SV	75% overall		
	I172N (c.515T>A; p.Ile172Asn) / I172N (c.515T>A; p.Ile172Asn)	2	-	1	1			SV	50%	One patient presented with an NC phenotype despite homozygous I172N genotype. Similar discordant cases were reported by Lao et al. (20) supporting phenotypic plasticity associated with moderate residual enzyme activity.	
	I172N (c.515T>A; p.Ile172Asn) / I2 splice (c.293-13C>G)	2	-	2	-			SV	100%		
Group 4	**Mild/non-classical variants**	6	-	-	6	24.0	20.95	NC	100%		
	I172N (c.515T>A; p.Ile172Asn) / V281L (c.844G>T; p.Val282Leu)	1	-	-	1			NC	100%		
	Q318X (c.955C>T; p.Gln319Ter) / I172N (c.515T>A; p.Ile172Asn) / V281L (c.844G>T; p.Val282Leu)	1	-	-	1			NC	100%		
	V281L (c.844G>T; p.Val282Leu) / P30L (c.92C>T; p.Pro31Leu)	1	-	-	1			NC	100%		
	P453S (c.1358C>T; p.Pro453Ser) / P453S (c.1358C>T; p.Pro453Ser)	2	-	-	2			NC	100%		
	P453S (c.1358C>T; p.Pro453Ser) / P453S (c.1358C>T; p.Pro453Ser) / p.R340H (c.1019G>A; p.Arg340His)	1	-	-	1			NC	100%		

Table V. Continued

Group	Detailed genotype classification	n	SW	SV	NC	Median basal 17-OHP (ng/mL)	Median basal cortisol (µg/dL)	Expected phenotype	Genotype-phenotype concordance	Comments	p value
Others	**Unclassified/novel variants or mutation-negative cases**	2	1	-	1	12.2	36.0				
	P30L (c.92C>T; p.Pro31Leu) / p.E321Afs*60 (c.961_962delGA); *p.Glu321Alafs*60)	1	-	-	1			SW/SV	-	Previously unreported compound heterozygous variant NC phenotype despite likely pathogenic frameshift variant; possible residual activity effect of P30L allele.	
	No mutation detected	1	1	-	-			Undefined	-	Clinically and hormonally confirmed SW phenotype despite negative molecular testing. Similar mutation-negative cases have been described in previous cohorts (6) and may reflect deep intronic variants or complex RCCX rearrangements undetectable by standard Sanger/MLPA methods.	
Total		22	8	6	8				90%	Strong overall genotype-phenotype correlation	*p=0.001

Variant nomenclature follows Human Genome Variation Society (HGVS) recommendations using the *CYP21A2* reference transcript NM_000500.9.
Median basal 17-OHP and cortisol values were calculated according to genotype severity groups using baseline hormonal measurements obtained at diagnosis.
*Genotype severity groups were significantly associated with observed clinical phenotypes (p=0.001). However, interpretation should be made cautiously because of the limited sample size and sparse subgroup distribution.
SW: Salt-Wasting, SV: Simple-Virilizing, NC: Non-classical, 17-OHP: 17-hydroxyprogesterone, MLPA: multiplex ligation-dependent probe amplification

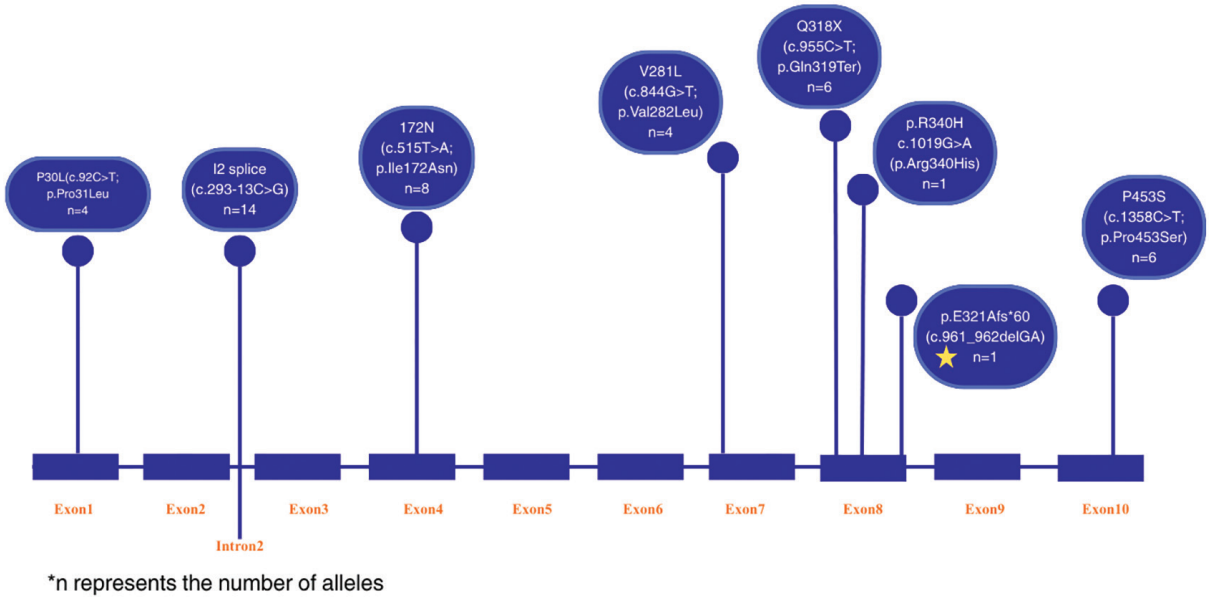

Figure 1. Lollipop plot demonstrating the distribution of *CYP21A2* variants identified in the cohort

Discussion

In this study, we evaluated genotype-phenotype correlation in pediatric patients with CAH due to 21-hydroxylase deficiency and identified a strong overall concordance between the *CYP21A2* genotype and the clinical phenotype. The I2 splice variant was the most frequently detected mutation, particularly among those patients with classical phenotypes, while I172N was mainly associated with the SV phenotype. A genotype-phenotype discordance was observed in one patient carrying homozygous I172N variants. In addition, we identified a previously unreported compound heterozygous variant, P30L (c.92C>T)/p.E321Afs*60 (c.961_962delGA), which may contribute to the expanding molecular spectrum of *CYP21A2*-related disease.

Complete phenotypic concordance was observed in Groups 1 and 2, which included null mutations and severe splice-site variants associated with an absence of enzyme activity, highlighting the importance of *CYP21A2* genotyping in predicting the severity of classical CAH. This finding is consistent with the strong correlation reported by New et al. (12) in their large cohort of 1,507 families. Furthermore, Speiser and White emphasized in their review that severe mutations were particularly strongly associated with the SW phenotype (16).

In our study, the most frequent variant was the I2 splice, which was consistent with previous reports (3,8,10,17). Baş et al. (6) identified the I2 splice as the most frequent variant (22%), primarily observed in patients with SW and SV forms. Similarly, in a large Turkish series of 124 patients reported by Kirac et al. (18), splice-site and deletion/conversion mutations represented the majority of the detected *CYP21A2* variants. The predominance of the I2 splice variant, particularly among those patients with classical phenotypes, further supports the phenotype-determining importance of this mutation in the Turkish population.

We detected the I172N variant as the second most frequent in our study. It was mostly associated with the SV phenotype, consistent with previous reports (10,16,19). In Turkish, Korean, and Chinese pediatric series, this variant has been shown to be frequently associated with the SV phenotype (3,6,10,18). One of the most striking findings in our cohort was the observation of a NC phenotype in a patient homozygous for the I172N variant (Case 15), rather than the expected SV clinical presentation. While the *CYP21A2* genotype is generally a robust predictor of clinical course, such discordances have been reported in 10% to 20% of cases in the literature (12,15,17). This phenotypic variability is particularly prevalent in variants with moderate residual enzyme activity, such as I172N. According to a comprehensive study by Lao et al. (20), which analyzed 457 individuals across the genotypic spectrum, the I172N variant represents a critical "threshold" in adrenal steroidogenesis. The I172N mutation typically retains approximately 1% to 2% of residual 21-hydroxylase activity, a level which is usually sufficient to prevent neonatal SW crises but resides on the borderline between the SV and NC phenotypes (20).

Consistent with our findings, Lao et al. (20) provided evidence that the clinical expression of homozygous I172N is not monolithic; they identified several instances where individuals shared this genotype yet presented with a milder NC phenotype rather than the classic SV form. This "phenotypic plasticity" suggests that the final clinical outcome is modulated by factors beyond the primary *CYP21A2* sequence. Phenotypic variability in those patients carrying the I172N variant may be influenced by alternative splicing efficiency, androgen receptor sensitivity, extra-adrenal steroid metabolism, and/or modifier genes affecting residual enzyme activity (2,12,16,21-23). Our study identified a patient with a SW phenotype in whom no mutation was detected by standard Sanger sequencing and MLPA (Case 8). While seemingly contradictory, this phenomenon has been reported in other cohorts, including those by Lao et al. (20) and Baş et al. (6). These cases likely represent the technical limitations of the current diagnostic panels in identifying promoter region variants, deep intronic mutations, or complex RCCX module rearrangements (2,5,12,24). In conclusion, our data, supported by the recent large-scale findings of Lao et al. (20), underscore that while genotyping is essential for prognostic prediction, it must be interpreted with caution. The existence of "grey zones", especially for moderate variants such as I172N, highlights the necessity of a multidisciplinary approach which integrates molecular data with longitudinal clinical observation and biochemical monitoring in order to optimize patient management.

The novel P30L (c.92C>T)/ p.E321Afs*60 (c.961_962delGA) variant identified in our cohort appears to be rare and, to the best of our knowledge, has not been previously reported in the literature. Numerous novel variants have been identified in different series in the Turkish population previously (1,18). The novel frameshift variant c.961_962delGA (p.Glu321Alafs*60) was classified as *likely pathogenic* according to ACMG criteria based on the presence of PVS1 and PM2 evidence. Since frameshift variants are generally expected to result in a severe loss of enzyme activity through the disruption of the normal protein sequence and premature termination, a more severe clinical phenotype might be anticipated (5,20). However, the patient carrying this variant presented with a NC phenotype in compound heterozygosity with the mild P30L (c.92C>T) variant. This observation suggests that residual enzymatic activity associated with the P30L allele may partially mitigate the clinical severity of this novel variant (20). In addition, the discrepancy between predicted molecular severity and observed phenotype further supports the concept that clinical expression in *CYP21A2* deficiency may be influenced by additional genetic and biological modifiers beyond the primary genotype alone.

Beyond genotype-phenotype concordance, our cohort also demonstrated distinct clinical and hormonal differences among the phenotype groups. Those patients with classical phenotypes were diagnosed at earlier ages and showed more severe virilization, greater skeletal maturation advancement, and increased rates of genital corrective surgery compared to those with the NC phenotype. In addition, the increase in BMI SDS observed during follow-up, particularly among those patients with classical phenotypes, may reflect the long-term metabolic consequences of chronic glucocorticoid exposure, which has been reported previously in patients with CAH (25,26). Together, these findings suggest that detailed longitudinal clinical characterization may provide important complementary information beyond molecular analysis alone.

Study Limitations

The main limitations of this study include its retrospective single-center design and its relatively small sample size. Functional analyses could not be performed for the novel variant. However, this study provides detailed longitudinal characterization, comprehensive genotype-phenotype analysis, and the identification of a novel likely pathogenic *CYP21A2* variant.

Conclusion

Our study demonstrates a strong genotype-phenotype correlation in those patients with CAH due to 21-hydroxylase deficiency, while also highlighting that moderate-severity and novel variants may present with unexpectedly mild or variable clinical phenotypes. The individualized evaluation of clinically discordant cases provides additional insight into the complexity of phenotypic expression in *CYP21A2* deficiency. Therefore, integrating molecular findings with clinical, hormonal, and longitudinal follow-up data remains essential for accurate phenotypic prediction and long-term patient management.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the Non-Interventional Research Ethics Committee of İzmir Katip Çelebi University Faculty of Medicine on September 29, 2016, with decision number 220.

Informed Consent: Retrospective cohort study.

Footnotes

Authorship Contributions

Concept: G.Ç., B.N.D., Design: N.M.B., G.Ç., B.N.D., Data Collection or Processing: N.M.B., C.K., Analysis or Interpretation: N.M.B., B.N.D., Literature Search: G.Ç., C.K., B.N.D., Writing: N.M.B., G.Ç., C.K.

Conflict of Interest: The authors declare no conflict of interest.

Financial Disclosure: The authors declare that no financial or material support was received for this study.

References

1. Gungor N, Saad R, Janosky J, Arslanian S. Validation of surrogate estimates of insulin sensitivity and insulin secretion in children and adolescents. J Pediatr. 2004; 144:47-55.
2. Concolino P, Falhammar H. Genetics in congenital adrenal hyperplasia due to 21-hydroxylase deficiency and clinical implications. J Endocr Soc. 2025; 9:bvaf018.
3. Yoo Y, Chang MS, Lee J, et al. Genotype-phenotype correlation in 27 pediatric patients in congenital adrenal hyperplasia due to 21-hydroxylase deficiency in a single center. Ann Pediatr Endocrinol Metab. 2013; 18:128-34.
4. Tang P, Zhang J, Peng S, et al. Genotype-phenotype correlation in patients with 21-hydroxylase deficiency. Front Endocrinol (Lausanne). 2023; 14:1095719.
5. Speiser PW, Azziz R, Baskin LS, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2010; 95:4133-60.
6. Baş F, Kayserili H, Darendeliler F, et al. CYP21A2 gene mutations in congenital adrenal hyperplasia: genotype-phenotype correlation in Turkish children. J Clin Res Pediatr Endocrinol. 2009; 1:116-28.
7. Kandemir N, Yordam N. Congenital adrenal hyperplasia in Turkey: a review of 273 patients. Acta Paediatr. 1997; 86:22-5.
8. Marumudi E, Sharma A, Kulshreshtha B, Khadgawat R, Khurana ML, Ammini AC. Molecular genetic analysis of CYP21A2 gene in patients with congenital adrenal hyperplasia. Indian J Endocrinol Metab. 2012; 16:384-8.
9. Espinosa Reyes TM, Collazo Mesa T, Lantigua Cruz PA, Agramonte Machado A, Domínguez Alonso E, Falhammar H. Genotype-phenotype correlation in patients with congenital adrenal hyperplasia due to 21-hydroxylase deficiency in Cuba. Int J Endocrinol. 2021; 2021:9316284.
10. Wang R, Yu Y, Ye J, et al. 21-hydroxylase deficiency-induced congenital adrenal hyperplasia in 230 Chinese patients: genotype-phenotype correlation and identification of nine novel mutations. Steroids. 2016; 108:47-55.
11. Fylaktou I, Mertzanian A, Farakla I, et al. Genetics of 21-OH deficiency and genotype-phenotype correlation: experience of the hellenic national referral center. Curr Issues Mol Biol. 2024; 46:10696-713.
12. New MI, Abraham M, Gonzalez B, et al. Genotype-phenotype correlation in 1,507 families with congenital adrenal hyperplasia owing to 21-hydroxylase deficiency. Proc Natl Acad Sci U S A. 2013; 110:2611-6.
13. Özgül E, Yücel A, Kaya F, Koca SB, Ertekin A. Determination of bone age and evaluating the applicability of Greulich-Pyle standards among the Turkish children. J Surg Med. 2022; 6:9-13.
14. Yang M, White PC. Genetics and pathophysiology of classic congenital adrenal hyperplasia due to 21-hydroxylase deficiency. J Clin Endocrinol Metab. 2025; 110(Suppl 1):S1-12.
15. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015; 17:405-24.
16. Speiser PW, White PC. Congenital adrenal hyperplasia. N Engl J Med. 2003; 349:776-88.
17. Kurzyńska A, Skalniak A, Franson K, Bistika V, Hubalewska-Dydejczyk A, Przybylik-Mazurek E. Molecular analysis and genotype-phenotype correlations in patients with classical congenital adrenal hyperplasia due to 21-hydroxylase deficiency from southern Poland-experience of a clinical center. Hormones (Athens). 2022; 21:241-9.
18. Kirac D, Guney AI, Akcay T, et al. The frequency and the effects of 21-hydroxylase gene defects in congenital adrenal hyperplasia patients. Ann Hum Genet. 2014; 78:399-409.
19. Yoon JY, Cheon CK. Genotype and clinical outcomes in children with congenital adrenal hyperplasia. Pediatr Int. 2021; 63:658-63.
20. Lao Q, Schulman A, Kulkarni S, et al. Clinical and biochemical phenotype across the genotypic spectrum of 21-hydroxylase deficiency in 457 individuals. J Clin Endocrinol Metab. 2026; 111:1098-113.
21. Santos-Silva R, Cardoso R, Lopes L, et al. CYP21A2 gene pathogenic variants: a multicenter study on genotype-phenotype correlation from a Portuguese pediatric cohort. Horm Res Paediatr. 2019; 91:33-45.
22. Narasimhan ML, Khattab A. Genetics of congenital adrenal hyperplasia and genotype-phenotype correlation. Fertil Steril. 2019; 111:24-9.
23. Einaudi S, Napolitano E, Restivo F, et al. Genotype, phenotype and hormonal levels correlation in non-classical congenital adrenal hyperplasia. J Endocrinol Invest. 2011; 34:660-4.
24. Khajuria R, Walia R, Bhansali A, Prasad R. The spectrum of CYP21A2 mutations in congenital adrenal hyperplasia in an Indian cohort. Clin Chim Acta. 2017; 464:189-94.
25. Paizoni L, Auer MK, Schmidt H, Hübner A, Bidlingmaier M, Reisch N. Effect of androgen excess and glucocorticoid exposure on metabolic risk profiles in patients with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. J Steroid Biochem Mol Biol. 2020; 197:105540.
26. Barbot M, Mazzeo P, Lazzara M, Ceccato F, Scaroni C. Metabolic syndrome and cardiovascular morbidity in patients with congenital adrenal hyperplasia. Front Endocrinol (Lausanne). 2022;13:934675.

J Pediatr Res 2026;13(3):220-6
DOI: 10.4274/jpr.galenos.2026.67026

Original Article

Blood Pressure Trends and Influencing Factors in the First Week of Life: A Prospective Study in Healthy Term Turkish Newborns

Ömer Güran[1], Sinan Uslu[2], Ali Bülbül[2], Evrim Kıray Baş[2], Muhittin Çelik[2], Umut Zübarioğlu[2]

[1]University of Health Sciences Türkiye, Ümraniye Training and Research Hospital, Clinic of Neonatology, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Şişli Hamidiye Etfal Training and Research Hospital, Clinic of Neonatology, İstanbul, Türkiye

ABSTRACT

Aim: To establish descriptive reference blood pressure (BP) values in healthy full-term Turkish newborns during the first seven postnatal days and to identify associated perinatal factors.

Materials and Methods: In this prospective observational study (April-September 2012), 419 healthy term newborns (≥37 weeks of gestation) underwent daily BP measurement using the oscillometric method. Exclusion criteria included neonatal intensive care unit admission, low 5th minute Apgar scores (<7), congenital anomalies, maternal pregnancy complications, or chronic disease.

Results: The newborns' mean birth weight, length, and gestational age were 3,302±398 g, 48.95±1.9 cm, and 39.28±1.19 weeks, respectively. The mean BP increased significantly from Day 1 to Days 5-7 (48.10±7.35 mmHg → 59.23±6.79 mmHg; p<0.0001). Sex and parity were not associated with BP. Caesarean delivery was linked to lower Day 1 BP. Birth length, the five-minute Apgar score, and maternal systolic BP were positively correlated with Day 1 BP, although correlation coefficients indicated weak associations (r<0.3).

Conclusion: This study provides descriptive early postnatal BP values for healthy Turkish newborns. Maternal BP, neonatal length, and Apgar score influence first-day BP, but their clinical relevance is limited due to weak correlations and the low explanatory power of regression models.

Keywords: Newborn, blood pressure, reference range, perinatal factors, term infants, Turkey

Introduction

Blood pressure (BP) monitoring plays a critical role in the management of sick newborns, both for detecting hypotension and recognizing hypertension. Accurate interpretation requires knowledge of normal BP values stratified by gestational age, postnatal age, sex, race/ethnicity, and other relevant factors (1,2).

Historically, various methods have been used to assess BP in newborns (3). Currently, oscillometric devices are the most commonly employed non-invasive method in neonatal units. However, intrinsic and extrinsic factors, such as software miscalculations, incorrect cuff size, and movement artefacts, may affect accuracy. In order to minimize these errors, several guidelines have been published (4,5).

Corresponding Author
Ömer Güran, MD, University of Health Sciences Türkiye, Ümraniye Training and Research Hospital, Clinic of Neonatology, İstanbul, Türkiye
E-mail: guran_omer@hotmail.com **ORCID:** orcid.org/0000-0002-5146-6949
Received: 21.04.2026 **Accepted:** 11.06.2026 **Epub:** 31.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Güran Ö, Uslu S, Bülbül A, Kıray Baş A, Çelik M, Zübarioğlu U. Blood pressure trends and influencing factors in the first five days of life: a prospective study in healthy term Turkish newborns. J Pediatr Res. 2026;13(3):220-6

Most neonatal reference ranges reported in textbooks are derived from the multicenter study by Zubrow et al. (6), which used oscillometric measurements. That study included neonatal intensive care unit (NICU) patients across a wide range of gestational ages, potentially influencing the reported values. In order to reduce such bias, subsequent investigators focused on healthy, term newborns in well-baby nurseries (7,8).

Compared with adults and older children, percentile curves for neonatal BP remain scarce. This may reflect the rarity of cardiovascular disease in the newborn period or historical limitations in measurement techniques. An increasing recognition that adult cardiovascular diseases may have origins in early life, or even in the fetal period, has shifted attention toward the neonatal stage (9). Careful evaluation of BP, one of the earliest parameters affected by a newborn's health status, allows for the early detection and management of septic, metabolic, or hemodynamic disturbances.

To date, no study has established reference BP values specifically for healthy term Turkish newborns. Previous reports from Türkiye have focused on NICU populations or preterm infants, limiting their applicability to healthy infants (10,11). The aim of the present study was therefore to provide descriptive BP values in healthy Turkish term newborns during the first seven postnatal days, to compare these values with previously published international data, and to examine the influence of perinatal factors on early postnatal BP.

Materials and Methods

Study Population

This prospective observational study was conducted between April and September 2012 at a tertiary hospital. A total of 419 healthy term newborns (≥37 weeks of gestation) were consecutively recruited. Ethical approval was obtained from the University of Health Sciences Türkiye, Şişli Hamidiye Etfal Training and Research Hospital Clinical Research Ethics Committee (approval no.: 33, date: 10.04.2012), and informed parental consent was secured. Exclusion criteria included maternal chronic or acute diseases, pregnancy complications, congenital anomalies, NICU admission, low 5th minute Apgar scores (<7), low birth weight (<2,500 g), fetal malnutrition (CAN score <26), or non-Turkish ethnicity.

Data Collection

Demographic and perinatal data were obtained from medical records and parental interviews. Variables included infant sex, gestational age, birth weight, birth length, Apgar scores, and maternal characteristics. Maternal weight was measured with a digital scale, and BP was assessed by the auscultatory method.

Blood Pressure Measurements

Neonatal BP was measured using an oscillometric device (DRÄGER Siemens SC 6002 XL, Siemens Co., NY, USA), calibrated according to manufacturer's recommendations. The first BP measurement was performed 12-24 hours after birth, followed by daily measurements between 9:00 AM and 12:00 PM while the infants remained hospitalized. After discharge, additional measurements were obtained during routine outpatient follow-up visits within the first week. The measurements were obtained while the infant was quietly awake or asleep, at least one hour after feeding, using the right upper arm. Cuff size was selected so that the inflatable bladder width was at least 40% of the mid-arm circumference and its length covered 80-100% of the arm circumference (3). Each measurement was repeated three times, and the average was used for analysis.

Statistical Analysis

Analyses were performed using NCSS 2007 (Number Cruncher Statistical System, Utah, USA). Continuous variables are expressed as mean ± standard deviation, median, and range; categorical variables as counts and percentages. Percentiles for systolic, diastolic, and mean BP (SBP, DBP, MBP) were calculated for each day.

Days 5-7 were combined due to the small number of infants measured on each of these days in order to ensure a sufficient sample size for statistical analysis. Repeated measurements were analyzed with paired variance analysis, and subgroup comparisons were made using the Newman-Keuls multiple comparison test. Independent t-tests were used to compare two continuous variables and chi-square tests were used for categorical variables. Correlations were assessed using Pearson's correlation coefficient. Multivariate linear regression analysis was performed in order to identify independent predictors of neonatal BP. Statistical significance was defined as $p<0.05$.

Results

A total of 419 healthy term newborns were enrolled. The number of infants with valid BP measurements declined from Day 1 to Days 5-7 (Figure 1) (401, 287, 169, 85, and 100, respectively). Baseline demographic and maternal characteristics are summarized in Table I. The mean gestational age was 39.3±1.2 weeks, and 51.3% of the

infants were male. The mean birth weight and length were 3,302.2±397.8 g and 49.0±1.9 cm, respectively. Cesarean delivery accounted for 49% of births, of which 65% were elective. Mean maternal age was 27.7±6.3 years, parity ranged from one to seven, and 15% of mothers reported smoking during pregnancy.

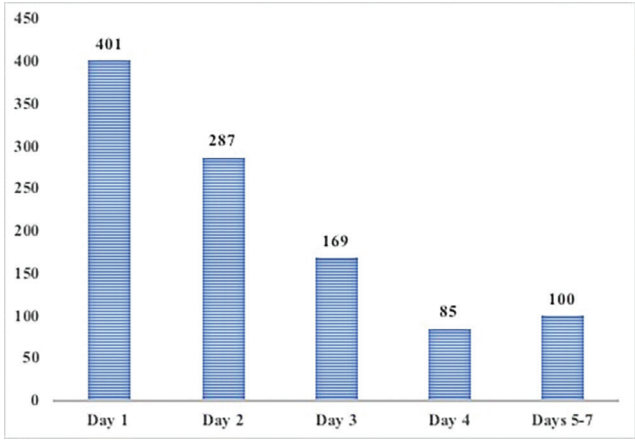

Figure 1. Number of infants measured per day (Days 5-7 combined)

Table I. Demographic and maternal characteristics of the study population (n=419)

Variable	Value	Min-Max
Neonatal characteristics		
Birth weight (g)	3,302.2± 397.8	2,520-4,580
Birth length (cm)	49.0±1.9	44-56
Male sex, n (%)	215 (51.3)	-
Gestational age (weeks)	39.3±1.2	37-42
1-min Apgar score	7.8±0.8	4-9
5-min Apgar score	9.4±0.6	7-10
Maternal characteristics		
Caesarean section, n (%)	209 (49.9)	-
Age (years)	27.8±6.2	15-50
Parity	2.1±1.1	1-7
Weight in the 1st trimester (kg)	65.1±25.6	40-120
SBP in the 1st trimester (mmHg)	103.3±10.5	70-140
DBP in the 1st trimester (mmHg)	65.2±9.2	50-110
Weight near delivery (kg)	76.0±12.2	45-120
SBP near birth (mmHg)	114.4±12.6	80-180
DBP near birth (mmHg)	72.9±9.4	37-120
Smoking during pregnancy, n (%)	61 (14.6)	-

SBP: Systolic blood pressure, DBP: Diastolic blood pressure

The percentile distributions of systolic, diastolic, and mean BP are illustrated in Figures 2-4. As shown in Table II, MBP values increased significantly across the first seven postnatal days (p=0.0001). Post-hoc pairwise comparisons (Table III) demonstrated significant differences between Day 1 and all subsequent days (p<0.05), indicating a rapid

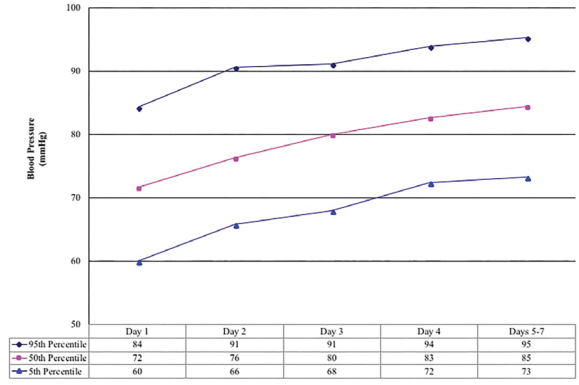

Figure 2. Systolic blood pressure percentiles of healthy term Turkish newborns during the first week of life

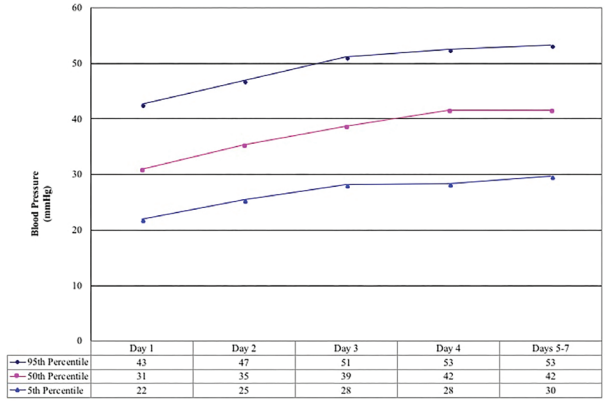

Figure 3. Diastolic blood pressure percentiles of healthy term Turkish newborns during the first week of life

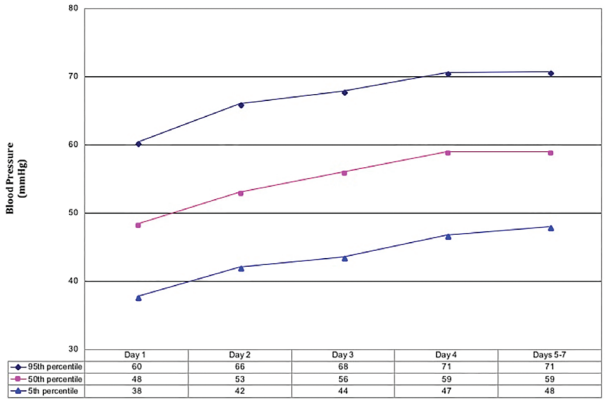

Figure 4. Mean blood pressure percentiles of healthy term Turkish newborns during the first week of life

Table II. Changes in mean blood pressure over the first five postnatal days

Day	Mean±SD (mmHg)	Median (IQR) (mmHg)
Day 1	48.1±7.35	48.33 (43.67-52.67)
Day 2	53.02±7.78	53.00 (48.33-57.67)
Day 3	55.91±7.52	56.00 (50.33-61.33)
Day 4	58.84±6.90	59.00 (54.17-63.84)
Days 5-7	59.23±6.79	59.00 (55.00-64.00)

Mean blood pressure increased significantly over the first five postnatal days (repeated measures ANOVA, F=12.27, p<0.001)
SD: Standard deviation, IQR: Interquartile range

Table III. Pairwise comparisons of mean blood pressure across postnatal days

Comparison	p value
Day 1 vs. Day 2	**0.013**
Day 1 vs. Day 3	**0.017**
Day 1 vs. Day 4	**<0.001**
Day 1 vs. Days 5-7	**<0.001**
Day 2 vs. Day 3	0.785
Day 2 vs. Day 4	**0.026**
Day 2 vs. Days 5-7	**0.003**
Day 3 vs. Day 4	0.056
Day 3 vs. Days 5-7	**0.002**
Day 4 vs. Days 5-7	0.358

Pairwise comparisons were performed using the Newman-Keuls multiple comparison test. Bold values indicate statistically significant correlations (p<0.05)

early rise. Beyond Day 2, significant increases were generally observed only on alternate days.

No statistically significant differences in MBP were detected by mode of delivery, sex, parity, or maternal smoking status. However, those infants born vaginally tended to have slightly higher MBP (by 1-2 mmHg) than those who were delivered by cesarean section on Days 1 and 2 (48.72±8.22 vs. 47.49±6.34 mmHg, p=0.09; and 54.13±7.37 vs. 52.33±7.97 mmHg, p=0.05, respectively).

Pearson correlation analyses revealed significant associations between neonatal MBP and several perinatal factors on Days 1 and 2 (Table IV). On Day 1, MBP correlated positively with birth weight (r=0.138, p=0.006), length (r=0.172, p=0.001), maternal near-birth SBP (r=0.259, p<0.001), maternal near-birth DBP (r=0.244, p<0.001), and 5th minute Apgar score (r=0.141, p=0.005). On Day 2, significant correlations persisted with gestational age (r=0.212, p<0.001), birth weight (r=0.165, p=0.005), and maternal near-birth SBP (r=0.132, p=0.032). From Day 3 onward, correlations weakened and lost statistical significance.

Multivariate stepwise linear regression analysis (Table V) identified maternal SBP near birth (β=0.143, p<0.001), birth length (β=0.605, p=0.001), and 5th minute Apgar score (β=1.26, p=0.007) as independent predictors of neonatal MBP on Day 1. Although statistically significant, the model explained only 0.1% of the variance (Adjusted R2=0.001), indicating that these predictors account for a minimal proportion of BP variability and have limited clinical relevance.

Table IV. Correlation between neonatal mean blood pressure and perinatal variables (r, p)

Variable	Day 1	Day 2	Day 3	Day 4	Days 5-7
Birth order	-0.008, 0.880	0.013, 0.827	0.005, 0.951	-0.063, 0.566	-0.040, 0.690
Birth weight	**0.138, 0.006**	**0.165, 0.005**	0.026, 0.737	0.013, 0.908	0.093, 0.359
Length	**0.172, 0.001**	0.095, 0.109	0.011, 0.886	0.017, 0.881	0.015, 0.885
Gestational age	0.043, 0.395	**0.212, <0.001**	0.125, 0.104	0.065, 0.553	-0.025, 0.808
Pre-delivery weight	0.081, 0.123	**0.135, 0.028**	0.038, 0.646	0.013, 0.910	0.029, 0.778
Pre-delivery SBP	0.098, 0.064	0.119, 0.056	-0.108, 0.185	-0.005, 0.964	0.110, 0.296
Pre-delivery DBP	**0.169, 0.001**	0.107, 0.085	-0.003, 0.968	0.086, 0.450	0.178, 0.089
Delivery weight	0.090, 0.089	**0.170, 0.006**	0.018, 0.826	0.096, 0.391	0.009, 0.928
Near birth SBP	**0.259, <0.001**	**0.132, 0.032**	**0.170, 0.035**	0.161, 0.152	0.166, 0.111
Near birth DBP	**0.244, <0.001**	0.074, 0.228	**0.196, 0.015**	0.139, 0.215	0.093, 0.374
Maternal age	0.091, 0.100	-0.020, 0.757	0.031, 0.727	-0.089, 0.483	0.071, 0.532
Apgar score (1 min)	0.023, 0.652	-0.005, 0.930	-0.015, 0.851	-0.112, 0.307	-0.025, 0.809
Apgar score (5 min)	**0.141, 0.005**	0.063, 0.293	-0.019, 0.810	-0.028, 0.797	0.019, 0.848

Bold values indicate statistically significant correlations (p<0.05)
SBP: Systolic blood pressure, DBP: Diastolic blood pressure

Table V. Multivariate stepwise linear regression analysis of day 1 mean blood pressure and perinatal parameters

Predictor variable	β coefficient	t value	p value
Constant	9.92	5.07	**0.0001**
Mothers' near birth SBP	0.143	4.95	**0.0001**
Birth length	0.605	3.21	**0.001**
5th minute Apgar score	1.26	2.74	**0.007**

Adjusted R^2=0.001, overall model p=0.001
SBP: Systolic blood pressure

Discussion

This single-center study provides preliminary reference ranges for BP in healthy term Turkish newborns during the first week of life, which may contribute to future multicenter studies. These values are broadly comparable to those reported in previous studies conducted in different ethnic populations. Notably, we observed a significant increase in BP, particularly between the first and second days of life. While maternal near-birth SBP, neonatal birth length, and 5th-minute Apgar scores were associated with MBP on the first day, these associations diminished in the subsequent days.

BP monitoring is one of the cornerstones of neonatal care, yet interpretation remains challenging due to variability in gestational age, birth weight, sex, and methodological differences. Only a limited number of studies have focused specifically on healthy term neonates (7,8,12). Kent et al. (7) reported median MBP values of 48-54 mmHg across the first four days in Australian infants, while Sadoh (8) found mean values of 48-52 mmHg in Nigerian neonates. Samanta et al. (12) reported a mean MBP of 60 mmHg on day 4 in a large cohort of term infants. Our values (48, 53, 56, and 59 mmHg) were slightly higher than those of Sadoh (8), but comparable to Samanta et al. (12). Differences may reflect ethnicity, maternal BP, environmental conditions, or device selection.

Consistent with prior research (6-8,12), we observed a significant rise in MBP during the early neonatal period, especially between the first and second days. Zubrow et al. (6) demonstrated a continuous increase in BP over the first five days across a range of gestational ages, independent of birth weight. This trend likely reflects the physiological transition from fetal to extrauterine circulation.

Although our data suggested a trend toward lower MBP in those infants born via cesarean section, this difference was not statistically significant. This contrasts with the findings from Salihoğlu et al. (11), who reported lower BP in neonates assessed within minutes of birth, likely reflecting the transient effects of anesthesia. In order to minimize confounding from perinatal asphyxia, we included only those infants with 5th minute Apgar scores ≥7. Interestingly, we found a correlation between 5th-minute Apgar scores and MBP on the first day. Given that Apgar scoring reflects neurobehavioral and physiological adaptation to extrauterine life, this association may reflect individual variability in neonatal cardiovascular responsiveness. However, we found no comparable data in the literature to support or refute this finding.

Sex-based differences in BP are well-documented in older populations, but remain inconclusive in neonates. While some studies have reported higher values in female newborns (11,13), others found no significant differences (6-8). The absence of pronounced hormonal influences in the neonatal period may explain the lack of consistent sex-related differences.

We identified a significant correlation between birth length and MBP on the first day. Kent et al. (7) did not report any associations between birth weight or length and BP, possibly due to their use of categorical variables. Sadoh (8) similarly categorized birth weights and found that large-for-gestational-age infants had higher BP than their small-for-gestational-age counterparts. Zubrow et al. (6) demonstrated that gestational age and birth weight were associated with SBP and DBP on day one, but post-conceptional age emerged as the primary determinant over time. Salihoğlu et al. (11) also found associations between BP and anthropometric measures including birth weight, length, and head circumference. In our study, birth length was significantly associated with MBP only on the first day. The lack of association with gestational age in this study may be due to our exclusive focus on term infants. We hypothesize that while maturation plays a role, other anthropometric or physiological factors may contribute to BP variability in healthy neonates.

In order to minimize confounding from maternal health conditions, we excluded those infants born to mothers with any known disease. We found a significant association between maternal near-birth SBP and neonatal MBP on the first day. Although BP tracking from infancy to adulthood has been documented (14-16), neonatal BP has not been shown to predict values at 6 months or 1 year (17). Neonatal BP appears to be more closely linked to perinatal and prenatal factors. Gillman et al. (18) described this as a "neonatal phenomenon," highlighting the influence of maternal age, third-trimester BP, birth weight, and postnatal age. They speculated that advanced maternal age may be associated with cardiovascular risk factors which affect placental

and fetal vascular development. Our findings support this hypothesis, demonstrating a link between maternal BP and neonatal MBP.

We used oscillometric devices, consistent with standard neonatal practice. While these devices have known limitations (2,19), invasive measurement was not ethically feasible in healthy infants. Adherence to published guidelines (4,5) minimized error, and our results remain clinically relevant given the widespread use of oscillometry.

Study Limitations

Several limitations should be acknowledged. First, attrition reduced sample size after Day 2, thus limiting generalizability. Second, the data were collected in 2012. No newer Turkish neonatal BP data exist, and many NICUs still rely on reference curves from 1995 (6). Thus, our findings remain valuable for current practice. Third, correlation coefficients were consistently below 0.3, indicating weak associations with perinatal factors. Finally, our descriptive percentile tables provide practical insight but lack the robustness of advanced modeling approaches (e.g., GAMLSS).

Conclusion

We present descriptive BP values for healthy term Turkish newborns during the early neonatal period. Maternal and perinatal factors influenced BP, particularly on the first day, but their clinical impact appears limited. Future multi-center studies with larger, contemporary cohorts and advanced statistical modeling are needed in order to establish robust reference curves and clarify long-term implications.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the University of Health Sciences Türkiye, Şişli Hamidiye Etfal Training and Research Hospital Clinical Research Ethics Committee (approval no.: 33, date: 10.04.2012).

Informed Consent: Informed parental consent was secured.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ö.G., Concept: Ö.G., S.U., A.B., Design: Ö.G., S.U., A.B., Data Collection or Processing: Ö.G., E.K.B., M.Ç., U.Z., Analysis or Interpretation: Ö.G., E.K.B., M.Ç., U.Z., Literature Search: Ö.G., Writing: Ö.G., S.U., A.B.

Conflict of Interest: The authors declare no conflicts of interest.

Financial Disclosure: The authors received no financial support for the conduct, authorship, or publication of this study.

References

1. Dore R, Barnes K, Bremner S, et al. Neonatal blood pressure by birth weight, gestational age, and postnatal age: a systematic review. Matern Health Neonatol Perinatol. 2024; 10:9.
2. Lalan SP, Warady BA. Discrepancies in the normative neonatal blood pressure reference ranges. Blood Press Monit. 2015; 20:171-7.
3. Nuntnarumit P, Yang W, Bada-Ellzey HS. Blood pressure measurements in the newborn. Clin Perinatol. 1999; 26:981.
4. Nwankwo M, Lorenz J, Gardiner J. A standard protocol for blood pressure measurement in the newborn. Pediatrics.1997; 6:E10.
5. Stebor DA. Basic principle of noninvasive blood pressure measurement in infants. Adv Neonatal Care. 2005; 5:252-61.
6. Zubrow A, Hulman S, Kushner H, Falkner B. Determinants of blood pressure in infants admitted to neonatal intensive care units: a prospective multicenter study. Philadelphia Neonatal Blood Pressure Study Group. J Perinatol. 1995 ;6:470-9.
7. Kent A, Kecskes Z, Shadbolt B, Falk M. Normative blood pressure data in the early neonatal period. Pediatr Nephrol. 2007; 9:1335-41.
8. Sadoh WE IS. Oscillometric blood pressure reference values of African full-term neonates in their first days postpartum. Cardiovasc J Afr. 2009; 20:344-7.
9. Barker DJP, Fall CHD. Fetal and infant origins of cardiovascular disease. Arch Dis Child. 1993; 68:797-9.
10. Perk O, Çakmak FN, Aliefendioğlu D. Determination of blood pressure in newborn diseases. Akd Med J. 2023; 9:271-6.
11. Salihoğlu O, Can E, Beşkardeş A, et al. Delivery room blood pressure percentiles of healthy, singleton, liveborn neonates. Pediatr Int. 2012; 2:182-9.
12. Samanta M, Mondal R, Ray S, et al. Normative blood pressure data for Indian neonates. Indian Pediatr. 2015; 52:669-73.
13. Emery E, Greenough A, Yuksel B. Effect of gender on blood pressure levels of very low birthweight infants in the first 48 hours of life. Early Hum Dev. 1993; 3:209-16.
14. Chen X, Wang Y. Tracking of blood pressure from childhood to adulthood: a systematic review and meta-regression analysis. Circulation. 2008; 25:3171-80.
15. Zinner S, Rosner B, Oh W, Kass E. Significance of blood pressure in infancy. Familial aggregation and predictive effect on later blood pressure. Hypertension. 1985;(3 Pt 1):411-6.

16. de Swiet M, Fayers P, Shinebourne E. Blood pressure survey in a population of newborn infants. Br Med J. 1976; 6026:9-11.

17. Kent A, Kecskes Z, Shadbolt B, Falk M. Blood pressure in the first year of life in healthy infants born at term. Pediatr Nephrol. 2007; 10:1743-9.

18. Gillman MW, Rich-Edwards JW, Rifas-Shiman SL, Lieberman ES, Kleinman KP, Lipshultz SE. Maternal age and other predictors of newborn blood pressure. J Pediatr. 2004; 2:240-5.

19. Dannevig I, Dale H, Liestø lK, Lindemann R. Blood pressure in the neonate: three non-invasive oscillometric pressure monitors compared with invasively measured blood pressure. Acta Paediatr. 2005; 2:191-6.

J Pediatr Res 2026;13(3):227-37
DOI: 10.4274/jpr.galenos.2026.29599

Original Article

Beyond Infections: Clinical and Genetic Spectrum of Pediatric Immune Dysregulation Disorders

Figen Çelebi Çelik, Aymen Hişmioğulları, Soner Günder, Gülçin Kaymakoğlu, Nesrin Gülez, Ferah Genel

University of Health Sciences Türkiye, İzmir Faculty of Medicine, Dr. Behçet Uz Pediatric Diseases and Surgery Training and Research Hospital, Department of Pediatrics, Division of Pediatric Allergy and Immunology, İzmir, Türkiye

ABSTRACT

Aim: Primary immune regulatory disorders represent a rapidly expanding subgroup of inborn errors of immunity. Unlike classical infection-predominant immunodeficiencies, these disorders may initially present with autoimmunity, lymphoproliferation, cytopenias, dermatologic disease, allergy, enteropathy, hemophagocytic lymphohistiocytosis, or malignancy.

Materials and Methods: We conducted a retrospective single-center observational cohort study of children with genetically confirmed diseases of immune dysregulation who had been followed at a tertiary pediatric immunology center between 2005 and 2025. Patients were included if their molecular diagnosis was classified under the International Union of Immunological Societies (IUIS) category of diseases of immune dysregulation. Demographic, clinical, genetic, therapeutic, transplant-related, and outcome data were extracted from the medical records and analyzed descriptively.

Results: Twenty-one children were included. Parental consanguinity was frequent (71.4%), and more than half of the cohort had a family history of primary immunodeficiency (57.1%). A substantial diagnostic delay was observed, with a median delay of 28 months (interquartile range, 8-105.5). The most common IUIS subcategory was regulatory T-cell defects (52.4%), followed by familial hemophagocytic lymphohistiocytosis syndromes with hypopigmentation (19.0%), autoimmune lymphoproliferative syndrome (14.3%), susceptibility to Epstein-Barr virus and lymphoproliferative conditions (9.5%), and immune dysregulation with colitis (4.8%). The main clinical features were lymphoproliferation (76.2%), hematologic abnormalities (66.7%) and autoimmunity (52.4%). Notably, malignancy was documented in 2 patients (9.5%). Antimicrobial prophylaxis was administered in 90.5% of the patients, immunoglobulin replacement in 66.7%, conventional immunosuppressive therapy in 38.1%, and biological or targeted therapy in 33.3%. Hematopoietic stem cell transplantation (HSCT) was performed in 8 patients (38.1%); immune reconstitution was achieved in all of the surviving transplant patients.

Conclusion: Pediatric diseases of immune dysregulation frequently present with non-infectious manifestations, particularly lymphoproliferation, cytopenias, autoimmunity, and dermatologic findings. Recognition beyond infection-centered warning signs, the integration of genetic testing, and the individualized use of targeted therapies or HSCT may improve care for this heterogeneous group of disorders.

Keywords: Primary immune regulatory disorders, diseases of immune dysregulation, inborn errors of immunity, autoimmunity, lymphoproliferation

Corresponding Author
Figen Çelebi Çelik, MD, University of Health Sciences Türkiye, İzmir Faculty of Medicine, Dr. Behçet Uz Pediatric Diseases and Surgery Training and Research Hospital, Department of Pediatrics, Division of Pediatric Allergy and Immunology, İzmir, Türkiye
E-mail: figencelebicelik@gmail.com **ORCID:** orcid.org/0000-0002-1100-4117
Received: 09.05.2026 **Accepted:** 11.06.2026 **Epub:** 31.07.2026 **Publication Date:** 18.09.2026

Cite this article as: Çelebi Çelik F, Hişmioğulları A, Günder S, Kaymakoğlu G, Gülez N, Genel F. Beyond infections: clinical and genetic spectrum of pediatric immune dysregulation disorders. J Pediatr Res. 2026;13(3):227-37

Introduction

Inborn errors of immunity (IEIs) comprise a rapidly expanding group of monogenic disorders which affect immune development, immune function, and immune homeostasis. The most recent update from the International Union of Immunological Societies (IUIS) classified IEIs across multiple disease categories and reported 508 disease-associated genes and 17 phenocopies, highlighting the accelerating pace of gene discovery in this field (1). Within this evolving classification, diseases of immune dysregulation represent one of the most dynamic and clinically challenging IEI subgroups, encompassing defects which impair immune tolerance, lymphocyte apoptosis, regulatory T-cell function, cytotoxic lymphocyte activity, cytokine signaling, and host control of lymphoproliferation (1,2).

Primary immune regulatory disorders (PIRDs) are defined by impaired immune homeostasis rather than isolated susceptibilities to infection. In children, they may present with autoimmunity, autoinflammation, lymphoproliferation, enteropathy, endocrinopathy, severe allergic or dermatologic disease, hemophagocytic lymphohistiocytosis, or malignancy, with or without recurrent infections (2,3). Therefore, non-infectious findings such as autoimmune cytopenias, persistent lymphadenopathy or splenomegaly, early-onset enteropathy, severe eczema or allergy, or unexplained hematologic abnormalities should raise suspicion for an underlying monogenic immune regulatory defect (2-4).

The clinical recognition of PIRDs has expanded substantially with the growing use of next-generation sequencing, particularly whole-exome sequencing (WES), which is now widely integrated into the diagnostic work-up of those children with suspected IEI. Genomic testing has not only expanded the known mutational spectrum of immune dysregulation disorders, but it has also enabled earlier and more precise diagnosis in children with overlapping immune phenotypes (5). This is particularly important in highly consanguineous populations, where autosomal recessive immune dysregulation disorders may be enriched and affected children may present with severe, early-onset, multisystem disease (6). Despite this progress, diagnostic delay remains common, partly because the initial presentation may mimic hematologic, rheumatologic, gastroenterologic, dermatologic, allergic, or oncologic disorders rather than a classical primary immunodeficiency (1-4).

Management of PIRDs differs substantially from that of conventional infection-predominant immunodeficiencies. In addition to antimicrobial prophylaxis and immunoglobulin replacement, many patients require corticosteroids, steroid-sparing immunosuppressive agents, biologic therapies, or mechanism-based targeted treatments (2,7,8). In selected severe disorders, hematopoietic stem cell transplantation (HSCT) remains a potentially curative option, although transplant decisions require careful consideration of genotype, disease severity, organ damage, infection burden, and treatment response (7,8).

As PIRDs are rare and phenotypically heterogeneous, single center and registry-based pediatric cohorts remain valuable in defining real-world clinical patterns, diagnostic pathways, treatment requirements, and outcomes. Such studies are particularly important in increasing awareness regarding the consideration of immune dysregulation disorders. In this study, we aimed to describe the clinical, genetic, therapeutic, and outcome characteristics of children with genetically confirmed diseases of immune dysregulation who had been followed at a tertiary pediatric immunology center.

Materials and Methods

Study Design and Setting

This retrospective single-center observational cohort study was conducted at University of Health Sciences Türkiye, Dr. Behçet Uz Pediatric Diseases and Surgery Training and Research Hospital. The medical records of those children diagnosed with genetically confirmed immune dysregulation disorders were retrospectively reviewed. This study aimed to evaluate the demographic characteristics, clinical presentation, genetic etiologies, disease course, treatment modalities, and outcomes of those children with immune dysregulation disorders. Ethical approval was obtained from the Clinical Research Ethics Committee of University of Health Sciences Türkiye, Dr. Behçet Uz Pediatric Diseases and Surgery Training and Research Hospital (approval no.: 511, date: 11.02.2021). This study was conducted in accordance with the principles of the Declaration of Helsinki. Due to the retrospective design of this study, informed consent requirements were obtained according to institutional ethics committee regulations.

Study Population

The source population was identified retrospectively from the departmental records of those patients followed with suspected IEIs between 2005 and 2025. During this period, the records of 1,586 patients with suspected IEIs were reviewed. Among them, 265 patients had a genetically

confirmed diagnosis of IEIs. Of these genetically confirmed patients, 21 patients fulfilled the inclusion criteria by having a genetically confirmed IEI classified under the IUIS category of "Diseases of Immune Dysregulation" according to the IUIS Primary Immunodeficiency Diseases Committee Report on IEI (2024) classification and so they were included in the final analysis (1). Diagnoses were based on compatible clinical manifestations and molecular confirmation by genetic testing.

Patients were eligible for inclusion if they fulfilled all of the following criteria:

- Age below 18 years at symptom onset or diagnosis;
- A genetically confirmed diagnosis which was included in the IUIS category of Diseases of Immune Dysregulation;
- Clinical features compatible with immune dysregulation, such as autoimmunity, autoinflammation, lymphoproliferation, hemophagocytic lymphohistiocytosis, severe or recurrent infections, growth failure, dermatologic manifestations, or hematologic abnormalities;
- Availability of sufficient clinical follow-up data.

Patients were excluded if they had incomplete medical records preventing clinical classification or if molecular confirmation was unavailable. Patients with immune dysregulation-like clinical manifestations but without a genetically confirmed diagnosis within the IUIS Diseases of Immune Dysregulation category were not included in this study.

Data Collection

Data were retrospectively extracted from the hospital records and departmental follow-up files of the primary immunodeficiency patients followed up during the period 2005-2025. The following variables were recorded: sex, current age or age at death, age at first symptoms, age at diagnosis of primary immunodeficiency, diagnostic delay, parental consanguinity, family history of primary immunodeficiency, presenting symptoms, genetic diagnosis, clinical findings, complications, treatments, HSCT, and survival status.

Genetic Diagnosis

Genetic diagnoses were obtained from molecular testing results available in the patient records. The patients were categorized according to the affected gene and related immune pathway where applicable. The cohort included those patients with pathogenic or clinically relevant variants in genes associated with immune dysregulation. Genetic results were interpreted together with clinical and immunologic findings by pediatric immunology specialists.

Definitions of Clinical Manifestations

Lymphoproliferation was defined as the presence of persistent or recurrent splenomegaly, hepatomegaly, and/or pathological lymphadenopathy documented by physical examination and/or imaging. Hematologic manifestations included thrombocytopenia, immune thrombocytopenia, autoimmune hemolytic anemia, neutropenia, lymphopenia, or other clinically relevant cytopenias recorded during presentation or follow-up. Autoimmunity was defined as physician-diagnosed autoimmune disease, including autoimmune cytopenias, autoimmune endocrinopathy, or other organ-specific autoimmune manifestations. Infectious complications were defined as clinically significant infections, or those resulting in sequelae such as bronchiectasis; opportunistic, viral, fungal, and mycobacterial infections were also recorded. Immune reconstitution after HSCT was defined according to transplant follow-up records based on survival after transplantation with donor-derived hematopoietic recovery and sufficient immune recovery for routine outpatient follow-up.

Treatment Data

Treatment modalities were recorded for each patient. These included antimicrobial prophylaxis, immunoglobulin replacement therapy, conventional immunosuppressive therapy, biological or targeted agents, and HSCT. HSCT was recorded with the underlying diagnosis and survival status when available.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 22.0. Categorical variables are summarized as frequencies and percentages, and continuous variables are reported as median and interquartile range (IQR). Percentages were calculated based on the available data for each variable. As this study was designed to describe the clinical, genetic, therapeutic, and outcome spectrum of genetically confirmed diseases of immune dysregulation, no comparative subgroup analysis or multivariable regression was performed.

Results

Demographic Characteristics of the Cohort

Of the 1,586 patients evaluated with suspected primary immunodeficiency, 265 had a genetically confirmed diagnosis. Among these, 21 patients fulfilled the IUIS criteria for "Diseases of Immune Dysregulation" and were included in the final cohort (Figure 1).

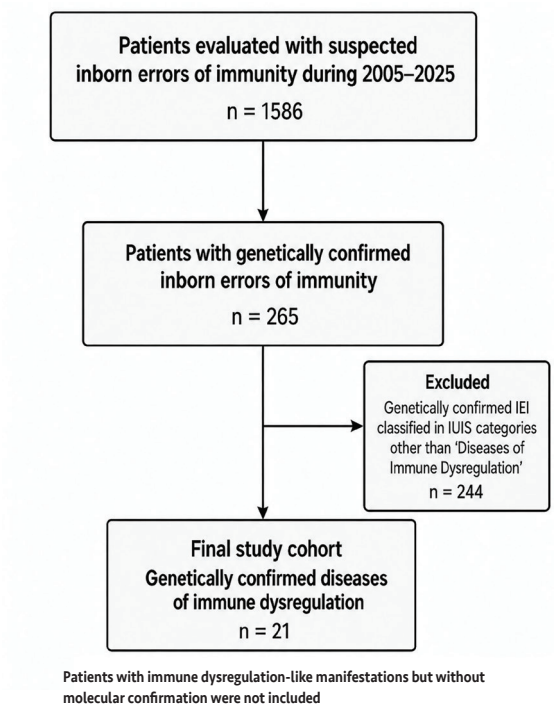

Figure 1. Patient selection flowchart

Of these, 10 patients were female (47.6%) and 11 were male (52.4%). Parental consanguinity was present in 15 patients (71.4%). A family history of primary immunodeficiency was documented in 12 patients (57.1%). Among the surviving patients, the median current age was 193 months (IQR, 134.25-257.5 months). The median age at first disease-related symptoms was 14 months (IQR, 2-49.5 months), whereas the median age at diagnosis of primary immunodeficiency was 91 months (IQR, 19.5-160.5 months). The median diagnostic delay was 28 months (IQR, 8-105.5 months). The median follow-up duration was 98 months (IQR, 63-138 months).

Distribution According to IUIS Diseases of Immune Dysregulation Subcategories

According to the IUIS classification, all patients had genetically confirmed disorders listed under "Diseases of Immune Dysregulation". The cohort represented five IUIS subcategories: familial hemophagocytic lymphohistiocytosis (FHL) syndromes with hypopigmentation, regulatory T-cell defects, immune dysregulation with colitis, autoimmune lymphoproliferative syndrome (ALPS), and susceptibility to Epstein-Barr virus (EBV) and lymphoproliferative conditions (Figure 2). The largest subgroup was regulatory T-cell defects, identified in 11 patients (52.4%). This group included LRBA deficiency in 5 patients, CTLA4 haploinsufficiency in 2 patients, FOXP3-related IPEX syndrome in 2 patients, and NBEAL2 deficiency in 2 patients. In the IUIS classification, *CTLA4, LRBA, FOXP3*, and *NBEAL2* are listed under the regulatory T-cell defect category. The second subgroup was FHL syndromes with hypopigmentation observed in 4 patients (19%). This subgroup included RAB27A deficiency/Griscelli syndrome type-2 in 2 patients, LYST deficiency/Chediak-Higashi syndrome in 1 patient, and AP3B1 deficiency/Hermansky-Pudlak syndrome type-2 in 1 patient. ALPS-related defects were present in 3 patients (14.3%). This subgroup included

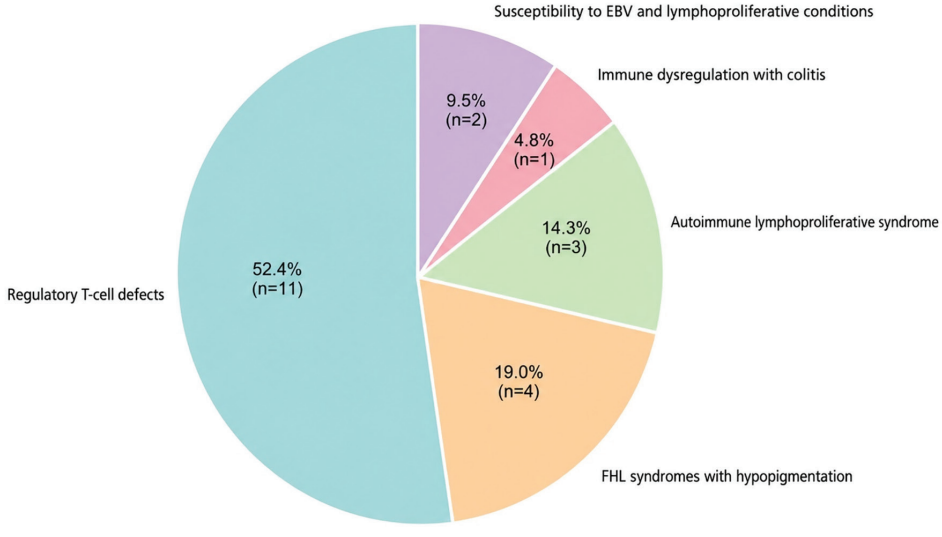

Figure 2. Distribution of patients according to IUIS diseases of immune dysregulation subcategories
EBV, Epstein-Barr virus, FHL: Familial hemophagocytic lymphohistiocytosis, IUIS: International Union of Immunological Societies

TNFRSF6/FAS-related ALPS in 2 patients and CASP10 deficiency in 1 patient. Immune dysregulation with colitis was represented by 1 patient (4.8%) with IL10RB deficiency. Finally, susceptibility to EBV and lymphoproliferative conditions was present in 2 patients (9.5%), both with TNFRSF9/CD137 deficiency.

Clinical Manifestations at Diagnosis and During Follow-up

Lymphoproliferative manifestations were documented in 16 patients (76.2%). The recorded lymphoproliferative findings included splenomegaly in 16 patients (76.2%), hepatomegaly in 13 patients (61.9%), and pathological lymphadenopathy in 13 patients (61.9%); these findings were not mutually exclusive. Hematologic manifestations were observed in 14 patients (66.7%). At initial evaluation, thrombocytopenia was present in 7 patients (33.3%), 5 of whom had immune thrombocytopenia. During the entire disease course, thrombocytopenia was documented in 10 patients (47.6%), 7 of whom had immune thrombocytopenia. Other hematologic abnormalities included neutropenia in 5 patients (23.8%), autoimmune hemolytic anemia in 5 patients (23.8%), and lymphopenia in 4 patients (19%). Autoimmunity was documented in 11 patients (52.4%), mainly autoimmune cytopenias in 9 patients (42.9%) and autoimmune endocrinopathies in 3 patients (14.3%). Recurrent sinopulmonary infections were present at diagnosis in 9 patients (42.9%). Dermatologic manifestations were observed in 7 patients (33.3%). Hypopigmented or silvery hair was present in 4 patients with cytotoxic trafficking defects, including 2 patients with RAB27A deficiency/Griscelli syndrome, 1 with Hermansky-Pudlak syndrome, and 1 with Chediak-Higashi syndrome. Eczematous skin involvement was documented in 2 patients with FOXP3/IPEX syndrome, and psoriasis was observed in 1 patient with CTLA4 haploinsufficiency. Allergic manifestations were recorded in 2 patients (9.5%), both with FOXP3/IPEX syndrome and food allergy. In addition, 2 patients (9.5%) were diagnosed through family screening because of previously identified primary immunodeficiencies in affected family members.

During follow-up, immune dysregulation-related and infectious complications were common and frequently involved multiple organ systems. Infection-related complications were observed in 7 patients (33.3%) during follow-up. Bronchiectasis was the most frequent infectious complication and was documented in 6 patients (28.6%) with CTLA4 haploinsufficiency, LRBA deficiency, TNFRSF6/FAS related disease, FOXP3/IPEX syndrome, IL10RB deficiency, and TNFRSF9/CD137 deficiency, and chronic lung disease was recorded in 1 patient with FOXP3/ IPEX syndrome. Abscess, pyoderma, or fungal skin infections occurred in 3 patients (14.3%) with CASP10, IL10RB, and LRBA deficiencies. Additional infectious complications included latent tuberculosis infection in 2 patients with LRBA deficiency, EBV disease in 1 patient with TNFRSF9/CD137 deficiency, and CMV infection in 1 patient with LRBA deficiency.

Hemophagocytic lymphohistiocytosis was recorded in 3 patients (14.3%), including patients with LYST deficiency and RAB27A/Griscelli syndrome.

Malignancy was recorded in 2 patients (9.5%). One patient with CASP10 deficiency developed lymphoma at 18 years of age, and one patient with TNFRSF9/CD137 deficiency developed EBV-associated B-cell lymphoma at 11 years of age. These findings indicate that, in this cohort, non-infectious manifestations, particularly lymphoproliferation, hematologic abnormalities and autoimmunity, were more frequent presenting features than infectious symptoms (Figure 3).

Therapeutic Management and Final Outcomes

Most patients required infection prophylaxis and/or immunomodulatory treatment during follow-up. Antimicrobial prophylaxis was used in 19 patients (90.5%), and immunoglobulin replacement therapy was administered to 14 patients (66.7%).

Conventional immunosuppressive treatment was administered in 8 patients (38.1%). The recorded agents included systemic corticosteroids, cyclosporine, and sirolimus. Corticosteroids were mainly used for autoimmune cytopenias in patients with LRBA deficiency, TNFRSF6/FAS-related disease, TNFRSF9/CD137 deficiency, LYST deficiency/Chediak-Higashi syndrome, and CASP10 deficiency. Cyclosporine was used in one patient with TNFRSF9/CD137 deficiency, and sirolimus was used in one patient with CTLA4 haploinsufficiency. Biological agents were administered to 7 patients (33.3%). Abatacept was given to 6 patients, including 5 with LRBA deficiency and 1 with CTLA4 haploinsufficiency. In addition, 1 patient with TNFRSF9/CD137 deficiency received rituximab and bortezomib because of EBV-associated lymphoproliferation.

HSCT was performed in 8 patients (38.1%), including 2 patients with RAB27A deficiency/Griscelli syndrome, 2 with FOXP3/IPEX syndrome, and one each with LRBA deficiency, TNFRSF9/CD137 deficiency, LYST deficiency/Chediak-Higashi syndrome, and IL10RB deficiency. At the last follow-up, 5 of the 8 transplant patients (62.5%) were alive, all of whom

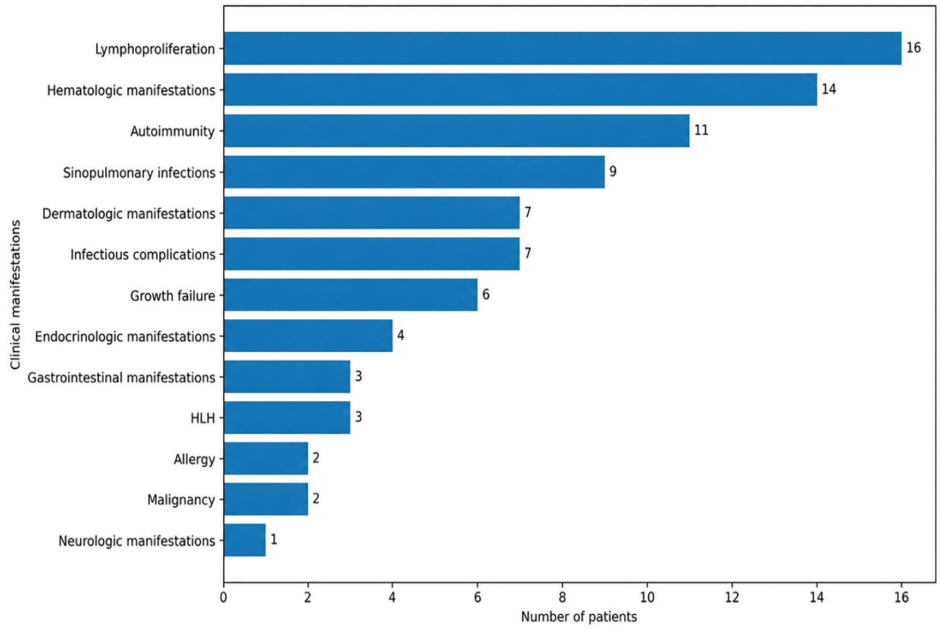

Figure 3. Presenting manifestations and immune dysregulation-related complications. The horizontal bar chart shows the frequency of major presenting manifestations and immune dysregulation-related complications in the study cohort
HLH: Hemophagocytic lymphohistiocytosis

had achieved immune reconstitution, while 3 had died. Deaths after HSCT occurred in patients with TNFRSF9/CD137 deficiency, LYST deficiency/Chediak-Higashi syndrome, and RAB27A deficiency/Griscelli syndrome. The patient with TNFRSF9/CD137 deficiency died due to septic shock.

At the last follow-up, 16 patients (76.2%) were alive, while 5 patients (23.8%) had died. The deceased patients had CTLA4 haploinsufficiency, LRBA deficiency, TNFRSF9/CD137 deficiency, LYST deficiency/Chediak-Higashi syndrome, and RAB27A deficiency/Griscelli syndrome. The median age at death was 68 months (IQR, 60-156 months). Comprehensive individual-level clinical, genetic, treatment-related, and outcome data for all of the patients are presented in Table I.

Table I. Clinical, genetic, therapeutic, and outcome characteristics stratified by IUIS diseases of immune dysregulation subcategories

No.	Genetic diagnosis/ disorder	Sex	Age at onset/ diagnosis, months	Clinical manifestations	Treatment	Outcome
FHL syndromes with hypopigmentation						
P10	RAB27A deficiency/ Griscelli syndrome type-2	M	12/22	Hypopigmented/silvery hair, hypothyroidism, hepatosplenomegaly, sinopulmonary infections, HLH	Antimicrobial prophylaxis, IVIG, HSCT	Alive; immune reconstitution after HSCT
P21	RAB27A deficiency/ Griscelli syndrome type-2	F	1/6	Hypopigmented/silvery hair, diarrhea, growth failure, HLH	Antimicrobial prophylaxis, IVIG, HSCT	Died after HSCT
P11	LYST deficiency/Chediak-Higashi syndrome	F	51/65	Hypopigmented/silvery hair, hepatosplenomegaly, neutropenia, thrombocytopenia, lymphopenia, sinopulmonary infections, HLH	Antimicrobial prophylaxis, corticosteroid, HSCT	Died after HSCT
P19	AP3B1 deficiency/ Hermansky-Pudlak syndrome type-2	M	1/7	Hypopigmented/silvery hair, sinopulmonary infections, growth failure, neutropenia, syndromic features, neuromotor retardation	Antimicrobial prophylaxis, IVIG	Alive

Table I. Continued

Regulatory T-cell defects						
P18	FOXP3 deficiency/IPEX syndrome	M	1/13	Eczema, growth failure, gastroenteritis, food allergy, chronic lung disease, bronchiectasis, eosinophilia	Antimicrobial prophylaxis, IVIG, HSCT	Alive; immune reconstitution after HSCT
P20	FOXP3 deficiency/IPEX syndrome	M	2/17	Eczema, growth failure, food allergy, eosinophilia	Antimicrobial prophylaxis, IVIG, HSCT	Alive; immune reconstitution after HSCT
P2	LRBA deficiency	F	6/199	Autoimmune hemolytic anemia, hepatosplenomegaly, lymphadenopathy, bronchiectasis	Antimicrobial prophylaxis, IVIG, abatacept	Died
P14	LRBA deficiency	M	12/82	Sinopulmonary infections, autoimmune hemolytic anemia, immune thrombocytopenia, splenomegaly, lymphadenopathy, latent tuberculosis infection	Antimicrobial prophylaxis, IVIG, corticosteroid, abatacept, HSCT	Alive; immune reconstitution after HSCT
No.	**Genetic diagnosis/ disorder**	**Sex**	**Age at onset/ diagnosis, months**	**Clinical manifestations**	**Treatment**	**Outcome**
P16	LRBA deficiency	M	14/185	Autoimmune hemolytic anemia, immune thrombocytopenia, chronic interstitial nephritis, splenomegaly	Antimicrobial prophylaxis, IVIG, corticosteroid, abatacept	Alive
P6	LRBA deficiency	F	3/107	Sinopulmonary infections, growth failure, Graves' disease, type 1 diabetes mellitus, hepatosplenomegaly, lymphadenopathy, CMV infection	Antimicrobial prophylaxis, IVIG, abatacept	Alive
P12	LRBA deficiency	M	2/40	Sinopulmonary infections, type 1 diabetes mellitus, Hashimoto thyroiditis, hepatosplenomegaly, lymphadenopathy, latent tuberculosis infection, pyoderma	Antimicrobial prophylaxis, IVIG, abatacept	Alive
P1	CTLA4 haploinsufficiency	F	48/54	Sinopulmonary infections, autoimmune hemolytic anemia, hepatosplenomegaly, lymphadenopathy	Antimicrobial prophylaxis, IVIG, abatacept	Died
P5	CTLA4 haploinsufficiency	M	24/130	Psoriasis, autoimmune hemolytic anemia, immune thrombocytopenia, hepatosplenomegaly, lymphadenopathy, bronchiectasis	Antimicrobial prophylaxis, IVIG, sirolimus	Alive
P7	NBEAL2 deficiency	M	70/129	ALPS-like disease, thrombocytopenia, hepatosplenomegaly, lymphadenopathy	Antimicrobial prophylaxis	Alive
P8	NBEAL2 deficiency	F	24/129	ALPS-like disease, thrombocytopenia, lymphopenia, hepatosplenomegaly, lymphadenopathy	Antimicrobial prophylaxis	Alive
Autoimmune lymphoproliferative syndrome						
P3	TNFRSF6/FAS-related ALPS	F	18/136	ALPS-like disease, immune thrombocytopenia, neutropenia, hypothyroidism, hepatosplenomegaly, lymphadenopathy, bronchiectasis	Antimicrobial prophylaxis, corticosteroid	Alive
P9	TNFRSF6/FAS-related ALPS	M	204/205	ALPS-like disease, immune thrombocytopenia, hepatosplenomegaly, lymphadenopathy	Antimicrobial prophylaxis, corticosteroid	Alive
P4	CASP10 deficiency	F	36/210	ALPS-like disease, neutropenia, immune thrombocytopenia, hepatosplenomegaly, lymphadenopathy, abscess, lymphoma	Corticosteroid, chemotherapy	Alive

Table I. Continued

Immune dysregulation with colitis						
P17	IL10RB deficiency	F	1/6	Sinopulmonary infections, growth failure, perianal abscess, rectovaginal fistula, colitis, bronchiectasis	Antimicrobial prophylaxis, IVIG, HSCT	Alive; immune reconstitution after HSCT
Susceptibility to EBV and lymphoproliferative conditions						
P15	TNFRSF9/CD137 deficiency	M	78/91	ALPS-like disease, immune thrombocytopenia, neutropenia, lymphopenia, splenomegaly, generalized lymphadenopathy, EBV-related B-cell lymphoma	Antimicrobial prophylaxis, IVIG, corticosteroid, cyclosporine, rituximab, bortezomib, chemotherapy, HSCT	Died after HSCT; due to septic shock
P13	TNFRSF9/CD137 deficiency	F	168/196	Sinopulmonary infections, hepatosplenomegaly, lymphopenia, lymphadenopathy, bronchiectasis	Denied treatment	Alive

ALPS: Autoimmune lymphoproliferative syndrome, CMV: Cytomegalovirus, EBV, Epstein-Barr virus, FHL: Familial hemophagocytic lymphohistiocytosis, HLH: Hemophagocytic lymphohistiocytosis, HSCT: Hematopoietic stem cell transplantation, IUIS: International Union of Immunological Societies, IVIG: Intravenous immunoglobulin

Discussion

In this single-center pediatric cohort, genetically confirmed diseases of immune dysregulation showed a broad and severe clinical spectrum, dominated by lymphoproliferation, hematologic manifestations, autoimmunity, and infection-related morbidity. The most frequent IUIS subcategory was regulatory T-cell defects, followed by FHL syndromes with hypopigmentation, autoimmune lymphoproliferative syndrome, immune dysregulation with colitis, and susceptibility to EBV and lymphoproliferative conditions. These findings support the argument that PIRDs are not merely infection-predominant immunodeficiencies, but also multisystem disorders in which failure of immune tolerance, abnormal lymphocyte activation, impaired apoptosis, defective cytotoxicity, and disturbed host–virus surveillance can shape the presenting phenotype (1-4).

The high rate of parental consanguinity in our cohort is consistent with the genetic architecture of IEI in populations where autosomal recessive disorders are enriched. In our series, consanguinity was present in 71.4% of the patients, and a family history of primary immunodeficiency was documented in more than half of the cohort. In the Turkish PIRD cohort reported by Aykut et al. (5), next-generation sequencing identified disease-causing variants across 15 PIRD genes, including several novel variants, further illustrating that PIRD cohorts from Türkiye and its neighboring regions may have a distinctive genetic burden related to consanguinity and founder effects. A recent national registry study from a highly consanguineous population reported parental consanguinity in 88.1% of pediatric PIRD cases and a family history of PIRDs in 45.7%, emphasizing the importance of family-based risk assessments in such settings (6).

Diagnostic delay remains a major challenge in PIRDs, largely because these disorders often present outside the classical infection-centered framework of primary immunodeficiency. In the national registry reported by Alajmi et al. (6), children with PIRDs were diagnosed relatively early, with a short median delay between symptom onset and diagnosis, probably reflecting registry-based ascertainment, high clinical awareness, and the predominance of severe early-onset phenotypes such as FHL-related disorders. However, our cohort suggests that children with broader immune dysregulation phenotypes may remain undiagnosed for a prolonged period, particularly when the initial presentation is dominated by hematologic, autoimmune, lymphoproliferative, dermatologic, allergic, or endocrine manifestations rather than recurrent infections. This observation is also consistent with the ESID Registry analysis of 16,486 patients with IEI which showed that non-infectious presentations are an important route to diagnosis and that relying exclusively on infection-centered warning signs would potentially miss a substantial proportion of these patients (9). Together, these data emphasize that the Jeffrey Modell warning signs, although useful for infection-predominant IEIs, are insufficient for detecting many PIRDs. Children with unexplained cytopenias, persistent lymphadenopathy or splenomegaly, early-onset autoimmunity, severe eczema or allergy, enteropathy, HLH-like inflammation, malignancy, or a positive family history should therefore be considered for early immunologic evaluation and WES-based genetic testing, even when their infection history is not striking (9-11).

The distribution of genetic diagnoses in our cohort showed that regulatory T-cell defects constituted the largest subgroup, including LRBA deficiency, CTLA4 haploinsufficiency, FOXP3/IPEX syndrome, and NBEAL2 deficiency. LRBA deficiency and CTLA4 haploinsufficiency are prototypic immune checkpoint disorders characterized by autoimmunity, lymphoproliferation, hypogammaglobulinemia, enteropathy, and recurrent infections (7,12,13). In line with previous reports showing that these disorders do not follow a uniform clinical pattern, our patients with LRBA deficiency and CTLA4 haploinsufficiency displayed overlapping CVID-like, ALPS-like, and autoimmune phenotypes (7,12,13). Although NBEAL2 deficiency has classically been recognized as the molecular cause of gray platelet syndrome, it is now classified among regulatory T-cell defects in the updated IUIS classification (1). A recent study showed reduced CTLA-4 expression in activated conventional T-cells from patients with NBEAL2 deficiency, providing a biological rationale for considering NBEAL2 deficiency as a CTLA-4 pathway-related immune dysregulation disorder (14).

Our cohort also illustrates the phenotypic diversity which may occur even within the same genotype or family. This was particularly evident in the siblings with TNFRSF9/CD137 deficiency. One sibling developed EBV-related B-cell lymphoma, required chemotherapy, rituximab, bortezomib, and HSCT, and died of septic shock after transplantation, whereas the other remains alive without disease-specific treatment. CD137 is critical for effective T-cell responses against EBV-infected cells, and TNFRSF9 deficiency has been associated with EBV-driven lymphoproliferation and lymphoma susceptibility (15,16). Such intrafamilial variability is increasingly recognized in IEIs and may reflect differences in viral exposure, immune maturation, modifying variants, somatic events, treatment timing, or other environmental factors (17). This sibling pair highlights that genotype alone may not be sufficient to predict disease severity and that long-term surveillance is warranted even in apparently mildly affected relatives.

The therapeutic profile of our cohort underscores the high morbidity of pediatric PIRDs. Management often extended beyond infection prevention and immunoglobulin replacement to include immunosuppressive agents, pathway-directed biological therapies, and, in selected severe cases, HSCT. Previous studies have shown that abatacept can control autoimmune and inflammatory manifestations in LRBA deficiency and CTLA4 insufficiency, although treatment responses may be incomplete and long-term therapy does not eliminate all risks, including infection, progressive organ damage, or malignancy (12,14,18). While CTLA-4 pathway disorders provide a clear example of mechanism-based replacement therapy with abatacept, other PIRDs require different individualized approaches according to the dominant pathogenic mechanism. In one patient with TNFRSF9/CD137 deficiency, rituximab and bortezomib were administered for EBV-associated lymphoproliferation, reflecting the need for tailored treatment strategies in EBV-driven immune dysregulation.

HSCT was performed in selected patients with severe disease phenotypes. More than half of these transplant patients were alive at the last follow-up, and immune reconstitution was achieved in all of the surviving transplant patients. However, three deaths occurred after HSCT, underscoring that transplantation can be curative but remains high-risk in children with advanced immune dysregulation, active infection, HLH, or severe organ involvement. HSCT is generally considered the definitive therapy for selected disorders such as FOXP3/IPEX syndrome, IL10RB deficiency with severe early-onset colitis, Griscelli syndrome type-2, Chediak-Higashi syndrome, and other cytotoxicity or trafficking defects with HLH susceptibility (19-22). However, HSCT is not uniformly indicated for all PIRD categories. In ALPS-FAS, for example, medical treatment with corticosteroids, mycophenolate mofetil, or sirolimus is generally preferred for autoimmune cytopenias and lymphoproliferation, while HSCT is exceptional (23). For LRBA deficiency and CTLA4 haploinsufficiency, HSCT may prevent ongoing disease progression in selected patients, but the decision requires careful assessment of disease activity, organ damage, infection burden, and response to CTLA-4-Ig or other immunomodulatory therapies (13,18,24).

Malignancy development in patients with PIRDs may result from chronic immune activation, defective apoptosis, impaired tumor immune surveillance, persistent viral stimulation, or a combination of these mechanisms, depending on the underlying genetic defect (25). In our cohort, lymphoma developed in two patients: one with CASP10 deficiency at 18 years of age and one with TNFRSF9/CD137 deficiency who developed EBV-associated B-cell lymphoma at 11 years of age. ALPS-related apoptosis pathway defects have long been associated with lymphoproliferation and lymphoma risk, whereas TNFRSF9/CD137 deficiency illustrates how impaired EBV immune control may predispose to EBV-associated lymphoproliferative disease and lymphomagenesis (15,16,25). Therefore, the follow-up of patients with PIRDs should not be limited to infection control or the management of acute autoimmune manifestations, but should also include long-term

surveillance for lymphoid malignancy, viral persistence, and progressive organ damage.

Study Limitations

The genetically confirmed diagnosis of all of the patients and their classification according to IUIS subcategories are among the main strengths of this study. The cohort provides real-world pediatric data from a tertiary immunology center in a population with a high rate of consanguinity and includes rare diagnoses such as NBEAL2 deficiency and TNFRSF9/CD137 deficiency. This study also reflects the therapeutic diversity of pediatric PIRDs, including conventional immunosuppression, CTLA-4-Ig therapy, anti-CD20/plasma cell- directed treatment, and HSCT. The retrospective design and the lack of standardized measures for disease activity, treatment response, and longitudinal organ involvement were the limitations of this study. The prospective use of tools such as IDDA2.1 scores can improve comparability among future PIRD cohorts (26). In addition, the absence of classical primary HLH genes such as *PRF1, UNC13D, STX11*, and *STXBP2* may reflect referral bias, as these patients are often diagnosed and managed primarily in pediatric hematology or transplant units. As pediatric diseases of immune dysregulation are rare and genetically heterogeneous, the number of patients included in this study was limited, which prevented robust subgroup comparisons, and the extended retrospective review period may have influenced case identification because diagnostic awareness, immunologic evaluation, and access to genetic testing improved over the study period.

Conclusion

In conclusion, this cohort shows that genetically confirmed pediatric diseases of immune dysregulation are characterized by early onset, substantial diagnostic delay, high consanguinity, frequent non-infectious manifestations, and a considerable need for advanced therapies. Lymphoproliferation, hematologic abnormalities, autoimmunity, dermatologic disease, HLH, and malignancy should prompt consideration of PIRDs, even when recurrent infections are not the dominant feature. Increasing awareness beyond infection-centered warning signs, integrating WES into the diagnostic pathway, and applying mechanism-based therapies or HSCT in carefully selected patients may improve outcomes in this expanding group of IEIs.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Clinical Research Ethics Committee of University of Health Sciences Türkiye, Dr. Behçet Uz Pediatric Diseases and Surgery Training and Research Hospital (approval no.: 511, date: 11.02.2021).

Informed Consent: Due to the retrospective design of this study, informed consent requirements were obtained according to institutional ethics committee regulations.

Footnotes

Authorship Contributions

Surgical and Medical Practices: F.Ç.Ç., N.G., F.G., Concept: F.Ç.Ç., A.H., Design: F.Ç.Ç., A.H., Data Collection or Processing: A.H., S.G., G.K., Analysis or Interpretation: F.Ç.Ç., A.H., Literature Search: F.Ç.Ç., A.H., Writing: F.Ç.Ç.

Conflict of Interest: The authors declare no conflict of interest.

Financial Disclosure: The authors declare that no financial or material support was received for this study.

References

1. Poli MC, Aksentijevich I, Bousfiha AA, 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.
2. Kolukısa B, Barış S. Primary immune regulatory disorders and targeted therapies. Turk J Haematol. 2021; 38:1-14.
3. Tsilifis C, Slatter MA, Gennery AR. Too much of a good thing: a review of primary immune regulatory disorders. Front Immunol. 2023; 14:1279201.
4. Chan AY, Torgerson TR. Primary immune regulatory disorders: a growing universe of immune dysregulation. Curr Opin Allergy Clin Immunol. 2020; 20:582-90.
5. Aykut A, Karakoc-Aydiner E, Kiykim A, et al. Primary immune regulatory disorders: expanding the mutation spectrum in Turkey and identification of sixteen novel variants. Immunol Res. 2024; 72:1129-41.
6. Alajmi A, Alsharidah S, Khalifa N, et al. Spectrum of primary immune regulatory disorders in children in a highly consanguineous population: report from a national registry. J Clin Immunol. 2026; 46:25.
7. Gámez-Díaz L, Seidel MG. Different apples, same tree: visualizing current biological and clinical insights into CTLA-4 insufficiency and LRBA and DEF6 deficiencies. Front Pediatr. 2021; 9:662645.
8. Makkoukdji N, Pundit V, Wyke M, et al. Targeted treatments for immune dysregulation in inborn errors of immunity. Explor Immunol. 2024; 4:310-29.

9. Thalhammer J, Kindle G, Nieters 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; 148:1332-41.e5.

10. Costagliola G, Peroni DG, Consolini R. Beyond infections: new warning signs for inborn errors of immunity in children. Front Pediatr. 2022; 10:855445.

11. Schiavo E, Martini B, Attardi E, et al. Autoimmune cytopenias and dysregulated immunophenotype act as warning signs of inborn errors of immunity: results from a prospective study. Front Immunol. 2022; 13:790455.

12. Jamee M, Hosseinzadeh S, Sharifinejad N, et al. Comprehensive comparison between 222 CTLA-4 haploinsufficiency and 212 LRBA deficiency patients: a systematic review. Clin Exp Immunol. 2021; 205:28-43.

13. Taghizade N, Babayeva R, Kara A, et al. Therapeutic modalities and clinical outcomes in a large cohort with LRBA deficiency and CTLA4 insufficiency. J Allergy Clin Immunol. 2023; 152:1634-45.

14. Delage L, Carbone F, Riller Q, et al. NBEAL2 deficiency in humans leads to low CTLA-4 expression in activated conventional T cells. Nat Commun. 2023; 14:3728.

15. Somekh I, Thian M, Medgyesi D, et al. CD137 deficiency causes immune dysregulation with predisposition to lymphomagenesis. Blood. 2019; 134:1510-6.

16. Zhao P, Chen K, Yang L, et al. Clinical and functional characterization of a novel TNFRSF9 variant causing CD137 deficiency and EBV-associated lymphoproliferative disease. Front Immunol. 2025; 16:1605221.

17. Similuk M, Kuijpers T. Nature and nurture: understanding phenotypic variation in inborn errors of immunity. Front Cell Infect Microbiol. 2023; 13:1183142.

18. Arnold DE, Chellapandian D, Leiding JW. The use of biologic modifiers as a bridge to hematopoietic cell transplantation in primary immune regulatory disorders. Front Immunol. 2021; 12:692219.

19. Kuskonmaz B, Ayvaz D, Gokce M, et al. Hematopoietic stem cell transplantation in children with Griscelli syndrome type 2: experience and outcomes. Pediatr Transplant. 2017; 21:e13041.

20. Barzaghi F, Amaya Hernandez LC, Neven B, et al. Long-term follow-up of IPEX syndrome patients after different therapeutic strategies: an international multicenter retrospective study. J Allergy Clin Immunol. 2018; 141:1036-49.e5.

21. Bacchetta R, Roncarolo MG. IPEX syndrome from diagnosis to cure, learning along the way. J Allergy Clin Immunol. 2024; 153:371-82.

22. Toro C, Nicoli ER, Malicdan MCV, Adams D, Introne WJ. Chediak-Higashi Syndrome. In: Adam MP, Feldman J, Mirzaa GM, et al. (eds). GeneReviews®. Seattle, University of Washington, 2023.

23. Rao VK, Pittaluga S, Uzel G. Beyond FAScinating: advances in diagnosis and management of autoimmune lymphoproliferative syndrome and activated PI3 kinase δ syndrome. Hematology Am Soc Hematol Educ Program. 2024; 2024:126-36.

24. Tsilifis C, Speckmann C, Lum SH, et al. Hematopoietic stem cell transplantation for CTLA-4 insufficiency across Europe: a European Society for Blood and Marrow Transplantation Inborn Errors Working Party study. J Allergy Clin Immunol. 2024; 154:1782-93.e7.

25. Oliveira JB, Bleesing JJ, Dianzani U, et al. Revised diagnostic criteria and classification for the autoimmune lymphoproliferative syndrome: report from the 2009 NIH International Workshop. Blood. 2010; 116:e35-40.

26. Seidel MG, Tesch VK, Yang L, et al. The immune deficiency and dysregulation activity (IDDA2.1 "Kaleidoscope") score and other clinical measures in inborn errors of immunity. J Clin Immunol. 2022; 42:484-98.

J Pediatr Res 2026;13(3):238-46
DOI: 10.4274/jpr.galenos.2026.56805

Original Article

Clinical and Electroencephalographic Characteristics of Childhood Occipital Epilepsies: Comparison of Self-Limited and Symptomatic Forms

Ayfer Arduç Akçay

Koç University Faculty of Medicine, Department of Pediatrics, Division of Pediatric Neurology, İstanbul, Türkiye

ABSTRACT

Aim: This study aimed to evaluate the clinical, electroencephalography, neuroimaging, and treatment characteristics of childhood occipital epilepsy spectrum disorders and to compare self-limited and symptomatic cases.

Materials and Methods: Sixty-four pediatric patients followed between October 2015 and January 2026 were retrospectively evaluated. The patients were classified into three groups according to clinical, developmental, and radiological findings: self-limited, symptomatic, and unclassified.

Results: Sixty-four patients were included, comprising 25 (39.1%) self-limited, 36 (56.3%) symptomatic/determinable, and 3 (4.7%) unclassified cases. The mean seizure onset age was 4.77±3.77 years and this was significantly earlier in the symptomatic group ($p<0.001$). Magnetic resonance imaging abnormalities were identified in 31 patients (48.4%). Autonomic symptoms (54.7%) and nocturnal seizures (37.5%) were common, with nocturnal seizures occurring more frequently in the self-limited group ($p=0.014$). Occipital epileptiform activity was observed in 55 patients (85.9%), including 24 out of the 25 (96.0%) patients in the self-limited group and 29 out of the 36 (80.6%) in the symptomatic/determinable group. The remaining nine patients (14.1%) demonstrated multifocal epileptiform activity with occipital involvement. Extra-occipital spread of occipital epileptiform discharges was observed in 56.3% of the patients, and generalized spike-wave discharges (31.3%) and spike-wave activation in sleep (9 patients) were identified in both groups. Initial antiseizure medication failure and polytherapy requirement were significantly more frequent in the symptomatic/determinable group.

Conclusion: Childhood occipital epilepsies represent a heterogeneous and dynamic electroclinical spectrum rather than sharply separated syndromes. The presence of extra-occipital epileptiform activity, generalized spike-wave discharges, and sleep-activated epileptiform activity supports the concept that childhood occipital epilepsies represent a dynamic electroclinical spectrum. Therefore, long-term neurocognitive and electroencephalographic follow-up remains important, particularly in those children with self-limited occipital epilepsy syndromes as electroclinical evolution may continue despite clinical seizure remission.

Keywords: Childhood occipital epilepsy, self-limited epilepsy with autonomic seizures, electroencephalography, generalized spike-wave discharges, symptomatic occipital epilepsy

Corresponding Author
Ayfer Arduç Akçay, MD, Koç University Faculty of Medicine, Department of Pediatrics, Division of Pediatric Neurology, İstanbul, Türkiye
E-mail: drayfer@gmail.com **ORCID:** orcid.org/0000-0002-2976-7112
Received: 19.05.2026 **Accepted:** 29.07.2026 **Epub:** 31.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Arduç Akçay A. Clinical and electroencephalographic characteristics of childhood occipital epilepsies: comparison of self-limited and symptomatic forms. J Pediatr Res. 2026;13(3):238-46

Introduction

Childhood-onset focal epilepsy syndromes were redefined under the category of "self-limited focal epilepsies" in the International League Against Epilepsy (ILAE) 2022 classification. Occipital epilepsy syndromes within this group are characterized by age-related onset, typical electroclinical features, and generally favorable prognosis. In the current terminology, Panayiotopoulos syndrome is referred to as "self-limited epilepsy with autonomic seizures (SeLEAS)," whereas Gastaut-type occipital epilepsy is referred to as "childhood occipital visual epilepsy (COVE)" (1).

SeLEAS typically begins between 3 and 6 years of age, although onset may range from 1 to 14 years, and it is characterized by seizures with prominent autonomic features. Ictal vomiting, eye deviation, altered consciousness, and frequently nocturnal seizures are among the characteristic clinical features of this syndrome. Electroencephalography (EEG) findings may initially be limited to occipital spikes but can demonstrate a multifocal distribution over time and with advancing age (2). Indeed, interictal electroencephalographic findings in SeLEAS have been reported to evolve with age; occipital spikes may predominate in early childhood, whereas a posterior-dominant multifocal distribution involving occipito-frontopolar and centro-parieto-temporal regions may emerge later in the course of the disease (2). These findings suggest that occipital epilepsies cannot be explained solely by localized occipital pathology and may instead represent an age-dependent multifocal electrophysiological spectrum. Although these syndromes are defined as self-limited, their prognosis is not always entirely benign (3). In addition, several studies have reported atypical or mixed forms and emphasized that the classical Panayiotopoulos–Gastaut distinction may not always be clear-cut in terms of clinical and electroencephalographic features (4). By contrast, symptomatic occipital epilepsies arise secondary to underlying structural pathology and, although their clinical features may overlap with self-limited forms, they may differ in terms of prognosis and treatment response (5). Several studies have demonstrated that the presence of a structural etiology may be a major determinant of seizure control and remission outcomes (6).

Most studies on childhood occipital epilepsies have focused on self-limited syndromes such as SeLEAS and COVE, whereas studies comparing self-limited and symptomatic cases are less common. The aim of the present study was to describe the clinical, electroencephalographic, and neuroimaging characteristics of childhood occipital epilepsies and to identify differences between self-limited and symptomatic cases.

Materials and Methods

This retrospective single-center cohort study included patients aged 0-18 years with occipital epilepsy spectrum disorders who were followed at the Pediatric Neurology Clinic of Koç University Hospital between October 2015 and January 2026. Ethical approval for this study was obtained from the Koç University Ethics Committee (approval no.: 2026.164.IRB2.088, date: 23.03.2026). Patients with occipital epilepsy spectrum disorders were retrospectively identified from the clinical records, EEG archives, and neuroimaging data. Eligible patients had at least 6 months of clinical follow-up after diagnosis, at least one interictal EEG recording, and sufficient clinical data for electroclinical classification. Occipital epilepsy spectrum disorder was defined as the presence of occipital-predominant interictal epileptiform discharges and/or seizure semiology suggestive of occipital onset, with or without supporting neuroimaging findings. As ictal EEG recordings were not available for all patients, classification relied on a comprehensive evaluation integrating seizure semiology, interictal EEG findings, neurodevelopmental assessment, magnetic resonance imaging (MRI) findings, and longitudinal clinical follow-up, with ictal EEG data incorporated whenever available. Owing to the fact that classical self-limited occipital epilepsy syndromes may not be fully expressed in very young children, patients younger than 3 years were classified according to these integrated electroclinical criteria rather than age alone. Those cases with insufficient clinical, EEG, neuroimaging, or follow-up data for electroclinical classification were excluded.

Data were retrospectively obtained from the hospital electronic medical records, EEG archives, and neuroimaging reports and recorded in a standardized data form. Demographic characteristics, seizure semiology, neurological and developmental status, EEG and MRI findings, treatment characteristics, and follow-up outcomes were evaluated retrospectively. Focal features were defined as focal motor, sensory, autonomic, or behavioral semiological manifestations suggestive of focal seizure onset, irrespective of the final seizure type classification. Autonomic symptoms included vomiting, pallor, flushing, hypersalivation, cardiorespiratory changes, or other autonomic manifestations occurring during seizures. Loss of consciousness referred to impaired awareness or unresponsiveness during the ictal or immediate postictal

period. Seizure recurrence within 24 hours was defined as two or more seizures occurring within a 24-hour period. Status epilepticus was defined according to the current ILAE operational definition.

MRI findings were classified according to structural characteristics. Perinatal/hypoxic-ischemic injury was analyzed as a separate neuroimaging category. Acquired structural lesions included postinfectious, posttraumatic, vascular, neoplastic, or other non-congenital structural abnormalities identified by MRI. For descriptive analysis, structural MRI abnormalities were classified according to their predominant anatomical distribution as occipital-predominant, posterior-predominant, multifocal, or diffuse. Each patient was assigned to the single category which best reflected their predominant distribution of structural abnormalities.

Occipital epileptiform activity was defined as epileptiform discharges involving the occipital regions. Extra-occipital spread was defined as the propagation of occipital epileptiform discharges to additional non-occipital regions during EEG recording. This category was distinct from multifocal epileptiform activity, which referred to the presence of independent epileptiform foci in two or more non-contiguous regions. Generalized spike-wave discharges (GSWD) were defined as bilaterally synchronous spike-wave complexes involving both hemispheres.

EEG background rhythm abnormality was defined as a slowing or disorganization of the age-appropriate posterior dominant rhythm. Initial antiseizure medication failure was defined as persistence of clinically documented seizures despite an adequate dose and duration of the first antiseizure medication, resulting in dose escalation, medication switch, or add-on therapy. EEG normalization was assessed during follow-up EEG recordings.

Spike-wave activation in sleep (SWAS) was defined as spike-wave activity occupying ≥50% of non-rapid eye movement (non-REM) sleep. Patients fulfilling this EEG criterion and demonstrating developmental, cognitive, language, behavioral, or motor stagnation/regression were considered to be within the ESES/EE-SWAS spectrum. Sleep EEG recordings were available for all patients, and adequate non-REM sleep segments were obtained for SWAS assessment. Spike-wave activation was assessed by a visual estimation of the spike-wave activity during non-REM sleep.

The patients were divided into three groups according to their clinical, EEG, neurodevelopmental, and neuroimaging features. Group 1 consisted of patients compatible with the self-limited occipital epilepsy spectrum, with normal neurodevelopmental evaluation and structural MRI findings together with occipital epileptiform activity. Group 2 included symptomatic cases with structural brain lesions, neurodevelopmental impairment, or identifiable genetic/metabolic/systemic etiologies. Group 3 consisted of those patients who could not be definitively classified into Group 1 or Group 2 based on the available data. As structural MRI abnormalities, developmental delay, and abnormal neurological examination findings were incorporated into the group classification framework, these variables were not included in the between-group statistical comparisons. Within Group 1, cases compatible with SeLEAS, COVE, and POLE were additionally identified; however, these subtypes were not included in separate comparative analyses due to clinical-electroencephalographic overlap and their limited subgroup sizes.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 29.0 (IBM Corp., Armonk, NY, USA). Continuous variables were assessed for normality using visual methods and the Shapiro-Wilk test. Normally distributed variables were compared using the Independent Samples t-test, whereas non-normally distributed variables were analyzed using the Mann-Whitney U test. Categorical variables were compared using the Pearson chi-square test or Fisher's exact test, as appropriate. Results are presented as mean ± standard deviation, median (range), or number (percentage), as appropriate. Analyses were performed using the available-case data, and missing values were not imputed. A two-sided p value <0.05 was considered statistically significant. Given the retrospective design and multiple unadjusted comparisons, p values should be interpreted with caution.

Results

A total of 64 patients were included in this study, comprising 33 males and 31 females. The demographic, clinical, radiological, electroencephalographic, and treatment characteristics of the cohort are summarized in Table I. The mean age was 10.76±3.90 years, and the mean age at diagnosis was 5.19±3.71 years. The mean follow-up duration was 30.1±25.3 months (range, 6-108 months). The mean seizure onset age was 4.77±3.77 years (n=62), with a median age of 5 years; seizure onset age data were unavailable for two patients. Of the patients, 39.1% were classified as Group 1 (compatible with the self-limited occipital epilepsy spectrum), 56.3% as Group 2 (symptomatic/determinable etiology), and 4.7% as Group 3 (unclassified/undetermined group).

Table I. Clinical, radiological, electroencephalographic, and treatment characteristics of the study cohort

Category and parameter	Number of patients (n)	Percentage (%)/Mean
Demographics and classification		
Male	33	51.6%
Female Mean age at study (years)	31 64	48.4% 10.76±3.90
Mean age at diagnosis (years)	64	5.19±3.71
Mean seizure onset age (years)	64	4.77±3.77
Median follow-up duration (months)	64	30.1±25.3
Group 1 (self-limited spectrum)	25	39.1%
Group 2 (symptomatic/determinable etiology)	36	56.3%
Group 3 (unclassified)	3	4.7%
Clinical and semiological features		
Focal semiological features	36	56.3%
Autonomic symptoms	35	54.7%
Automatisms	37	57.8%
Loss of consciousness	57	89.1%
Ictal vomiting	28	43.8%
Seizure recurrence (within 24 hours)	12	18.8%
Nocturnal seizures	24	37.5%
History of status epilepticus	25	39.1%
Visual symptoms	3	4.7%
Abnormal neurological examination	32	50.0%
Developmental delay	29	45.3%
Brain MRI findings		
Normal MRI	33	51.6%
Perinatal/hypoxic-ischemic injury	15	23.4%
Cortical atrophy	6	9.4%
Cerebral infarction	2	3.1%
Cortical dysplasia	2	3.1%
Multiple cavernomas	1	1.6%
Posterior reversible encephalopathy (PRES)	1	1.6%
Tuberous sclerosis complex	1	1.6%
Hydrocephalus	1	1.6%
Sleep EEG findings		
Occipital epileptiform activity	55	85.9%
Bilateral occipital involvement	26	40.6%
Extra-occipital spread	36	56.3%
Multifocal epileptiform activity	9	14.1%
Generalized spike-wave discharges	20	31.3%
EEG normalization (during follow-up)	24	37.5%
Spike-wave activation in sleep	9	14.1%

Table I. Continued

Category and parameter	Number of patients (n)	Percentage (%)/Mean
Treatment and outcomes		
Monotherapy	39	60.9%
Polytherapy	24	37.5%
Initial medication failure	35	54.7%
Most common medications		
Levetiracetam	46	71.9%
Valproate	20	31.3%
Clobazam	19	29.7%
Carbamazepine	9	14.1%
Oxcarbazepine	6	9.4%
Steroid	5	7.8%

Percentages may exceed 100% because individual patients could be included in more than one category
MRI: Magnetic resonance imaging, EEG: Electroencephalography

Three patients (4.7%) were classified as Group 3 because their electroclinical features did not allow for confident assignments to either the self-limited or symptomatic/determinable categories. None of these 3 patients had structural MRI abnormalities. Two of these patients had developmental delay and abnormal neurological examination findings, while one of them had normal neurodevelopment and neurological examination. All three showed occipital epileptiform activity with extra-occipital spread, and one of them also exhibited GSWD.

Focal semiological features were present in 56.3% of patients, autonomic symptoms in 54.7%, and automatisms in 57.8%. Loss of consciousness occurred in 89.1%, vomiting in 43.8%, and seizure recurrence within 24 hours in 18.8%. Nocturnal seizures were observed in 37.5%, while a history of status epilepticus was present in 39.1% overall, including 36% (9/25) of Group 1 and 39% (14/36) of Group 2 patients. Of these, two patients in Group 2 had non-convulsive status epilepticus, whereas all of the remaining cases were convulsive. Visual symptoms were rare and identified in only three patients, whose clinical features were compatible with the COVE spectrum. Neurological examination abnormalities were detected in 50% of cases and developmental delay in 45.3%. Family history revealed epilepsy in nine patients and febrile seizures in four. Four patients had a personal history of febrile seizures, while fever-triggered seizures were reported in 11 cases. Consanguinity was present in four patients.

Brain MRI was normal in 33 of 64 patients (51.6%). MRI abnormalities were identified in 31 patients (48.4%), most commonly perinatal/hypoxic-ischemic injury (15 patients, 23.4%) and cortical atrophy (6, 9.4%), followed by cortical dysplasia (2, 3.1%), cerebral infarction (2, 3.1%), posterior reversible encephalopathy syndrome (2, 3.1%), hydrocephalus (1, 1.6%), multiple cavernomas (1, 1.6%), tuberous sclerosis complex (1, 1.6%), and delayed myelination (1, 1.6%). Regarding lesion distribution, occipital-predominant lesions were present in 4 patients, posterior-predominant lesions in 10, multifocal lesions in 11, and diffuse cerebral involvement in 6.

EEG evaluation demonstrated occipital epileptiform activity in 55 patients (85.9%), including bilateral occipital involvement in 26 cases. The remaining nine patients (14.1%) exhibited multifocal epileptiform activity with occipital involvement. Extra-occipital spread was observed in 56.3% of the patients and was similarly frequent in Group 1 and Group 2 (40% vs. 63.9%). GSWD were detected in 20 patients (31.3%), including 28% (7/25) of Group 1 cases. EEG normalization was documented in 24 patients (37.5%), with a mean time to normalization of 24.8±19.9 months (range, 6–66 months). Electrographic seizures were recorded in four patients, and SWAS was identified in nine (5 in Group 1, 3 in Group 2, and 1 in Group 3). Photosensitivity was observed in two patients, both considered compatible with the POLE spectrum. Both photosensitive patients also demonstrated GSWD.

Additional systemic disorders were present in 19 patients (29.6%). Among these, four (6.3%) had organ transplantation or severe chronic systemic disease, 10 (15.6%) had metabolic-genetic/syndromic disorders, and five (7.8%) had other systemic comorbidities.

Overall, 60.9% of the patients were managed with monotherapy, whereas 37.5% required polytherapy; one patient remained untreated. Levetiracetam was the most commonly used antiseizure medication (71.9%), followed by valproate (31.3%) and clobazam (29.7%). Oxcarbazepine and carbamazepine were used in 9.4% and 14.1% of the patients, respectively. Steroid treatment was administered in five patients; however, steroids were prescribed for associated epileptic encephalopathies rather than occipital epilepsy itself, including West syndrome (n=3) and ESES/EE-SWAS (n=2). Those patients with a previous history of West syndrome were included because they subsequently developed electroclinical features compatible with the occipital epilepsy spectrum during follow-up and fulfilled the predefined inclusion criteria.

Initial antiseizure medication failure, requiring treatment modification or additional therapy, was observed in 35 patients (54.7%).

The comparative clinical, electroencephalographic, and treatment-related characteristics of Group 1 and Group 2 patients are presented in Table II. Comparative analyses were performed between Group 1 (n=25) and Group 2 (n=36), excluding Group 3 cases. Those patients in Group 2 had significantly earlier seizure onset (3.34 vs. 6.63 years, p<0.001) and earlier diagnosis ages (3.81 vs. 6.96 years, p<0.001) compared with Group 1. No significant difference was observed between the groups regarding status epilepticus. Autonomic symptoms were present in 68% of Group 1 and 50% of Group 2 patients; however, this difference was not statistically significant (p=0.165). Similarly, vomiting was observed in 52% of Group 1 and 41.7% of Group 2 patients, without any significant difference (p=0.428). Nocturnal seizures were significantly more frequent in Group 1 than in Group 2 (56% vs. 22.2%, p=0.014).

Neither occipital epileptiform activity nor multifocal epileptiform activity differed significantly between Group 1 and Group 2. By contrast, EEG background rhythm abnormalities were observed exclusively in Group 2 patients (p<0.001). Although EEG normalization was more frequent in Group 1 than in Group 2 (48% vs. 33.3%), the difference was not statistically significant (p=0.294). Similarly, no significant difference was found between Group 1 and Group 2 regarding extra-occipital EEG spread (p=0.075). In addition, EEG normalization was not associated with multifocal epileptiform activity (p=0.464), MRI abnormalities (p=0.561), or developmental delay (p=0.080).

Among those patients with normal development, 66.7% were managed with monotherapy, whereas 66.7% of those with developmental delay required polytherapy (p=0.010).

Initial antiseizure medication failure was significantly more frequent in Group 2 than in Group 1 (69.4% vs. 28.0%, p=0.001), and the need for polytherapy was also higher in this group (47.2% vs. 20.0%, p=0.038). No association was found between initial medication failure and multifocal epileptiform activity (p=0.166), MRI abnormalities (p=0.075), EEG background rhythm abnormalities (p=0.069), or extra-occipital EEG spread (p=0.155). However, developmental delay was significantly associated with a higher rate of initial medication failure (p<0.001).

Table II. Comparative analysis of self-limited (Group 1) and symptomatic (Group 2) occipital epilepsy cases

Parameters	**Group 1 (n=25)**	**Group 2 (n=36)**	**p value**
Mean seizure onset age (years)	6.63±3.52	3.34±3.41	<0.001
Mean age at diagnosis (years)	6.96±3.41	3.81±3.44	<0.001
Nocturnal seizures (%)	56.0% (14/25)	22.2% (8/36)	0.014
Autonomic symptoms (%)	68.0% (17/25)	50.0% (18/36)	0.165
Ictal vomiting (%)	52.0% (13/25)	41.7% (15/*36)*	0.428
Extra-occipital EEG spread (%)	40.0% (10/25)	63.9% (23/36)	0.075
Generalized spike-wave discharges	28.0% (7/25)	36% (13/36)	0.594
EEG background rhythm abnormality (%)	0.0% (0/25)	36.1% (17/36)	<0.001
EEG normalization	48.0% (12/25)	33.3% (12/36)	0.294
Status epilepticus history (%)	36.0% (9/25)	39.0% (14/36)	0.812
Initial medication failure (%)	28.0% (7/25)	69.4% (25/36)	0.001
Polytherapy requirement (%)	20.0% (5/25)	47.2% (17/36)	0.038

EEG: Electroencephalography

Although systemic disorders were more frequent in Group 2, the difference between the groups was not statistically significant (p=0.089). Both Group 1 patients with systemic comorbidity had type 1 diabetes mellitus. By contrast, Group 2 patients exhibited more heterogeneous and clinically severe comorbidities, including organ transplantation, congenital heart disease, and genetic syndromes.

Discussion

In this study, the clinical, electroencephalographic, and neuroimaging features of childhood occipital epilepsy spectrum disorders were evaluated, and differences between self-limited spectrum cases and symptomatic/determinable etiologies were identified. Symptomatic/determinable cases showed earlier seizure onset and greater treatment burden, including higher rates of initial antiseizure medication failure and polytherapy. Additionally, multifocality and extra-occipital EEG spread were not limited to symptomatic cases, supporting the concept that childhood occipital epilepsies represent a heterogeneous and dynamic electroclinical spectrum.

Regarding seizure semiology, autonomic symptoms and ictal vomiting were among the commonly observed features in our cohort and were largely compatible with the SeLEAS spectrum in the self-limited group. Ictal vomiting and other autonomic manifestations have been reported as characteristic clinical features of SeLEAS (1,7). Aksoy et al. (4) similarly identified vomiting (80%) and eye deviation (87%) as the most frequent findings in Panayiotopoulos syndrome. Although autonomic symptoms and vomiting were more frequent in Group 1 in our cohort, the differences between groups did not reach statistical significance. This may be related to clinical overlap, the presence of autonomic manifestations in symptomatic cases, and/or the limited sample size.

Visual symptoms were rare, suggesting an autonomic-dominant spectrum rather than classic COVE; such manifestations may be underreported in young children due to their brevity and description difficulties (8). Additionally, the higher frequency of nocturnal seizures in Group 1 is consistent with SeLEAS patterns. These results, alongside noted overlaps between Panayiotopoulos and Gastaut syndromes, reinforce the concept of childhood occipital epilepsies as a dynamic, overlapping electroclinical spectrum (4).

The significantly earlier seizure onset in the symptomatic group (3.34 vs. 6.63 years) highlights the association between structural etiologies and a more severe clinical course. The mean onset age of 4.77 years in our cohort is consistent with previous reports supporting the age-dependent nature of occipital epilepsies (9). Early onset has been recognized as one of the strongest predictors of poor prognosis, treatment resistance, and structural MRI abnormalities (6,7,10).

The frequent presence of perinatal/hypoxic injury in our cohort, together with the selective vulnerability of the occipital lobe to neonatal hypoglycemia (11), suggests that structural posterior abnormalities may contribute to epileptogenesis in a subset of patients. The higher need for polytherapy and the presence of EEG background rhythm abnormalities in the symptomatic group further emphasize that etiological classification is a major determinant of treatment strategy and prognosis (1,6).

In recent years, the term "*self-limited occipital epilepsy*" has been preferred over "*benign occipital epilepsy*" due to risks such as cognitive impairment and the development of electrical status epilepticus during sleep (ESES) (12). The presence of multifocality, extra-occipital spread, and SWAS in our cohort supports this approach. The current literature suggests that SWAS/ESES is not limited to SeLECTS and may occur at similar rates in SeLEAS. İriş et al. reported a SWAS rate of 9.23% in a SeLEAS group, comparable to that observed in SeLECTS (3). Therefore, those patients presenting with autonomic seizures may require follow-up with sleep EEG recordings.

Prolonged seizures and focal status epilepticus have been reported to occur frequently in SeLEAS, affecting approximately one-third of patients (12). In Panayiotopoulos syndrome, nearly half of seizures may present as autonomic status epilepticus (13), while convulsive status at onset has not been found to be directly associated with prognosis (14). The 36% rate of status epilepticus in Group 1 in our cohort supports the inclusion of these presentations within the self-limited occipital epilepsy spectrum. Although this relatively high rate may partly reflect referral bias related to our tertiary center setting, our findings indicate that self-limited occipital epilepsies do not always present with a mild clinical course and that the term "*self-limited*" does not exclude dramatic seizure presentations during the acute phase.

In the SeLEAS spectrum, interictal electroencephalographic findings are not limited to the occipital region, and shifting multifocal discharges are frequently observed (12). The literature further emphasizes that multifocality and extra-occipital spread are among the

main features supporting the use of the term "*self-limited*" rather than "*benign*". The 31.3% rate of GSWD in our cohort is consistent with previously reported rates of 17-35% (15-17).

The relatively high frequency of GSWD in our cohort may reflect broader network involvement in a subset of patients, although the retrospective design and limited ictal EEG data did not allow firm mechanistic conclusions. In COVE series excluding photosensitive cases according to newer criteria, markedly lower GSWD rates (5%) have been reported, suggesting a relationship between photosensitivity and generalized epileptiform activity (18). Verrotti et al. (14) also emphasized that occipital epileptiform activity may evolve into multifocal patterns during the course of the disease. The extra-occipital and multifocal findings in our cohort are consistent with the concept that childhood occipital epilepsies represent a dynamic electroclinical spectrum. In addition, as SWAS recurrence has been reported even years after EEG normalization (3), these findings suggest that selected patients, particularly those with persistent EEG abnormalities, developmental concerns, sleep-activated epileptiform activity, or treatment difficulties, may benefit from longitudinal clinical and EEG follow-up.

The ability to manage self-limited occipital epilepsy cases with monotherapy supports the pharmaco-sensitive nature of this group (19), whereas symptomatic cases more often require polytherapy. Although carbamazepine and valproate have traditionally been preferred in the literature (1,18,19), the predominance of levetiracetam usage (71.9%) in our cohort likely reflects the current tendency in pediatric practice toward well-tolerated broad-spectrum antiseizure medications.

Study Limitations

This study had several limitations. Its retrospective design may have introduced selection bias and limited the completeness of the clinical data, including seizure semiology. Neurocognitive outcomes were not systematically assessed using a standardized longitudinal protocol, and ictal EEG recordings were available for only a small proportion of the patients. The absence of multivariable analyses limited our ability to determine whether the observed associations were independent of underlying etiological or developmental factors. In addition, as this was a tertiary referral center cohort, referral bias may have resulted in an overrepresentation of more complex or treatment-resistant cases.

Conclusion

Childhood occipital epilepsies demonstrated heterogeneous clinical and electroencephalographic characteristics in this retrospective cohort. Self-limited spectrum cases more frequently exhibited autonomic symptoms and nocturnal seizures, whereas symptomatic/determinable cases showed earlier seizure onset and greater treatment burden. The extra-occipital epileptiform activity, GSWD, and sleep-activated epileptiform activity observed in a subset of our patients further support the concept that childhood occipital epilepsies may represent a dynamic electroclinical spectrum rather than sharply separated syndromes. These findings suggest that selected patients, particularly those with sleep-activated epileptiform activity, persistent EEG abnormalities, developmental concerns, or treatment difficulties, may benefit from longitudinal clinical and EEG follow-up. These results should be interpreted in light of the retrospective design and the limited subgroup sizes in this study.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the Koç University Ethics Committee (approval no.: 2026.164.IRB2.088, date: 23.03.2026).

Informed Consent: Informed consent was obtained.

Footnotes

Acknowledgements

The author would like to thank Professor Serap Uysal, MD, for her helpful comments and discussions regarding the evaluation of selected patients included in this study. The author also sincerely thanks all the patients and their families for their valuable cooperation and for consenting to the use of their clinical data for research purposes.

Financial Disclosure: The author received no funding from any agency or individual.

References

1. Specchio N, Wirrell EC, Scheffer IE, et al. International League Against Epilepsy classification and definition of epilepsy syndromes with onset in childhood: position paper by the ILAE task force on nosology and definitions. Epilepsia. 2022; 63:1398-442.
2. Oguni H. A reappraisal of interictal EEG characteristics in self-limited epilepsy with autonomic seizures, formerly known as Panayiotopoulos syndrome or early-onset benign occipital epilepsy. Eur J Paediatr Neurol. 2023; 47:18-24.

3. İriş M, Yaşgüçlükal MA, Yalçınkaya C, Demirbilek V. Evolution into spike-and-wave activation in sleep in patients with self-limited focal epilepsies. Seizure. 2026; 135:34-8.

4. Aksoy A, Haliloğlu G, Yalnızoğlu D, Turanlı G. Childhood epilepsy with occipital paroxysm: classification, atypical evolution and long-term prognosis in 35 patients. Turk J Pediatr. 2015; 57:439-52.

5. Çavuşoğlu, D. Childhood occipital epilepsy: clinical evaluation of 19 patients. Kocatepe Tıp Dergisi. 2021; 22:168-72.

6. Özkul D, Tekin HG. Childhood epilepsies with occipital discharges: evaluation of 84 patients. J Pediatr Res. 2025; 12:96-101.

7. Tata G, Guveli BT, Dortcan N, et al. Panayiotopoulos syndrome and symptomatic occipital lobe epilepsy of childhood: a clinical and EEG study. Epileptic Disord. 2014; 16:197-202.

8. Taylor I, Scheffer IE, Berkovic SF. Occipital epilepsies: identification of specific and newly recognized syndromes. Brain. 2003; 126:753-69.

9. Gauci MC, Gauci R, Ruggieri M, Polizzi A, Praticò AD. Clinical, genetic, EEG, neuroimaging insights and conservative treatment in pediatric focal epilepsy: a retrospective observational study. J Clin Med. 2025; 14:2234.

10. Schrader D, Shukla R, Gatrill R, Farrell K, Connolly M. Epilepsy with occipital features in children: factors predicting seizure outcome and neuroimaging abnormalities. Eur J Paediatr Neurol. 2011; 15:15-20.

11. Hosseinzadeh S, Vagharmousavi D, Rajabzadeh R, Bayani G, Babaei M. Determining potential risk factors for epilepsy in children with neonatal hypoglycemia. Iran J Child Neurol. 2025; 19:53-62.

12. Barua D, Anilkumar AC. Self-limited focal epilepsies in childhood. Pract Neurol. 2018; 18:64-8.

13. Adcock JE, Panayiotopoulos CP. Occipital lobe seizures and epilepsies. J Clin Neurophysiol. 2012; 29:397-407.

14. Verrotti A, Sebastiani M, Giordano L, et al. Panayiotopoulos syndrome with convulsive status epilepticus at the onset: a long-term study. Seizure. 2014; 23:728-31.

15. Wakamoto H, Nagao H, Fukuda M, et al. Idiopathic childhood occipital epilepsy of Gastaut: report of 12 patients. Pediatr Neurol. 2011; 44:183-6.

16. Caraballo RH, Cersósimo RO, Fejerman N. Childhood occipital epilepsy of Gastaut: a study of 33 patients. Epilepsia. 2008; 49:288-97.

17. Verrotti A, Laino D, Rinaldi VE, et al. Clinical dissection of childhood occipital epilepsy of gastaut and prognostic implication. Eur J Neurol. 2016; 23:241-6.

18. Atacan Yaşgüçlükal M, Güleç B, Emre DH, Elmalı AD, Ertürk Çetin Ö, Demirbilek AV. Clinical and electrophysiological characteristics and prognosis of childhood occipital visual epilepsy in light of current ILAE criteria. Clin EEG Neurosci. 2026; 57:293-8.

19. Guerrini R, Pellacani S. Benign childhood focal epilepsies. Epilepsia. 2012; 53(Suppl 4):9-18.

J Pediatr Res 2026;13(3):247-53
DOI: 10.4274/jpr.galenos.2026.66664

Original Article

Heart Rate Variability in Children Receiving Palliative Care: An Exploratory Cross-Sectional Study

Meryem Beyazal, Ganime Ayar

University of Health Sciences Türkiye, Ankara Bilkent City Hospital, Clinic of Pediatrics, Ankara, Türkiye

ABSTRACT

Aim: Heart rate (HR) variability (HRV) is simply a measure of the variation in time between each heartbeat. This variation is controlled by the autonomic nervous system. The aim of this study was to evaluate HRV parameters in children who were hospitalized in palliative care.

Materials and Methods: Twenty-eight palliative patients and 30 age- and gender-matched healthy control participants were included in this study. 24-hour Holter monitoring results were compared.

Results: The minimum and mean HRs were significantly higher, while the maximum HR was significantly lower in the palliative care patients. Additionally, PR duration was shorter in the palliative care patients than in the control participants. Moreover, all time and frequency domain HRV values were lower in the palliative care patients compared to the control participants. Lastly, the low frequency/high frequency ratio was almost significantly higher in the palliative care patients than in the controls. In the analysis of only palliative cases, the mean root mean square of successive RR interval differences and the percentage of successive RR intervals which differed by more than 50 ms (pNN50) were significantly higher in those patients with tracheostomy, while the same parameters were significantly lower in those with gastrostomy.

Conclusion: HRV parameters may provide complementary information regarding autonomic regulation in pediatric palliative care settings.

Keywords: Heart rate variability, palliative care, children

Introduction

Heart rate (HR) variability (HRV) is a marker which provides information regarding the health of the autonomic nervous system. HRV has also been accepted as a parameter representing the complex interaction between the brain and the cardiovascular system (1). HRV consists of time and frequency domain parameters. Among these parameters, some of them, such as standard deviation of normal-to-normal intervals (SDNN) and low frequency (LF), indicate both sympathetic and parasympathetic health, while root mean square of successive RR interval differences (rMSSD), pNN50 and high frequency (HF) are more influenced by the parasympathetic system (2).

The autonomic nervous system can offer insights into the biological aspects of psychological states and coping in the context of palliative medicine, since it mediates the body's response to stressors (3). Studies on how HRV is affected in pediatric patients receiving palliative care are limited. However, the significant relationship between SDNN and acute well-being and also higher HRV values in

Corresponding Author
Meryem Beyazal, MD, University of Health Sciences Türkiye, Ankara Bilkent City Hospital, Clinic of Pediatrics, Ankara, Türkiye
E-mail: meryembeyazal@gmail.com **ORCID:** orcid.org/0000-0002-5626-9055
Received: 11.02.2026 **Accepted:** 24.07.2026 **Epub:** 01.09.2026 **Publication Date:** 18.09.2026

Cite this article as: Beyazal M, Ayar G. Heart rate variability in children receiving palliative care: an exploratory cross-sectional study. J Pediatr Res. 2026;13(3):247-53

adult patients discharged from palliative care have been reported (4,5). In addition, HRV has been used in a limited number of palliative care studies, primarily as a potential predictor of mortality (6,7).

In this study, it was aimed to determine whether the HRV parameters of children receiving palliative care differed from those of age- and gender-matched healthy children and also to evaluate the clinical importance of HRV in palliative care.

Materials and Methods

Study Design

This prospective cross-sectional study was conducted in children between eight and 18 years old. The inclusion criteria were as follows: a stay in a pediatric palliative care clinic for at least two weeks, full recovery from acute inflammation and infection, 24-hour Holter monitoring due to suspected rhythm problems on electrocardiography and voluntary participation in this study. Those patients with clinically significant arrhythmias, such as sustained atrial or ventricular arrhythmias, were excluded from this study. In addition, those patients with progressive neurological disease, fever, inflammatory disease, head trauma, those with a Glasgow Coma score <13 and those using medication such as sedatives, opioids, antiepileptics, and beta-blockers were also excluded from this study. Age- and gender-matched healthy control groups were selected from asymptomatic patients who visited the pediatric cardiology outpatient clinic for examination before participating in sport, and had applied for 24-hour Holter monitoring due to their family history.

Written informed consent was obtained from the parents of all of the participants, and this study was approved by the University of Health Sciences Türkiye, Ankara Bilkent City Hospital Clinical Research Ethics Committee No. 2 (approval no.: E2-23-5880, date: 27.12.2023).

Data Collection

The patients' data were obtained from the electronic records. Age, gender, vital signs such as HR, respiratory rate, systolic and diastolic blood pressure (BP), primary and secondary pathology, the presence of tracheostomy, gastrostomy or medical device, Karnofsky/Lansky performance score, and hospital stay duration were noted. The Karnofsky and Burchenal (8) and Lansky Performance Scores (9) were obtained during their stay in the palliative care unit simultaneously with the 24-hour Holter monitoring.

24-Hour Holter Monitoring

The DM Software CardioScan 79A was used. All participants, including the control group, underwent Holter monitoring under standardized daily living conditions. They were instructed to maintain their usual routines while avoiding strenuous physical activity. Posture and activity-related variations were minimized by ensuring comparable daytime activity patterns and nighttime rest periods between the groups. All participants were instructed to maintain their usual daily activities during the monitoring period. However, they were specifically asked to avoid strenuous physical exercise, maintain similar daytime activity patterns, and follow their habitual sleep-wake schedule in order to minimize circadian and activity-related variability. In addition, parents/caregivers were instructed to ensure the consistency of daily routines as much as possible during the recording period. During HRV processing, ectopic beats and artifacts were identified and managed according to standard procedures. Specifically, segments containing frequent ectopic beats or artifacts were either automatically filtered by the analysis software or manually reviewed and excluded when necessary. Only NN intervals were used for HRV calculations in order to maintain data accuracy and consistency. 24-hour Holter parameters including rhythm, PR, QTC duration, minimum, maximum and mean HR, and HRV parameters such as SDNN, SDANN (the standard deviation of all five-min interval means), SDNNi (mean of the standard deviation of all of the NN intervals for each five-minute segment), rMSSD, pNN50, triangular index, total power, very-LF, LF and HF were calculated and recorded.

Statistical Analysis

All data were analyzed using the Statistical Package for Social Sciences for Windows 25.0 program (IBM Corp., Armonk, N.Y., USA). After descriptive statistics and normality analysis, normally distributed continuous variables are reported as mean ± standard deviation, non-normally distributed continuous variables are given as median with range, and categorical variables are shown as count with the percentage of the total. Chi-square or Fisher's exact test was used to compare categorical or ordinal variables, and Student's t-test or the Mann-Whitney U test was used to compare continuous variables.

Results

Twenty-eight palliative patients and 30 age- and gender-matched healthy control participants were included in

Table I. Clinical characteristics of the palliative care patients	
	Data*
Age (year)	9.05±5.99
Gender Male Female	 16 (57.1%) 12 (42.9%)
Height (cm)	117.71±40.37
Weight (kg)	28.37±21.93
Respiratory rate (per min)	25.43±7.66
BP systolic (mmHg)	98.79±19.55
BP diastolic (mmHg)	59.29±14.58
Primary pathology Trauma Syndromic disease Neurometabolic disease Operated CHD Malignancy Lupus PM, BPD	 10 (35.7%) 2 (7.1%) 4 (14.2%) 5 (17.8%) 5 (17.8%) 1 (3.6%) 1 (3.6%)
Hospitalization duration	30 (15-90)
Lansky-Karnofsky score	33.92±18.72
Tracheostomy	22 (78.6%)
Gastrostomy	4 (14.3%)
Medical device MV BIPAP	 11 (39.3%) 17 (60.7%)

*Data are presented as n (%) for categorical variables and median (interquartile range) or mean± standard deviation for continuous variables
CHD: Congenital heart disease, PM: Prematurity, BPD: Broncho pulmonary dysplasia, MV: Mechanical ventilation, BIPAP: Bi-level positive airway pressure

this study. While 46.7% (n=14) of palliative patients were female, this rate was 57.1% (n=16) in the healthy control participants. The mean ages for the palliative patients and healthy control participants were 9.04±6.02 and 9.10±4.03, respectively. Table I shows the clinical characteristics of those patients who were hospitalized in the palliative care unit.

When comparing the palliative care patients to the healthy control participants, the minimum and mean HRs were significantly higher, while the maximum HR was significantly lower in the palliative care patients. Additionally, the PR duration was shorter in the palliative care patients than in the control participants. Moreover, all time and frequency domain HRV values were lower in the palliative care patients compared to the control participants. Lastly, the LF/HF ratio was almost significantly higher in the palliative care patients than in the control group. Detailed 24-hour Holter monitoring results are given in Table II.

When it comes to the analysis of only the palliative cases, median rMSSD and pNN50 were significantly higher in those patients with tracheostomy with a value of 32 (10-175) for rMSSD and of 7 (0-33) for pNN50, as opposed to 20 (10-29) and 1.5 (0-7) for those patients without this condition, as shown in Figure 1 (p=0.044 and 0.042, respectively). In addition, the medians of the same parameters were significantly lower in those with gastrostomy, with a value of 14 (10-3) versus 29 (10-175) for rMSSD and 1 (0-6) versus 7 (0-33), as shown in Figure 2 (p=0.037 and 0.043, respectively). Moreover, interesting differences were observed when those patients on mechanical ventilation were compared to those supported with Bi-level positive airway pressure (BIPAP). For instance, the former group's median rMSSD [48 (19-175)] was significantly higher than the latter group's [26 (10-45)], while the median SDNNi and pNN50 were almost significantly higher [52 (13-132) versus 29 (10-52) for SDNNi and 22 (0-33) versus 6 (0-19) for pNN50], as shown in Figure 3 (p=0.005, 0.057, 0.058, respectively). However, no correlations were found between the duration of hospital stay and the Karnofsky and Lansky scores and the HRV parameters.

Table II. Comparison of 24-holter monitoring				
Data*	**Total (n=58)**	**Patients (n=28)**	**Control (n=30)**	**p value**
HR				
Minimum	56 (39-100)	64.50 (40-100)	51.5 (39-85)	**0.001****
Maximum	161.3±22.2	156.3±21.9	165.9±21.8	**0.049****
Mean	98.5±18.8	107.2±21.6	90.4±10.9	**0.001****
ECG				
PR duration (ms)	140 (100-200)	120 (100-160)	140 (120-180)	**0.049****
QTC (msn)	395.6±23.7	393.3±23.1	394.9±24.7	0.651

Table II. Continued

Data*	Total (n=58)	Patients (n=28)	Control (n=30)	p value
Time domain parameters				
SDNN (msec)	122.5 (36-260)	73.5 (36-210)	155.0 (67-260)	**0.0001****
SDANN (msec)	102.5 (28-244)	61.5 (28-193)	123.5 (58-244)	**0.0001****
SDNNi (msec)	53.5 (10-215)	33.5 (10-132)	77.0 (33-215)	**0.0001****
rMSSD (msec)	44 (10-191)	27.5 (10-175)	73.5 (25-191)	**0.0001****
pNN50 (msec)	19 (1-71)	6.5 (0-33)	29.5 (4-71)	**0.0001****
Triangular index	27.7±13.3	21.0±13.4	33.9±9.9	**0.0001****
Frequency domain parameters				
TP (msec^2)	2,605.9 (43.4-7,936)	896.9 (43.4-4,463)	3,138.0 (893-7,936)	**0.0001****
VLF (msec^2)	1,234.5 (18.4-5,418)	571.0 (18.4-3204)	2,031.5 (495-5,418)	**0.0001****
LF (msec^2)	407.2 (10.1-1899)	179.7 (10.1-731)	608.5 (96-1,899)	**0.0001****
HF (msec^2)	300.1 (11.8-1,223)	68.1 (11.8-905)	558.7 (37.6-1,223)	**0.0001****
LF/HF ratio	2.06±1.31	2.48±1.62	1.66±0.78	0.058

*Data are presented as n (%) for categorical variables and median (minimum-maximum) or mean ± standard deviation for continuous variables
**Significant difference (p<0.05)
HRV: Heart rate variability, SDNN: Standard deviation of NN intervals, SDANN: The standard deviation of all 5-min interval means, SDNNi: Mean of the standard deviation of all the NN intervals for each 5-minute segment of a 24 hours HRV recording, rMSSD: Root mean square of successive RR interval differences, pNN50: Percentage of successive RR intervals that differ by more than 50 ms, TP: Total power, HF: High frequency, LF: Low frequency, VLF: Very-LF

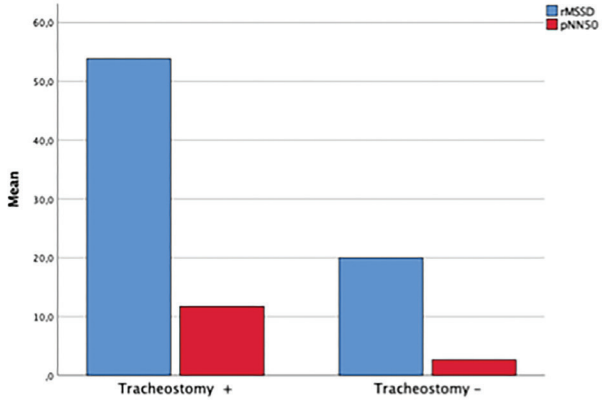

Figure 1. Comparison of rMSSD and pNN50 between patients with and without tracheostomy (p=0.044 and 0.042, respectively)
rMSSD: Root mean square of successive RR interval differences, pNN50: Percentage of successive RR intervals that differ by more than 50 ms

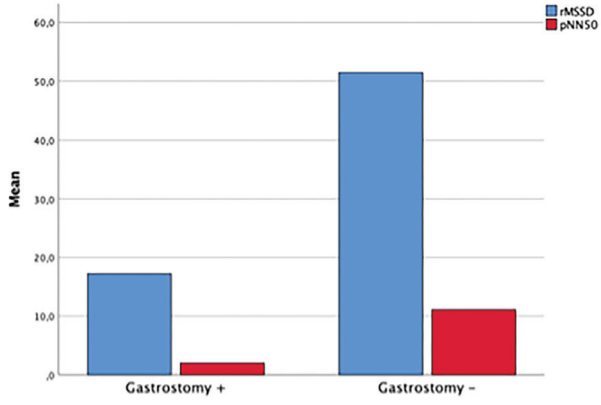

Figure 2. Comparison of rMSSD and pNN50 between patients with and without gastrostomy (p=0.037 and 0.043, respectively)
rMSSD: Root mean square of successive RR interval differences, pNN50: Percentage of successive RR intervals that differ by more than 50 ms

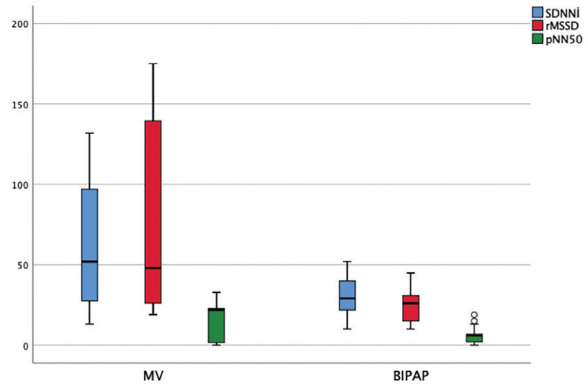

Figure 3. Comparison of rMSSD, pNN50 and SDNNi between mechanically ventilated and BIPAP patients (p=0.005, 0.057, 0.058, respectively)
rMSSD: Root mean square of successive RR interval differences, pNN50: Percentage of successive RR intervals that differ by more than 50 ms, SDNNi: Mean of the standard deviation of all the NN intervals for each 5-minute segment of a 24 hours HRV recording, BIPAP: Bi-level positive airway pressure, MV: Mechanical ventilation

Discussion

According to the current study, in those patients who were hospitalized in palliative care for at least 15 days, overall HRV parameters and maximum HR were found to be lower, minimum and mean HRs were higher and PR intervals were shorter than for those of the control participants. Analysis of only palliative care patients showed that while parasympathetic system-related parameters were higher in those with tracheostomy, the same parameters were lower in those with gastrostomy. In addition, similarly, parasympathetic system-related parameters were higher in those on mechanical ventilation compared to BIPAP.

The lower overall HRV parameters and the decreased difference between maximum and minimum HRs observed in this cohort may be consistent with altered autonomic regulation. However, given the heterogeneity of the underlying conditions and the cross-sectional design, it is not possible to determine whether these findings are attributable to the palliative care status itself, the primary diseases, or other clinical factors. Various factors, including trauma, cardiac surgery, and cancer, can impair HRV (10-12). Since most of the patients had conditions such as trauma, surgery, or cancer in their etiology, the decrease in HRV may be related to the primary etiology. In addition, hospitalization can cause anxiety and poorer sleep quality in the context of chronic stressor exposure and these conditions result in impaired HR regulation (13,14). On the other hand, a recent study reported that patients' posture does not have an effect on HRV parameters (15). Therefore, impaired HRV parameters in palliative care patients may be attributable to a combination of factors, including their primary illness and emotional state. In order to establish this, an assessment of the patients' emotional state with scales such as anxiety and sleep quality may be considered.

The finding that the mean HR and LF/HF ratios of palliative patients were higher than those of the control group may reflect a relative shift of autonomic balance toward sympathetic predominance. Since LF is affected by both the sympathetic and parasympathetic systems and HF is affected more by the parasympathetic system, an increase in the LF/HF ratio can be used as an indicator of sympathetic activity (16). A relatively more active sympathetic system in palliative care patients could be related to exposure to stressors such as pain and anxiety. However, it should be noted that the interpretation of the LF/HF ratio as a direct marker of sympathetic activity remains controversial in the current HRV literature (17,18) and should therefore be interpreted with caution. In addition, as HR increases, the PR interval tends to shorten (19), which may partly explain the shorter PR intervals observed in this cohort.

Another finding of this study was that the Karnofsky and Lansky scores used as performance scores were not associated with HRV parameters. These scores assess whether the patient is performing an activity, such as walking, playing, or performing basic needs, independently or with support. To the best of our knowledge, there are no studies investigating the relationship between performance scores and HRV. Our results may suggest that the level of functional independence alone does not necessarily correspond to differences in HRV parameters; however, this finding should be interpreted cautiously. Further studies with larger patient populations are needed in order to better understand this relationship.

The fact that parasympathetic system-related HRV parameters were lower in those with gastrostomy may be explained by vagal nerve influence. Vital aspects of autonomic physiology, such as breathing, HR, BP, and gut motility, as well as reflexes such as coughing and swallowing and survival behaviors such as eating, drinking, and reacting to illness, are all regulated by the vagus nerve, a body-brain connection (20). The link between vagus nerve activity and the HF component of HRV has been well-established (21). Patients with gastrostomy demonstrated lower rMSSD and pNN50 values. Although impaired vagal function may represent one possible explanation, the small subgroup size and potential influence of underlying neurological disease preclude definitive conclusions regarding the

mechanisms involved. It has also been reported that vagus nerve stimulating devices can improve oral feeding in patients with gastrostomy (22). However, this association should be interpreted with caution, as the severity of the underlying neurological condition may also have contributed to reduced parasympathetic activity and the need for gastrostomy. Therefore, this association should be considered hypothesis-generating.

The most obvious difference between spontaneous breathing and mechanical ventilation is that intrathoracic volume decreases in spontaneous breathing, while it increases with mechanical ventilator support (23). During mechanical ventilation, the sudden expansion of the lungs activates the afferent arm of a depressor reflex, which produces negative inotropic and chronotropic responses, in addition to arterial vasodilation. The receptors are sensitive to stretch and the afferent pathway runs predominantly in the vagus nerves. Vagal feedback from pulmonary stretch receptors is obligatory for the generation of a neurally mediated respiratory sinus arrhythmia in awake individuals (24,25). Accordingly, the higher rMSSD and pNN50 values observed in those patients with tracheostomy and mechanical ventilation support may represent an exploratory observation which is potentially related to vagally mediated physiological mechanisms. However, given the limited sample size, these findings should be interpreted cautiously and require confirmation in larger studies.

Study Limitations

This study had several limitations. First, it was conducted at a single center with a relatively small sample size, which may limit the generalizability of the findings. Second, the heterogeneity of the underlying diagnoses in the palliative care population may have influenced HRV parameters, making it difficult to isolate the effects of palliative care itself from those of the primary disease. Although patients with conditions known to markedly affect autonomic function, such as head trauma and chronic inflammatory diseases, were excluded, the study cohort still comprised a range of clinical conditions, including neurological disorders, post-surgical states, and malignancies, each of which may affect autonomic regulation through different pathophysiological mechanisms. Therefore, residual confounding related to disease-specific effects cannot be excluded and may have contributed to variability in HRV measurements. In addition, the lack of disease-specific subgroup analyses further limits the ability to draw condition-specific conclusions. Subgroup analyses involving tracheostomy, gastrostomy, and respiratory support modalities were based on relatively small numbers of patients; therefore, these findings should be considered exploratory and hypothesis-generating rather than confirmatory. Future multicenter studies with larger and more homogeneous patient populations are needed in order to validate these observations. Third, emotional and psychological factors such as anxiety, pain perception, and sleep quality, all of which can affect autonomic function and HRV, were not systematically assessed or controlled for. Moreover, the cross-sectional design also precludes conclusions about causality or longitudinal changes in HRV over time. The wide age range of the pediatric participants may have influenced HRV parameters, as HRV is known to vary with age. Furthermore, the absence of follow-up data prevents assessment of whether HRV parameters are associated with clinical outcomes such as symptom burden, functional decline, or survival. Lastly, the selection of control participants who underwent Holter monitoring due to family history may have introduced a potential selection bias, which should be considered when interpreting the results.

Conclusion

In this exploratory study, children receiving palliative care demonstrated lower HRV parameters and higher mean HRs compared with healthy controls. These findings may be consistent with altered autonomic regulation; however, because of the heterogeneous patient population and cross-sectional design, the underlying mechanisms and clinical significance remain uncertain. Differences observed according to gastrostomy status, tracheostomy, and mechanical ventilatory support should be regarded as exploratory observations. Although physiological mechanisms involving vagal activity may contribute to these findings, larger studies are required before definitive interpretations can be made. HRV parameters may provide complementary information regarding autonomic regulation in pediatric palliative care settings; however, further studies are needed before their clinical utility can be established.

Ethics

Ethics Committee Approval: Ethics committee approval was obtained from the University of Health Sciences Türkiye, Ankara Bilkent City Hospital Clinical Research Ethics Committee No. 2 (approval no.: E2-23-5880, date: 27.12.2023).

Informed Consent: Written informed consent was obtained from the parents of all of the participants.

Footnotes

Authorship Contributions

Concept: M.B., Design: M.B., Data Collection or Processing: G.A., Analysis or Interpretation: M.B., Literature Search: G.A., Writing: M.B.

Conflict of Interest: The authors declare no conflicts of interest.

Financial Disclosure: The authors received no financial support for the conduct, authorship, or publication of this study.

References

1. Ernst G. Heart-rate variability—more than heart beats? Front Public Health. 2017; 5:240.
2. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health. 2017; 5:258.
3. Taylor MR, Scott SR, Steineck A, Rosenberg AR. Objectifying the subjective: the use of heart rate variability as a psychosocial symptom biomarker in hospice and palliative care research. J Pain Symptom Manage. 2021; 62:e315-21.
4. Warth M, Kessler J, Bardenheuer HJ. Insomnia and autonomic function predict well-being in patients receiving palliative care. J Palliat Med. 2017; 20:1395-9.
5. Masel EK, Huber P, Schur S, et al. Predicting discharge of palliative care inpatients by measuring their heart rate variability. Ann Palliat Med. 2014; 3:244-9.
6. Kloter E, Barrueto K, Klein SD, Scholkmann F, Wolf U. Heart rate variability as a prognostic factor for cancer survival: a systematic review. Front Physiol. 2018; 9:623.
7. Chiang JK, Kuo TB, Fu CH, Koo M. Predicting 7-day survival using heart rate variability in hospice patients with non-lung cancers. PLoS One. 2013; 8:e69482.
8. Karnofsky DA, Burchenal JH. The clinical evaluation of chemotherapeutic agents in cancer. In: MacLeod CM, ed. Evaluation of chemotherapeutic agents. New York, NY: Columbia University Press; 1949; 191-205.
9. Lansky LL, List MA, Lansky SB, Cohen ME, Sinks LF. Toward the development of a play-performance scale for children (PPSC). Cancer. 1985; 56:1837-40.
10. Florez-Perdomo WA, García-Ballestas E, Moscote-Salazar LR, et al. Heart rate variability as a predictor of mortality in traumatic brain injury: a systematic review and meta-analysis. World Neurosurg. 2021; 148:80-9.
11. Heragu NP, Scott WA. Heart rate variability in healthy children and in those with congenital heart disease both before and after operation. Am J Cardiol. 1999; 83:1654-7.
12. Palma S, Keilani M, Hasenoehrl T, Crevenna R. Impact of supportive therapy modalities on heart rate variability in cancer patients: a systematic review. Disabil Rehabil. 2020; 42:36-43.
13. Cheng YC, Su MI, Liu CW, Huang YC, Huang WL. Heart rate variability in patients with anxiety disorders: a systematic review and meta-analysis. Psychiatry Clin Neurosci. 2022; 76:292-302.
14. da Estrela C, McGrath J, Booij L, Gouin JP. Heart rate variability, sleep quality, and depression in the context of chronic stress. Ann Behav Med. 2021; 55:155-64.
15. Kumar P, Das AK, Halder S. Statistical heart rate variability analysis for healthy persons: influence of gender and body posture. J Electrocardiol. 2023; 79:81-8.
16. Chiang JK, Lin YC, Hung TY, Kao HH, Kao YH. The impact on autonomic nervous system activity during and following exercise in adults: a meta-regression study and trial sequential analysis. Medicina (Kaunas). 2024; 60:1223.
17. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health. 2017; 5:258.
18. Billman GE. The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. Front Physiol. 2013; 4:26.
19. Danter WR, Carruthers SG. The heart rate–PR interval relationship: a model for evaluating drug actions on SA and AV nodal function. Br J Clin Pharmacol. 1990; 30:490-2.
20. Prescott SL, Liberles SD. Internal senses of the vagus nerve. Neuron. 2022; 110:579-99.
21. Capilupi MJ, Kerath SM, Becker LB. Vagus nerve stimulation and the cardiovascular system. Cold Spring Harb Perspect Med. 2020; 10:a034173.
22. Badran BW, Jenkins DD, Cook D, et al. Transcutaneous auricular vagus nerve stimulation–paired rehabilitation for oromotor feeding problems in newborns: an open-label pilot study. Front Hum Neurosci. 2020; 14:77.
23. Shekerdemian L, Bohn D. Cardiovascular effects of mechanical ventilation. Arch Dis Child. 1999; 80:475-80.
24. Taha BH, Simon PM, Dempsey JA, Skatrud JB, Iber C. Respiratory sinus arrhythmia in humans: an obligatory role for vagal feedback from the lungs. J Appl Physiol (1985). 1995; 78:638-45.
25. Gerritsen RJS, Band GPH. Breath of life: the respiratory vagal stimulation model of contemplative activity. Front Hum Neurosci. 2018; 12:397.

J Pediatr Res 2026;13(3):254-62
DOI: 10.4274/jpr.galenos.2025.55453

Original Article

Prevalence and Risk Factors Related to Cerebral Edema Associated with Diabetes Ketoacidosis in Pediatrics with Type-1 Diabetes Mellitus

Masoumeh Sarvari[1], Maedeh Sarvari[2], Reihaneh Mohsenipour[3], Mahsa Soti Khiabani[4,5], Mahya Sadat Mohammadi[5]

[1]Tehran University of Medical Sciences Endocrine and Metabolism Research Institute (EMRI), Tehran, Iran
[2]Qazvin University of Medical Science School of Medicine, Qazvin, Iran
[3]Tehran University of Medical Sciences, Growth and Development Research Center, Children's Medical Center, Department of Pediatrics, Tehran, Iran
[4]Children's Medical Center, Pediatrics Center of Excellence, Tehran, Iran
[5]Tehran University of Medical Sciences, Children's Medical Center, School of Medicine, Department of Pediatrics, Tehran, Iran

ABSTRACT

Aim: Cerebral edema associated with diabetes ketoacidosis (CEDKA) is the most important clinical challenge in children with type-1 diabetes mellitus (T1DM). CEDKA is associated with a high rate of morbidity and mortality. We investigated the prevalence and risk factors of CEDKA in children with T1DM.

Materials and Methods: This observational study was conducted at the Children's Medical Center Hospital, Tehran, in children up to the age of 18 years with T1DM who presented with DKA from December 2022 to June 2024. All of the patients' demographic, clinical, and laboratory data were collected and analyzed. A p value <0.05 was considered significant.

Results: In total, of the 427 patients included in this study, 204 (47.8%) were male, and their mean and standard deviation (SD) age was 8.11 (4.09) years. 100 patients (23.31%) had mild DKA, 149 (34.73%) had moderate DKA, and 178 (41.49%) had severe DKA. In this study, 8 patients (1.86%) had CE, 6 of these patients (75%) were male, and 2 of them (25%) were female, and their mean ± SD age was 9.12±2.64 years. In the analysis, factors such as pH levels [odds ratio (OR): 0.001, 95% confidence interval (CI): 0.00-0.44, p=0.00] and bicarbonate levels (OR: 0.55, 95% CI: 0.36-0.83, p=0.00) had an independent and significant relationship with the incidence of CE.

Conclusion: The prevalence of CE was higher in severe DKA patients. Rapid diagnosis and management of DKA patients, especially those with neurological symptoms, is the most critical step in T1DM.

Keywords: Type-1 diabetes mellitus, ketoacidosis, cerebral edema, pediatric, Iran

Corresponding Author
Masoumeh Sarvari, MD, Tehran University of Medical Sciences Endocrine and Metabolism Research Institute (EMRI), Tehran, Iran
E-mail: maasoomehsarvari@yahoo.com **ORCID:** orcid.org/0009-0000-4009-8485
Received: 29.07.2025 **Accepted:** 24.11.2025 **Epub:** 25.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Sarvari M, Sarvari Mae., Mohsenipour R, Khiabani MS, Mohammadi MS. Prevalence and risk factors related to cerebral edema associated with diabetes ketoacidosis in pediatrics with type-1 diabetes mellitus. J Pediatr Res. 2026;13(3):254-62

Introduction

Failure to properly manage blood glucose levels in diabetes patients can lead to critical conditions, including hypoglycemia and diabetes ketoacidosis (DKA). DKA is a combination of many pathophysiological cascades, such as hyperglycemia, insulin deficiency, metabolic acidosis, the accumulation of ketone bodies, and increased anion gap (1). Infection, female gender, newly diagnosed diabetes, lower socio-economic status, and depression are risk factors for DKA (2). Balancing fluids and electrolytes, controlling blood glucose levels, and eventually preventing complications are all crucial components of a multidisciplinary approach to DKA treatment (3).

Mismanagement or prolonged DKA leads to multi-organ damage and other life-threatening complications, such as cerebral edema (CE) (4). The prevalence of CE has been reported to be between 0.5-1%; however, some studies have reported a higher rate of 1.44% to 1.8% (5,6). CE is the most common cause of mortality in diabetic children, with a rate between 30-60% and morbidity with a rate between 21-35% (7,8). The detailed pathogenesis of CE in DKA is undetermined. However, there are some explanations in this area, including cellular osmolarity changes during fluid replacement for rehydration therapy, cerebral ischemia-hypoperfusion and reperfusion injury, and other cytotoxic and vasogenic events during DKA therapy (9-13). Unfortunately, only limited and heterogeneous information is available regarding aspects of CE, such as epidemiology and related risk factors in various ethnicities of DKA patients. Therefore, the current study aimed to investigate the prevalence of CE and its possible related risk factors in Iranian children with DKA.

Materials and Methods

Study Design and Data Gathering

This observational study was conducted at the Children's Medical Center Hospital in Tehran, Iran, from December 2022 to June 2024. All patients during this period were included as part of the census in our study. One of the inclusion criteria for patients was being aged under 18 years. Patients older than 18 years were excluded. Other conditions, including metabolic disorders and congenital abnormalities with similar manifestations of DKA, were ruled out in all age groups.

DKA was defined as a serum glucose concentration greater than 11 mmol/L (200 mg/dL), ketonemia (blood ß-hydroxybutyrate ≥3 mmol/L) or moderate or large ketonuria (urine ketones ≥2+), blood pH below 7.3, or a serum bicarbonate level below 18 mmol/L (14). Additionally, DKA can be classified into three categories based on the severity of acidosis: mild (venous pH ≤7.30; bicarbonate concentration ≤18 mmol/L), moderate (pH ≤7.2; bicarbonate ≤10), and severe (pH ≤7.1; bicarbonate ≤5) (3,13).

All procedures were performed based on the ethical standards of the institutional and national research committees. Written informed consent was obtained retrospectively from the parents or legal guardians of the pediatric patients included in the study and was documented in the patients' medical records. Ethical approval was obtained from Research Ethics Committees of Children's Medical Center-Tehran University of Medical Sciences (approval no.: IR.TUMS.CHMC.REC.1400.202, date: 23.12.2021).

Study Protocol

The medical records of those patients admitted to the hospital based on a diagnosis of ketoacidosis were used to collect the data. These medical records included all of the information we needed for our analysis. In order to extract the information required for this study, we designed a special form. We then completed the patients' hospital records, extracted the relevant information from our medical records, and entered it into the designed form. Finally, we transferred the data to a statistics software package. All of the patients included had venous blood lab tests assessing their baseline biochemical parameters carried out before any therapeutic intervention was provided.

CE Determination and Assessment

The diagnostic criteria for DKA were based on a paper from 2004 by Muir et al. (14). The determination of CE in the presence of DKA was conducted according to one diagnostic criterion, two major criteria, or one major and two minor criteria. The diagnostic criteria was abnormal motor or verbal response to pain, decorticate or decerebrate posture, cranial nerve palsy (especially III, IV, and VI), or abnormal neurogenic respiratory pattern (i.e. grunting, tachypnea, Cheyne-Stokes respiration, or apneusis). The major criteria were altered mental status, fluctuating levels of consciousness, sustained heart rate deceleration (a decrease of more than 20 beats per minute) not due to improved hydration or sleep, and age-inappropriate incontinence. The minor criteria were vomiting, headache, lethargy, diastolic blood pressure >90 mmHg, or age <5 years (14).

Fluid and Electrolyte Administration

In those children with volume depletion, fluid resuscitation was initiated with N/S (0.9%) infusion at 10-20 mL/kg for 20-30 min. Subsequent deficit fluids were replaced with 0.45-0.9% saline. One hour after intravenous fluid therapy, insulin infusion was initiated a dose of 0.05-0.1 U/kg/h of regular (soluble) insulin (15). Glucose-corrected blood sodium in mEq/L was based on the following formula: actual sodium in mEq/L-[(blood glucose in mg/dL-100) x 1.6/100] (14).

Statistical Analysis

The Statistical Package of Social Science Software (SPSS version 21, Chicago, USA) was used to analyze the data. Continuous and categorical variables are expressed as the mean and standard deviation and frequency (percentage), respectively. An independent t-test was used to compare continuous variables between the study groups. Chi-square and Fisher's exact tests were applied to compare categorical variables appropriately. A p value ≤0.05 was considered statistically significant. Additional analyses were conducted in order to determine relationships between the demographic and laboratory variables of the patients with CE.

Results

A total of 427 DKA patients were included in this study. The laboratory and demographic information on the severity of ketoacidosis is summarized in Table I. Mild ketoacidosis was present in 100 patients (23.35%), moderate ketoacidosis in 149 patients (34.9%), and severe ketoacidosis in 178 patients (41.68%). The severity of DKA was significantly greater in the male gender ($p<0.05$). Most patients in the age groups under one year and over ten years were in the severe category ($p<0.05$) with rates of 58.33% and 45.71%, respectively. The average age of those patients in the moderate group was greater, with a mean of 8.68±3.96 ($p<0.05$). The two most common symptoms, polyuria and polydipsia, had respective rates of 65.3% and 60.7%. Of all of the patients, 54 (12.58%) experienced signs of altered mental status, featuring obtundation or lethargy. Of all of the patients, only one individual had the symptom of papilledema and was concomitantly diagnosed with a severe form of DKA. One hundred and ninety-one patients (44.7%) were admitted to intensive care units (ICU). More details are shown in Table I. Laboratory data, including pH, pCO_2, bicarbonate, sodium, potassium, glucose, creatinine, and blood urea nitrogen levels, are given in Table I based on the severity of DKA.

Prevalence of CE in DKA

Table II presents the complications of the patients categorized according to the severity of ketoacidosis. In all of the examined patients, 8 patients (1.87%) were diagnosed with CE, with one patient (12.5%) having moderate DKA, and 7 patients (87.5%) being diagnosed with severe DKA ($p<0.05$). Fortunately, none of the CE patients expired during this investigation. Regardless of altered mental status and ICU admission, none of the CE cases needed to be intubated. Seven patients had CE symptoms during admission, and only one patient developed CE hours after. One of the CE cases had a history of DKA ($p<0.05$). After establishing CE diagnosis, all patients received mannitol, but none of them received bicarbonate therapy. Table III shows the demographic, clinical, and laboratory factors in the DKA patients according to CE. Levels of pH were significantly related to the occurrence of CE ($p<0.05$) (Table III).

In the analysis, it was shown that only the factors of pH [odds ratio (OR): 0.001, 95% confidence interval (CI): 0.00-0.44) and bicarbonate (OR: 0.55, 95% CI: 0.36-0.83) had an independent relationship with the occurrence of CE associated with DKA (CEDKA) ($p=0.000$ for both). Additionally, the age group of patients younger than 5 years old (OR: 2.71, 95% CI: 0.33-22.28, $p=0.032$) was significantly related to CE incidence (Table IV).

Table I. Demographic, clinical symptoms, and laboratory findings of the included patients

Variable	Total (n=427)	Mild DKA (n=100)	Moderate DKA (n=149)	Severe DKA (n=178)	p value
Demographic					
Age, (mean ± SD)	8.11 (4.09)	7.17 (4.18)	8.68 (3.96)	8.17 (4.07)	<0.05
Age groups, n (%)					
<1 year	24 (5.59)	6 (6)	4 (2.68)	14 (7.86)	<0.05
1 year-less than 5 years	94 (21.91)	34 (34)	27 (18.12)	33 (18.53)	
5 years-less than 10 years	171 (39.85)	36 (36)	68 (45.63)	67 (37.64)	
10 years or older	140 (32.63)	24 (24)	50 (33.55)	64 (35.95)	
Sex					
Male, n (%)	204 (47.8)	58 (58)	52 (25.5)	94 (46.1)	<0.05
Female, n (%)	223 (52.2)	42 (42)	97 (43.5)	84 (37.7)	
Weight, (kg) (SD)	28.55 (14.16)	26.57 (13.31)	30.59 (15.37)	27.93 (13.40)	<0.05
New onset of T1DM, n (%)	291 (67.73)	72 (72)	95 (63.75)	124 (69.66)	0.334
History of DM >1 month, n (%)	136 (31.70)	28 (28)	54 (36.24)	54 (30.33)	0.381
Misdiagnosis in EM, n (%)	0	0	0	0	
Symptoms		**n=100**	**n=149**	**n=178**	
Polyuria, n (%)	279 (65.3)	56 (20.1)	85 (30.4)	138 (49.4)	<0.001
Polydipsia, n (%)	259 (60.7)	44 (17)	77 (29.7)	138 (53.3)	<0.001
Polyphagia, n (%)	191 (44.7)	20 (10.4)	56 (29.3)	115 (60.2)	<0.001
Vomiting, n (%)	55 (12.9)	7 (12.7)	21 (38.2)	27 (49)	0.128
Headache, n (%)	64 (14.9)	4 (6.2)	12 (18.7)	48 (75)	<0.001
Lethargy/obtundation, n (%)	54 (12.58)	3 (5.5)	8 (14.8)	43 (79.6)	<0.001
Fatigue/irritability, n (%)	58 (13.6)	9 (15.5)	22 (37.9)	27 (46.5)	0.332
Abdominal pain, n (%)	112 (26.2)	22 (19.6)	31 (27.6)	59 (52.6)	0.022
Papilledema, n (%)	1 (0.2)	0	0	1 (0.2)	0.421
Focal neurological signs, n (%)	0	0	0	0	
ICU admission, n (%)	191 (44.7)	20 (10.4)	56 (29.3)	115 (60.2)	<0.001
Laboratory					
pH (unit) mean ± SD	7.19±0.14	7.33±0.06	7.23±0.07	7.07±0.12	<0.001
pCO_2 (mm Hg) mean ± SD	13.2±7.02	18.2±6.32	13.6±4.57	8.7±1.84	0.005
Bicarbonate (mmol/l) mean ± SD	9.04±4.98	15.4±3.73	9.6±3.16	5.0±2.02	<0.001
Sodium (mmol/L) mean ± SD	135.35±5.03	135.06±4.82	135.70±4.72	135.22±5.40	0.553
Potassium (mmol/dL) mean ± SD	4.28±0.7	4.36±0.67	4.24±0.66	4.27±0.72	0.441
Glucose (mg/dL) mean ± SD	462.4±142.9	432±152.2	445.5±130.8	493.6±141.7	<0.001
BUN (mg/dL) mean ± SD	13±3.36	13±3.08	12±2.87	13±3.01	0.092
Creatinine (mg/dl) mean ± SD	1.06±3.84	0.8±0.22	4.24±6.48	0.9±0.32	<0.001

DKA: Diabetic ketoacidosis, SD: Standard deviation, T1DM: Type 1 diabetes mellitus, DM: Diabetes mellitus, EM: Emergency department, ICU: Intensive care unit, pCO_2: Partial pressure of carbon dioxide, BUN: Blood urea nitrogen

Table II. Prognosis of DKA patients

Variable	Total n=427	Mild DKA n=100	Moderate DKA n=149	Severe DKA n=178	p value
Hospitalization duration (days), mean ± SD (range)	3.39±1.6 (range: 1-16)	2.88±1.46 (1-9)	3.2±1.21 (1-8)	3.86±1.81 (1-16)	<0.001
DKA resolution, (days), mean ± SD (range)	1.53±0.85 (range: 0.5-9)	1.22±0.6 (0.5-4)	1.43±0.7 (1-4)	1.8±1.02 (1-9)	<0.001
ICU admission, (days) mean ± SD (range)	0.93±1.41 (range: 0-10)	0.38±0.9 (0-4)	0.72±1.11 (0-5)	1.42±1.7 (1-10)	<0.001
Cerebral edema, n (%)	8 (1.87%)	0	1 (12.5%)	7 (87.5%)	0.023
Death, n (%)	0	0	0	0	

DKA: Diabetic ketoacidosis, ICU: Intensive care unit, SD: Standard deviation

Table III. Prevalence of cerebral edema according to treatment fluid protocol

Variable	Absent CE (n=419)	Present CE (n=8)	p value
Sex Male, n (%) Female, n (%)	 198 (47.25) 221 (52.74)	 6 (75) 2 (25)	0.115
Age, mean ± SD	8.09±4.11	9.12±2.64	0.481
Age range years, n (%) 1-5 years, n (%) 5-10 years, n (%) 10-18 years, n (%)	 24 (5.72) 93 (22.19) 167 (39.85) 135 (32.21)	 0 1 (12.50) 4 (50) 3 (37.5)	0.782
New cases of diabetes, n (%)	286 (68.25)	5 (62.5)	0.490
Altered mental status, n (%)	46 (10.97)	8 (100)	<0.001
BS reduction of more than 100, n (%)	207 (49.40)	5 (62.5)	0.354
BS less than 70 (mg/dL), n (%)	48 (11.45)	2 (25)	0.233
Initial insulin (IU/kg/h) 0.01-0.02, n (%) 0.03-0.05, n (%) 0.06-0.08, n (%) 0.09-0.1, n (%)	 3 (0.71) 88 (21) 45 (10.73) 283 (67.54)	 0 0 0 8 (100)	0.286
Na quartile (mmol/L) 110-134, n (%) 135-145, n (%) 146-160, n (%)	 174 (41.52) 234 (55.84) 11 (2.62)	 6 (75) 2 (25) 0	0.163
pH quartile 6.6-7.1, n (%) 7.11-7.2, n (%) 7.21-7.3, n (%)	 82 (19.57) 114 (27.20) 223 (53.22)	 7 (87.5) 1 (12.5) 0	<0.001
HCO_3 quartile (mmol/L) 1-4, n (%) 5-9, n (%) 10-25, n (%)	 357 (85.20) 47 (11.21) 15 (3.57)	 8 (100) 0 0	0.501
Potassium quartile (mmol/L) 1.0-3.5, n (%) 3.6-5.5, n (%) 5.6-15, n (%)	 51 (12.17) 362 (86.39) 6 (1.43)	 2 (25) 6 (75) 0	0.524

CE: Cerebral edema, SD: Standard deviation, BS: Blood sugar, IU: International unit, Na: Sodium, HCO_3: Bicarbonate, pH: Potential of hydrogen

Table IV. Adjusted odds ratios and 95% confidence intervals with significant statistical results (p<0.050) for cerebral edema occurrence

Risk factor	OR (95% CI)	p value
New case of type-1 diabetes	0.77 (0.18-3.29)	0.496
Sex	3.34 (0.66-16.78)	0.115
Age <5-year-old	2.71 (0.33-22.28)	0.032
Weight	1.02 (0.97-1.07)	0.381
Insulin	2.02 (0.00-8.21)	0.224
pH	0.001 (0.00-0.44)	<0.001
HCO_3	0.55 (0.36-0.83)	<0.001
pCO_2	0.83 (0.78-1.12)	0.980
BS less than 70 mg/dL	0.39 (0.076-1.98)	0.238
Hyponatremia (Na <135 mmol/L)	0.23 (0.05-1.18)	0.063
Hypokalemia (K <3.5 mmol/L)	0.416 (0.08-2.11)	0.261

OR: Odds ratio, CI: Confidence interval, HCO_3: Bicarbonate, pCO_2: Partial pressure of carbon dioxide, pH: Potential of hydrogen, BS: Blood sugar

Discussion

As noted, despite extensive research on CE in DKA patients, the evidence from several domains including epidemiology and risk factors associated with the incidence of CEDKA is markedly variable. This variability leads to challenges in managing this patient group. As of the time of writing, no specific study had focused on the incidence of CEDKA in Iran, making this a relatively comprehensive report on Iranian children with DKA in this field. In the current study, we investigated the prevalence of CEDKA among children under 18 years and its association with various risk factors. We reviewed data from 427 patients with DKA, with an average age of 8.11 years and a female predominance. Most of these cases, 291 (67.73%), were of new-onset diabetes. The severity of DKA was significantly associated with male gender, age, and weight. In this study, 8 patients (1.87%) experienced CEDKA; of these, 7 had severe DKA, and one had moderate DKA. Loss of consciousness, being aged <5 years, bicarbonate levels and pH levels were significantly related to CE incidence in this study.

Age is one of the significant factors in DKA occurrence. The mean age in our study was 8.11 years, consistent with other findings: Del Pozo et al. (16) reported an average age of 7.2 years, and Sultana et al. (17) in Pakistan reported 7.86 years. Conversely, an Indian study noted an average DKA onset age of 9.5 years (18). The average age in another study in Jordan was reported to be 11.03±3.88 (19). Furthermore, a Polish investigation demonstrated that children aged 0-2 years had the highest incidence of DKA and that age was one of the factors linked to DKA (20).

Various rates of new cases of diabetes have been reported in different articles. For instance, 28.7% of patients newly diagnosed with type-1 diabetes mellitus (T1DM) had DKA in the study by McKenna et al. (21). In other studies, with findings similar to ours, Del Pozo et al. (16) and Aminzadeh et al. (22) reported DKA rates of 67.4% and 67.3% in newly diagnosed patients, respectively. New onset of T1DM has been highlighted as one of the risk factors for CEDKA in the literature (23,24); however, our analyses showed that there was no association between them. A study by Sultana et al. (17) also reported a 59% rate of new onset diabetes with no association between the incidence of new cases of diabetes and CE (p=0.27) (17).

Aligning with our results, Sehgal et al.'s (25) study emphasized a positive relation for the age group of younger than 5 years with CEDKA. Also focusing on the issue of age, Edge et al. (24) found an average patient age of 8.5 years as being an independent risk factor.

The prevalence of CEDKA reported in previous studies varies. For example, in Edge et al.'s (24) study of 2,940 DKA cases, 34 (1.15%) had CE. Another study reported prevalence rates of 1.44% (15). González Pannia et al. (5) studied 693 DKA patients aged 1-18 years and found 10 cases of CE. Totally, CE occurs in about 0.3% to 1% of DKA cases (26). In Iran, no large-scale studies on DKA and CE in pediatrics have been conducted, so comparisons are limited. The different prevalence rates across studies are likely due to heterogeneity in risk factors. Our study aimed to explore most of these factors, including demographic, clinical, and laboratory data. Additionally, as an academic referral center, we see the most complex

cases from across the country, which may have influenced our findings.

The acidosis mechanisms proposed include hypercapnia caused by metabolic acidosis, leading to vasoconstriction, dehydration, reduced cerebral blood flow, and cytotoxic edema, as well as reperfusion injury and vasogenic edema during rehydration (27,28). Laboratory studies support these mechanisms, showing that hypoperfusion and reperfusion during DKA treatment can alter diffusion-weighted imaging (DWI) in magnetic resonance imaging (MRI), indicating vasogenic edema (29). DWI-MRI changes related to vasogenic edema have been observed in children with DKA, supporting the concept of reperfusion injury.

In our study, pH (OR: 0.001, 95% CI: 0.00-0.44) and HCO_3 (OR: 0.55, 95% CI: 0.36-0.83) were associated with increased CEDKA risk. All patients with CEDKA had bicarbonate levels between 1-4 mmoL/dL. A body of literature has reported a similar relation between acidosis and CEDKA. For instance, Yaneva et al. (30) studied 256 children with DKA and found a direct relationship between acidosis severity and CE, although their reported prevalence was 8.6%, higher than most other studies, with unclear reasons. Also, their research showed that the severity of DKA was significantly associated with low bicarbonate levels and with higher initial blood glucose (30). Kumar et al. (31) reported that CEDKA was significantly associated with pH and bicarbonate levels. Prior studies, such as Marcin et al. (32), found that the severity of neurological complications was related to low bicarbonate levels, and other factors, such as low pCO_2 (<22 mmoL/dL), influenced outcomes. Glaser et al. (33) also reported similar findings in 2001 with 6,977 children: 0.9% experienced CE, and factors such as high initial nitrogen, low pCO_2, and bicarbonate therapy were related to CEDKA. Regarding fluid therapy, our analyses showed no significant relationship between therapy duration and CEDKA occurrence, consistent with the studies by Kadhim (28) and Lam et al. (29), which also found no significant link.

While hypoperfusion and cerebral ischemia are considered primary mechanisms for CE, the precise pathogenesis remains unclear. The roles of fluid therapy and osmotic shifts have been investigated with conflicting results. Some studies, such as Glaser et al. (34) in 2008, suggest that osmotic changes contribute to CE, whereas others highlight widespread osmolality alterations in DKA patients (35), with hyperosmolar states leading to CEDKA (36). Rapid and precise fluid management is crucial, yet no comprehensive protocol exists, as fluid therapy varies in type, tonicity, dose, and duration across studies, complicating the assessment of osmolality and volume effects in these patients (37).

There is a strong link between CE and ICU admission in DKA patients ($p<0.001$). Patients who present with severe metabolic disturbances, dehydration, and altered mental status require ICU care. Other studies have emphasized that CEDKA patients require longer hospital stays (31,38).

Study Limitations

For the first time in Iran, we investigated the prevalence of CE in a relatively large-scale cohort of children with DKA. In this study, a relatively large number of risk factors for CE, including low sodium levels, acidosis, low bicarbonate levels, and serum urea nitrogen concentrations, were examined. Among the limitations of our study, firstly, we could not establish a cause-and-effect relationship or analyze changes over time due to this study's cross-sectional design. Additionally, the lack of a control group limited the ability to better compare results, and the single-center population prevented us from exploring the effects of different protocols and patient demographics on the occurrence of CE.

Conclusion

The predominance of CE is higher in severe DKA patients. In this study, pH levels and bicarbonate levels had an independent and significant relationship with the incidence of CE. Conducting large-scale epidemiologic and multicenter studies may shed further light on this field, as well as help in developing a comprehensive treatment protocol for fluid therapy and determining the exact prevalence of CE and other complications of DKA.

Ethics

Ethics Committee Approval: Ethical approval was obtained from Research Ethics Committees of Children's Medical Center-Tehran University of Medical Sciences (approval number: IR.TUMS.CHMC.REC.1400.202, date: 23.12.2021).

Informed Consent: Written informed consent was obtained retrospectively from the parents or legal guardians of the pediatric patients included in the study and was documented in the patients' medical records.

Acknowledgments

I want to express my profound gratitude to all the professors and colleagues of the Children's Hospital at Tehran University of Medical Sciences who contributed to the completion of this project.

Footnotes

Authorship Contributions

Concept: M.S., M.S.M., Design: M.S.K., Analysis or Interpretation: R.M., Literature Search: Mae.S., Writing: Mae.S.

Conflict of Interest: The authors have no conflicts of interest to declare. All co-authors have seen and agree with this manuscript's contents, and there is no financial interest to report. We certify that this submission is original work and is not under review at any other publication.

Financial Disclosure: The authors received no financial support for the research, authorship, and/or publication of this article.

References

1. Azad C, Kaur R, Singh DK, Arya A. Stroke in pediatric diabetic ketoacidosis: case series. J Trop Pediatr. 2017; 63:324-7.
2. Ahuja W, Kumar N, Kumar S, Rizwan A. Precipitating risk factors, clinical presentation, and outcome of diabetic ketoacidosis in patients with type 1 diabetes. Cureus. 2019; 11:e4789.
3. Glaser N, Fritsch M, Priyambada L, et al. ISPAD clinical practice consensus guidelines 2022: diabetic ketoacidosis and hyperglycemic hyperosmolar state. Pediatr Diabetes. 2022; 23:835-56.
4. Scutca AC, Nicoară DM, Mang N, Jugănaru I, Brad GF, Mărginean O. Correlation between neutrophil-to-lymphocyte ratio and cerebral edema in children with severe diabetic ketoacidosis. Biomedicines. 2023; 11:2976.
5. González Pannia P, Balboa R, Navarro R, Nocita MF, Ferraro M, Mannucci C. Prevalence of cerebral edema among diabetic ketoacidosis patients. Arch Argent Pediatr. 2020; 118:332-6.
6. Ferraro M, Ozuna B, Mazza C, et al. Estudio multicéntrico de cetoacidosis diabética en servicios pediátricos de Argentina. Rev Soc Argent Diabetes. 2009; 43:127-36.
7. Agrawal S, Baird GL, Quintos JB, et al. Pediatric diabetic ketoacidosis with hyperosmolarity: clinical characteristics and outcomes. Endocr Pract. 2018; 24:726-32.
8. Cook AM, Morgan Jones G, Hawryluk GWJ, et al. Guidelines for the acute treatment of cerebral edema in neurocritical care patients. Neurocrit Care. 2020; 32:647-66.
9. Alzamzami AA, Alshab DA, Alrabiea MA, et al. Diabetic ketoacidosis in pediatric patients: fluid management and cerebral edema risk. J Healthc Sci. 2025; 5:213-20.
10. Watts W, Edge JA. How can cerebral edema during treatment of diabetic ketoacidosis be avoided?. Pediatr Diabetes. 2014; 15:271-6.
11. Szmygel Ł, Kosiak W, Zorena K, Myśliwiec M. Optic nerve and cerebral edema in the course of diabetic ketoacidosis. Curr Neuropharmacol. 2016; 14:784-91.
12. Vavilala MS. Treating cerebral edema in diabetic ketoacidosis: caveats in extrapolating from traumatic brain injury. Pediatr Crit Care Med. 2013; 14:723-4.
13. Wolfsdorf JI, Glaser N, Agus M, et al. ISPAD clinical practice consensus guidelines 2018: diabetic ketoacidosis and the hyperglycemic hyperosmolar state. Pediatr Diabetes. 2018; 19(Suppl 27):155-77.
14. Muir AB, Quisling RG, Yang MC, Rosenbloom AL. Cerebral edema in childhood diabetic ketoacidosis: natural history, radiographic findings, and early identification. Diabetes Care. 2004; 27:1541-6.
15. Wolfsdorf JI, Allgrove J, Craig ME, et al. ISPAD Clinical Practice Consensus Guidelines 2014. Diabetic ketoacidosis and hyperglycemic hyperosmolar state. Pediatr Diabetes. 2014; 15(Suppl 20):154-79.
16. Del Pozo P, Aránguiz D, Córdova G, et al. Clinical profile of children with diabetic ketoacidosis in fifteen years of management in a Critical Care Unit. Rev Chil Pediatr. 2018; 89:491-8.
17. Sultana N, Bari A, Sarwar M, et al. Risk factors of cerebral edema in diabetic ketoacidosis admitted in PICU of a tertiary care hospital. Pak Pediatr J. 2021; 45:457-62.
18. Raghunathan V, Jevalikar G, Dhaliwal M, et al. Risk factors for cerebral edema and acute kidney injury in children with diabetic ketoacidosis. Indian J Crit Care Med. 2021; 25:1446-51.
19. Odeh R, Gharaibeh L, Daher A, et al. Frequency, clinical characteristics and predictors of ketoacidosis at diagnosis of type one diabetes mellitus in children and adolescents from Jordan. J Clin Res Pediatr Endocrinol. 2023; 15:46-54.
20. Szypowska A, Dżygało K, Wysocka-Mincewicz M, et al. High incidence of diabetic ketoacidosis at diagnosis of type 1 diabetes among Polish children aged 10-12 and under 5 years of age: a multicenter study. Pediatr Diabetes. 2017; 18:722-8.
21. McKenna A, Sandys N, Ryder K, et al. Single centre experience - clinical presentation and frequency of paediatric diabetic ketoacidosis (DKA) at diagnosis over a 5-year period. Ir Med J. 2018; 111:714.
22. Aminzadeh M, Navidi N, Valavi E, Aletayeb SMH. Childhood onset type 1 diabetes at a tertiary hospital in south-western Iran during 2000-2015: Rapid increase in admissions and high prevalence of DKA at diagnosis. Prim Care Diabetes. 2019; 13:43-8.
23. Jafari H, Dadashian P, Ghaffari J, Zamanfar. Predisposing factors of cerebral edema in pediatric diabetic ketoacidosis: a systematic review. J Pediatr Rev. 2022; 10:29-38.
24. Edge JA, Hawkins MM, Winter DL, Dunger DB. The risk and outcome of cerebral oedema developing during diabetic ketoacidosis. Arch Dis Child. 2001; 85:16-22.
25. Sehgal M, Batra M, Jha P, Sanchez O. Risk factors and laboratory findings associated with diabetic ketoacidosis in hospitalized pediatric patients. Cureus. 2022; 14:e25410.
26. Azova S, Rapaport R, Wolfsdorf J. Brain injury in children with diabetic ketoacidosis: review of the literature and a proposed pathophysiologic pathway for the development of cerebral edema. Pediatr Diabetes. 2021; 22:148-60.
27. Glaser NS, Wootton-Gorges SL, Kim I, et al. Regional brain water content and distribution during diabetic ketoacidosis. J Pediatr. 2017; 180:170-6.
28. Kadhim FI. The etiology of cerebral edema in DKA managed patients [Bachelor's thesis]. Diyala, Iraq: University of Diyala, College of Medicine; 2022.
29. Lam TI, Anderson SE, Glaser N, O'Donnell ME. Bumetanide reduces cerebral edema formation in rats with diabetic ketoacidosis. Diabetes. 2005; 54:510-6.

30. Yaneva NY, Konstantinova MM, Iliev DI. Risk factors for cerebral oedema in children and adolescents with diabetic ketoacidosis. Biotechnol Biotechnol Equip. 2016; 30:1142-7.

31. Kumar R, Kumari S, Chandra V, Singh A, Ansari A, Jha G. Initial laboratory characteristics and treatment related factors in relation to occurrence of cerebral edema in children with diabetic ketoacidosis-experience of a tertiary care center. Int J LifeSci Biotechnol Pharma Res. 2023; 12:1457-61.

32. Marcin JP, Glaser N, Barnett P, et al.; The Pediatric Emergency Medicine Collaborative Research Committee. Factors associated with adverse outcomes in children with diabetic ketoacidosis-related cerebral edema. J Pediatr. 2002; 141:793-7.

33. Glaser N, Barnett P, McCaslin I, et al. Risk factors for cerebral edema in children with diabetic ketoacidosis. The Pediatric Emergency Medicine Collaborative Research Committee of the American Academy of Pediatrics. N Engl J Med. 2001; 344:264-9.

34. Glaser NS, Marcin JP, Wootton-Gorges SL, et al. Correlation of clinical and biochemical findings with diabetic ketoacidosis-related cerebral edema in children using magnetic resonance diffusion-weighted imaging. J Pediatr. 2008; 153:541-6.

35. Long B, Koyfman A. Emergency medicine myths: cerebral edema in pediatric diabetic ketoacidosis and intravenous fluids. J Emerg Med. 2017; 53:212-21.

36. Hassan EM, Mushtaq H, Mahmoud EE, et al. Overlap of diabetic ketoacidosis and hyperosmolar hyperglycemic state. World J Clin Cases. 2022; 10:11702-11.

37. Patino-Galarza D, Duque-Lopez A, Cabra-Bautista G, Calvache JA, Florez ID. Fluids in the treatment of diabetic ketoacidosis in children: a systematic review. J Evid Based Med. 2024; 17:317-28.

38. Patel A, Singh D, Bhatt P, Thakkar B, Akingbola OA, Srivastav SK. Incidence, trends, and outcomes of cerebral edema among children with diabetic ketoacidosis in the United States. Clin Pediatr (Phila). 2016; 55:943-51.

J Pediatr Res 2026;13(3):263-71
DOI: 10.4274/jpr.galenos.2026.58630

Original Article

Virus-Specific Patterns of Early Hospitalization after Pediatric Emergency Department Discharge: A Real-World Cohort Study

Fırat Bedir, Mahmut Altuntaş

Bahçelievler Medipol Hospital, Clinic of Pediatrics, İstanbul, Türkiye

ABSTRACT

Aim: To compare virus-specific clinical patterns predictive of early hospitalization within 7 days following emergency department (ED) discharge among children with polymerase chain reaction (PCR)-confirmed coronavirus disease-2019 (COVID-19), respiratory syncytial virus (RSV), and influenza.

Materials and Methods: This retrospective cohort study included children aged 0-18 years with PCR-confirmed COVID-19, RSV, or influenza who presented to the pediatric ED of a tertiary-care university hospital between January, 2020 and December, 2024 and who were initially discharged at the index visit. The primary outcome was hospitalization within 7 days following ED discharge. Multivariable logistic regression models were constructed separately for each viral etiology in order to identify virus-specific predictors of early hospitalization within 7 days.

Results: Among the 6,700 children included, hospitalization within 7 days occurred in 7.8% of COVID-19 cases, 9.0% of RSV cases, and 7.0% of influenza cases. In the COVID-19 cases, C-reactive protein ≥33 mg/L [adjusted odds ratio (aOR) 2.8, 95% confidence interval (CI): 1.7-4.3] and persistent fever longer than 3 days (aOR 3.2, 95% CI: 1.9-4.9) were independently associated with early hospitalizations. In the RSV cases, respiratory distress was the strongest predictor (aOR 3.7, 95% CI: 2.2-5.9). In the influenza cases, persistent systemic symptoms were independently associated with hospitalizations (aOR 3.1, 95% CI: 1.8-4.7).

Conclusion: Early hospitalization following ED discharge shows distinct clinical patterns across viral etiologies in children. Incorporating etiology-informed, readily available clinical features at ED presentation into discharge planning may improve the safety of outpatient management and support targeted early follow-up.

Keywords: Respiratory viral infections, early hospitalization, children, emergency department, risk factors, COVID-19, respiratory syncytial virus, influenza

Introduction

Respiratory viral infections are one of the leading causes of emergency department (ED) visits and hospitalizations among children worldwide (1,2). In the post-pandemic era, coronavirus disease-2019 (COVID-19), respiratory syncytial virus (RSV), and influenza represent the most common causes of pediatric respiratory viral illnesses. Although these infections often present with overlapping clinical features at initial evaluation, recent evidence demonstrates that their early clinical trajectories differ substantially (3-5).

Unplanned revisits or hospitalizations shortly after ED discharge have emerged as a marker of suboptimal early

Corresponding Author
Fırat Bedir, MD, Bahçelievler Medipol Hospital, Clinic of Pediatrics, İstanbul, Türkiye
E-mail: leheng65.fb@gmail.com **ORCID:** orcid.org/0000-0002-8100-8498
Received: 03.03.2026 **Accepted:** 17.07.2026 **Epub:** 31.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Bedir F, Altuntaş M. Virus-specific patterns of early hospitalization after pediatric emergency department discharge: a real-world cohort study. J Pediatr Res. 2026;13(3):263-71

etiology-specific risk assessment rather than disease severity alone. Population-based data have demonstrated that the timing and location of pediatric ED revisits vary substantially, with the majority occurring within the first 72 hours and frequently at institutions different from the index visit (6). Moreover, critical revisits [those resulting in intensive care unit (ICU) admission or death] disproportionately affect children who received respiratory diagnoses at the index visit (7). Although earlier studies have identified predictors of return visits in pediatric emergency populations (8,9), the first week following the initial presentation represents a clinically vulnerable period during which subtle clinical deterioration may emerge, particularly in those children initially classified as having a mild disease.

Pediatric COVID-19 has been associated with systemic inflammatory responses and prolonged fever, even in the absence of significant respiratory involvement (10-12). Emerging post-pandemic data from large electronic health record cohorts have further characterized the acute and post-acute clinical burden of severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) infection in children, confirming that illness severity and younger age are associated with higher risks of complications (13). By contrast, RSV-related deterioration is largely driven by host-virus interaction at the airway level, resulting in progressive lower respiratory tract involvement, particularly in young infants (14,15). Influenza shows an intermediate clinical pattern, with persistent systemic symptoms frequently leading to repeat healthcare visits rather than hypoxemia (16-18).

However, most of the existing studies evaluate these viral infections in isolation and are often limited to single-center cohorts, thus limiting direct comparison. The 2022 simultaneous respiratory viral surge highlighted the clinical and operational challenges of concurrent pediatric viral epidemics in emergency settings (19). Consequently, early post-discharge risk assessment is frequently extrapolated across viral etiologies, despite the limited evidence supporting the validity of a uniform approach. In order to address this gap, we conducted a retrospective cohort study to compare virus-specific clinical patterns predictive of early hospitalizations within a standardized 7-day timeframe among children with COVID-19, RSV, and influenza.

Materials and Methods

Ethical Considerations

This study was approved by the İstanbul Medipol University Non-Interventional Clinical Research Ethics Committee (approval no.: E-10840098-202.3.02-8228, date: 04.12.2025) and conducted in accordance with the principles of the Declaration of Helsinki.

Study Design and Setting

This retrospective cohort study included children aged 0-18 years with polymerase chain reaction (PCR)-confirmed COVID-19, RSV, or influenza who presented to a pediatric ED between January 2020 and December 2024. This study was conducted at a tertiary-care university hospital complex. Although data were obtained from two hospital campuses, both are affiliated to the same academic institution, share unified clinical protocols, and were therefore analyzed as a single-institution cohort. Discharge decisions were based on institutional pediatric ED protocols including clinical stability, oxygen saturation, oral intake adequacy, and physician assessment.

Participants

Children with PCR-confirmed COVID-19, RSV, or influenza who were initially discharged from the ED at the index visit were eligible for inclusion. Only the first presentation per infection episode was analyzed; a new infection episode was defined as an acute respiratory illness with PCR-confirmed viral detection occurring at least 14 days after a prior confirmed episode. Those patients who were admitted at the initial visit, transferred to another facility, or had incomplete outcome data were excluded. A total of 8,942 children with PCR-confirmed COVID-19, RSV, or influenza were identified during the study period. Of these, 1,204 were admitted at the initial visit, 386 were transferred to another facility, 491 had incomplete follow-up data, and 161 (1.8%) had multiple simultaneous viral detections, yielding a final analytic cohort of 6,700 children (Supplementary Figure 1). Those children with a simultaneous detection of more than one respiratory viral pathogen were excluded from the primary analysis in order to ensure that clinical patterns could be attributed to a single viral etiology. The patient selection process is summarized in a flow diagram (Supplementary Figure 1).

Data Collection

Demographic characteristics, comorbidities, and acute clinical findings assessed during the index ED visit, including fever duration, respiratory distress, oxygen saturation, and feeding difficulties, were extracted from the electronic medical records. Respiratory distress was defined as the presence of tachypnea (based on age-adjusted thresholds), subcostal or intercostal retractions, nasal flaring, or the use of the accessory muscles of breathing. Feeding difficulties

were defined as documented oral intake of less than 50% of the usual intake at the time of ED assessment. Laboratory parameters, including C-reactive protein (CRP), were recorded where available. Systemic symptoms were defined as a composite variable reflecting the presence of at least one of the following at the index ED visit: persistent fever longer than 48 hours, myalgia, severe headache, marked malaise, or the documented need for repeated antipyretic use. Fever duration longer than 3 days was additionally analyzed as a separate standalone predictor variable, distinct from the composite systemic symptom definition.

CRP testing was not protocolized and was performed at the discretion of the treating emergency physician. Missing data were handled using complete-case analysis, and the proportion of missing laboratory data reflected routine physician-directed testing rather than a protocolized assessment. This pragmatic approach reflects real-world clinical practice and enhances the external validity of CRP-based findings, although it introduces the possibility of indication bias in CRP measurements. CRP-tested and non-tested patients were not found to differ systematically in baseline clinical severity, as assessed by the attending physician's documented rationale; however, formal propensity adjustment was not performed.

Outcome Definition

The primary outcome was hospitalization within 7 days following the initial ED discharge. The 7-day timeframe was selected a priori in order to capture early post-discharge clinical deterioration most plausibly related to ED disposition decisions, rather than later disease progression or unrelated healthcare utilization. This 7-day interval has been commonly used in pediatric emergency medicine studies evaluating early revisit and hospitalization outcomes (6-9). Hospitalizations were considered outcome events only when clinically related to the index viral respiratory illness based on the discharge diagnosis documentation in the electronic health record. Outcome attribution was determined by a structured review of the discharge diagnostic codes in the institutional EHR; no external adjudication panel was used, and those cases with ambiguous diagnostic coding were reviewed by a senior clinician in order to confirm relatedness to the index viral illness. Outcome data were ascertained through the integrated institutional electronic health record system, which covers all hospital campuses and affiliated units within the university network.

Statistical Analysis

Multivariable logistic regression models were constructed separately for each viral etiology in order to identify virus-specific predictors of early hospitalizations. All models were adjusted for age, sex, and comorbidity status. Age-stratified analyses (<2 years vs. ≥2 years) were performed. The CRP cut-off value of 33 mg/L was determined using receiver operating characteristic curve analysis within the COVID-19 cohort [area under the curve (AUC) 0.73, 95% confidence interval (CI): 0.70-0.76]. Model discrimination was assessed using the AUC separately for each viral cohort: the COVID-19 model demonstrated acceptable performance (AUC 0.73, 95% CI: 0.70-0.76), the RSV model showed acceptable performance (AUC 0.76, 95% CI: 0.73-0.79), and the influenza model showed acceptable-to-good performance (AUC 0.79, 95% CI: 0.76-0.82). These values are consistent with the exploratory nature of the analysis and the inherent heterogeneity of clinical presentations across viral etiologies. Sensitivity analyses restricted to patients with available CRP measurements demonstrated similar direction and magnitude of associations. All statistical analyses were performed using SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). A two-sided p value <0.05 was considered statistically significant. Age was evaluated both as a continuous variable and as a dichotomized variable (age <2 years versus ≥2 years) in order to better capture clinically relevant pediatric risk stratification; the dichotomized parameterization was retained in the final models as it reflected a clinically meaningful threshold established in the pediatric respiratory illness literature and was consistent with the age-related vulnerability patterns observed across viral etiologies. Given the exploratory nature of the analysis with separate models for each viral etiology, no formal correction for multiple comparisons was applied; the results should be interpreted as hypothesis-generating and individual associations should be considered in conjunction with their effect sizes and CIs rather than for the p-values alone. The number of outcome events per viral cohort (COVID-19, n=156; RSV, n=198; influenza, n=175) satisfied recommended events-per-variable assumptions for stable multivariable logistic regression estimation.

Results

Study Population

A total of 6,700 children with PCR-confirmed viral respiratory infections were included in the final analysis (COVID-19, n=2,000; RSV, n=2,200; influenza, n=2,500). All patients were discharged from the ED at the index visit.

Baseline demographic and clinical characteristics varied across the viral etiologies, particularly with respect to age distribution, fever prevalence, and respiratory findings (Table I).

Primary Outcome

Hospitalization within 7 days following ED discharge occurred in 156 of the 2,000 children with COVID-19 (7.8%), 198 of the 2,200 children with RSV (9.0%), and 175 of the 2,500 children with influenza (7.0%). Early post-discharge hospitalization rates differed across the viral etiologies despite similar discharge decisions at the index ED visit.

Clinical and Laboratory Characteristics

The clinical and laboratory features of children with and without hospitalization within 7 days are compared in Table II. In the COVID-19 cohort, those children who were hospitalized within 7 days more frequently had prolonged fever and elevated inflammatory markers at the index ED visit. CRP measurements were available in 68% of COVID-19, 52% of RSV, and 61% of influenza cases, reflecting routine physician-directed testing; higher CRP values were associated with an increased likelihood of early hospitalization.

In the RSV cohort, early hospitalization was predominantly associated with respiratory findings. Those children who were hospitalized within 7 days more commonly presented with respiratory distress and oxygen desaturation at the index ED visit, consistent with lower airway involvement following discharge.

By contrast, respiratory compromise was less commonly observed as a driver of early hospitalizations in influenza. Instead, persistent systemic symptoms, including prolonged fever, marked malaise, and myalgia, were more commonly observed among those children who were hospitalized shortly after discharge.

Multivariable Analyses

In multivariable logistic regression analyses adjusted for age, sex, and comorbidity status, distinct virus-specific predictors of early hospitalization were identified (Table III). All models additionally included fever duration, respiratory distress, SpO_2, feeding difficulties, and systemic symptoms as candidate variables. In the cases of COVID-19, CRP ≥33 mg/L [adjusted odds ratio (aOR) 2.8, 95% CI: 1.7-4.3] and persistent fever longer than 3 days (aOR 3.2, 95% CI: 1.9-4.9) were independently associated with hospitalization within 7 days. In the RSV cases, respiratory distress emerged as the strongest independent predictor (aOR 3.7, 95% CI: 2.2-5.9). In the influenza patients, the presence of persistent systemic symptoms was independently associated with early hospitalization (aOR 3.1, 95% CI: 1.8-4.7). Age-stratified analyses demonstrated that these virus-specific associations were consistent across those children younger than 2 years and those aged 2 years or older (data not shown). These associations remained significant after adjustment for age.

A summary of virus-specific predictors of early hospitalization is presented in Figure 1, illustrating the distinct clinical pathways leading to early post-discharge deterioration across viral etiologies.

Table I. Baseline characteristics of children by viral etiology

Characteristic	COVID-19 (n=2,000)	RSV (n=2,200)	Influenza (n=2,500)	p value
Age, months, median (IQR)	36 (18-72)	8 (3-18)	48 (24-96)	<0.001
Age <2 years, n (%)	580 (29.0)	1,540 (70.0)	375 (15.0)	<0.001
Male, n (%)	1,060 (53.0)	1,232 (56.0)	1,275 (51.0)	0.28
Comorbidity, n (%)	240 (12.0)	264 (12.0)	325 (13.0)	0.52
Fever at index visit, n (%)	1,600 (80.0)	1,320 (60.0)	2,125 (85.0)	<0.001
Respiratory distress, n (%)	200 (10.0)	660 (30.0)	250 (10.0)	<0.001
SpO_2 <94%, n (%)	80 (4.0)	286 (13.0)	100 (4.0)	<0.001
Feeding difficulty, n (%)	140 (7.0)	528 (24.0)	175 (7.0)	<0.001
CRP, mg/L, median (IQR)*	12 (5-28)	8 (3-18)	10 (4-22)	0.003

*CRP available in 68% of COVID-19, 52% of RSV, and 61% of influenza cases. Group comparisons were performed using the Kruskal-Wallis test for continuous variables and the chi-square test for categorical variables. Statistically significant overall p values ($p<0.05$) indicate differences across at least one pair of viral groups; post-hoc pairwise comparisons confirmed significant differences between all three groups for age, fever, respiratory distress, SpO_2, feeding difficulties, and CRP (all $p<0.05$ after Bonferroni correction)
COVID-19: Coronavirus disease-2019, IQR: Interquartile range, SpO_2: Peripheral oxygen saturation, CRP: C-reactive protein, RSV: Respiratory syncytial virus

Table II. Clinical characteristics of children with and without hospitalization within 7 days, by viral etiology

	COVID-19		RSV		Influenza	
	Hospitalized (n=156)	**No Hosp. (n=1,844)**	**Hospitalized (n=198)**	**No Hosp. (n=2,002)**	**Hospitalized (n=175)**	**No Hosp. (n=2,325)**
Fever >3 days, %	58.3	22.1	31.8	18.4	48.6	21.3
Resp. distress, %	18.6	9.1	52.5	27.4	14.3	9.4
SpO_2 <94%, %	8.3	3.6	25.3	11.6	7.4	3.7
CRP ≥33 mg/L, %*	42.3	14.8	18.2	12.1	22.9	13.5
Systemic sx, %	35.9	18.2	20.7	14.3	54.3	22.6
Feeding diff., %	14.7	6.3	38.9	22.1	12.0	6.5
Comorbidity, %	19.9	11.1	17.7	11.4	20.6	12.3

Hospitalized: hospitalized within 7 days. *Among patients with available CRP measurements. All within-etiology comparisons p<0.05
COVID-19: Coronavirus disease-2019, SpO_2: Peripheral oxygen saturation, CRP: C-reactive protein, RSV: Respiratory syncytial virus

Table III. Multivariable predictors of hospitalization within 7 days after emergency department discharge, by viral etiology

Viral etiology	Predictor	aOR (95% CI)	p value
COVID-19	CRP ≥33 mg/L	2.8 (1.7-4.3)	<0.001
	Persistent fever >3 days	3.2 (1.9-4.9)	<0.001
	Age <2 years	1.6 (1.1-2.4)	0.02
RSV	Respiratory distress	3.7 (2.2-5.9)	<0.001
	SpO_2 <94%	2.4 (1.5-3.8)	<0.001
	Age <2 years	2.1 (1.3-3.5)	0.004
Influenza	Persistent systemic symptoms	3.1 (1.8-4.7)	<0.001
	Prolonged fever >3 days	2.1 (1.3-3.4)	0.003
	Comorbidity	1.8 (1.1-2.9)	0.02

All models adjusted for age, sex, and comorbidity status
COVID-19: Coronavirus disease-2019, SpO_2: Peripheral oxygen saturation, CRP: C-reactive protein, RSV: Respiratory syncytial virus, aOR: Adjusted odds ratio, CI: Confidence interval

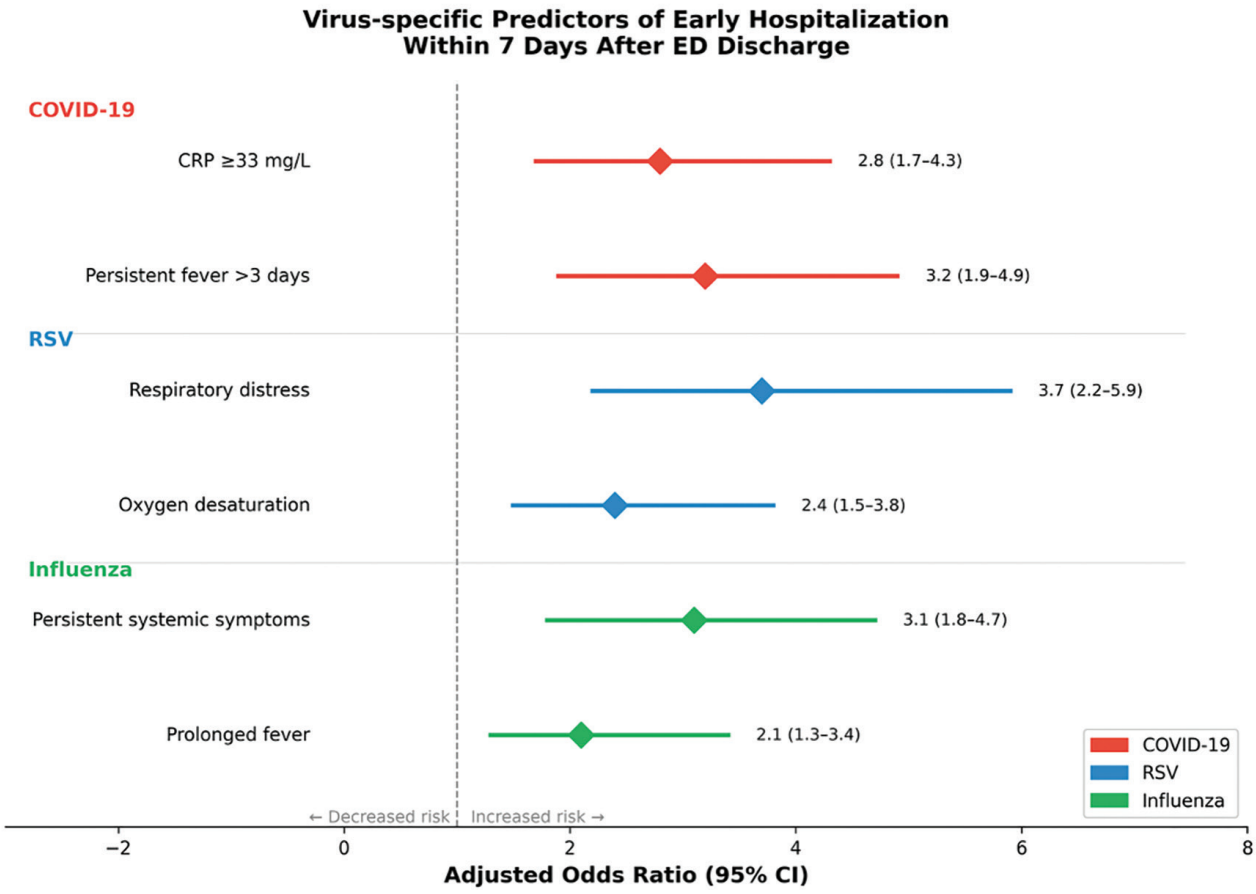

Figure 1. Virus-specific clinical predictors of early hospitalization within 7 days after emergency department discharge in children. Adjusted odds ratios with 95% confidence intervals from separate multivariable logistic regression models adjusted for age, sex, and comorbidity status
COVID-19: Coronavirus disease-2019, RSV: Respiratory syncytial virus, CI: Confidence interval, ED: Emergency department, CRP: C-reactive protein

Discussion

This retrospective cohort study demonstrates that early hospitalization following ED discharge reflects distinct, virus-specific clinical deterioration patterns in children with viral respiratory infections. Although the outcome (hospitalization within 7 days) is shared across etiologies, the clinical drivers of early deterioration differ substantially. In a large multi-state population-based analysis, critical ED revisits in children, defined as those resulting in ICU admission or death, were disproportionately associated with respiratory diagnoses at the index visit (7).

In pediatric COVID-19 cases, early hospitalization was primarily associated with systemic inflammatory features and prolonged fever rather than respiratory compromise, supporting prior observations that host inflammatory response plays a central role in early disease progression (10-12). Data from the RECOVER consortium have further demonstrated that illness severity is an important determinant of post-acute complications in children with SARS-CoV-2, reinforcing the relevance of inflammatory markers at early clinical encounters (13). While these studies did not directly evaluate post-discharge outcomes, the association between CRP elevation and clinical severity supports the biological plausibility of CRP as a possible indicator of early deterioration risk in febrile children being considered for ED discharge. It should be noted, however, that CRP was measured in only a subset of our patients (68% of COVID-19 cases) and was subject to indication bias; accordingly, the role of CRP as a predictive marker should be interpreted with caution and requires prospective validation before informing routine clinical practice. Sensitivity analyses restricted to patients with available CRP measurements demonstrated similar direction and magnitude of associations, partially mitigating but not fully excluding the influence of indication bias.

By contrast, RSV-related early deterioration was dominated by respiratory distress and hypoxemia,

consistent with classical bronchiolitis phenotypes described in contemporary cohorts (14,15). This finding underscores the well-established pattern of progressive lower airway obstruction which may not be fully apparent at the initial ED presentation, particularly in young infants. Although age was evaluated both as a continuous and as a dichotomized variable (<2 years) in our multivariable models, the clinical risk associated with age in RSV is not linear even within this youngest group; infants younger than 6 months represent a distinctly vulnerable subgroup in whom even mild respiratory distress at ED presentation may herald rapid deterioration, a pattern well-recognized in the bronchiolitis literature. Clinicians should exercise heightened caution when considering discharge in this youngest age group, even when initial oxygen saturation remains above threshold values.

Influenza exhibited an intermediate pattern, with persistent systemic symptoms rather than progressive respiratory failure driving early hospitalizations (16-18). This observation has practical implications for ED discharge counseling, as families may underestimate the significance of persistent constitutional symptoms in the absence of overt respiratory distress.

Our findings are particularly relevant in the context of the post-pandemic respiratory viral landscape. The 2022 simultaneous respiratory viral surge demonstrated that pediatric ED systems face substantial operational strain during concurrent epidemics, with increased visit volumes associated with higher revisit rates across diverse emergency settings (19). These observations collectively support the need for virus-specific, rather than uniform, approaches to ED discharge risk assessment in the current epidemiological era. The independent association between elevated CRP and early hospitalizations in our cohort is consistent with observations from pooled laboratory analyses demonstrating that elevated inflammatory markers are characteristic of more severe pediatric COVID-19 presentations (20,21). The identification of respiratory distress as the strongest predictor of early hospitalization in the RSV cohort reinforces current recommendations for close respiratory monitoring following discharge, especially during the first 48-72 hours (22,23). These findings support the notion that early post-discharge vigilance in children with influenza should prioritize the monitoring of systemic symptom persistence rather than focusing exclusively on respiratory parameters (24). Post-pandemic surveillance data have further demonstrated a substantial surge in RSV-related hospitalizations and respiratory support requirements, with older children and those with fewer comorbidities disproportionately affected during the 2022-2023 season (25,26). Recent surveillance data from the COVID-19-Associated Hospitalization Surveillance Network similarly identified chronic comorbidities and younger age as key determinants of severe disease among children hospitalized with COVID-19 during 2022-2024, underscoring the continuing clinical relevance of these risk factors in the post-pandemic era (27).

Importantly, the objective of this study was not to derive or validate a clinical risk score, but rather to identify clinically observable, etiology-specific warning patterns relevant to early post-discharge vigilance and ED disposition decisions. The distinct predictor profiles identified across the three viral etiologies suggest that a one-size-fits-all approach to discharge risk assessment may be suboptimal and that virus-specific clinical trajectories should be considered when counseling families at the time of ED discharge. This study was designed as an exploratory etiologic analysis rather than a derivation of clinical prediction rules, and the findings are intended to generate hypotheses for future prospective validation studies. The selected outcome, namely hospitalization within 7 days restricted to the index viral illness, reflects clinically meaningful early deterioration directly relevant to ED disposition quality, rather than incidental or unrelated healthcare utilization.

Clinical Implications

These findings have direct and practical implications for ED discharge planning. In pediatric COVID-19 cases, discharge counseling should emphasize close monitoring of fever trajectory, as prolonged fever emerged as a key marker of early post-discharge deterioration. CRP measurement, when obtained during routine ED evaluation, may provide additional risk stratification for febrile children being considered for discharge.

For RSV cases, respiratory monitoring remains paramount, particularly in young infants. Families should be specifically educated about early warning signs of progressive respiratory distress, including increased work in breathing, feeding difficulties, and changes in respiratory rate or effort, even when initial oxygen saturation appears acceptable at the time of discharge.

In children with influenza, symptoms such as severe myalgia, marked malaise, or sustained fever may warrant re-evaluation after discharge, even in the absence of overt respiratory deterioration.

These virus-specific clinical patterns may inform more tailored post-discharge follow-up strategies in routine pediatric emergency care.

Study Limitations

This study had limitations inherent to its retrospective design, including potential selection bias and residual confounding. Laboratory testing was performed at physician discretion rather than protocolized assessment. Virological data such as viral load or genomic characterization were unavailable. This study was limited to a single academic institution, which may restrict its generalizability to community-based pediatric emergency settings. Although the cohort was derived from tertiary-care EDs, the reliance on routinely collected clinical variables enhances the applicability of the findings to similar pediatric emergency care settings. The findings should be interpreted as hypothesis-generating and therefore require prospective validation. Furthermore, hospitalization outcomes may partially reflect physician decision-making variability inherent to real-world emergency care rather than clinical deterioration alone. The COVID-19 cohort spans the entire pandemic period (2020-2024), encompassing multiple SARS-CoV-2 variants with potentially different clinical profiles in children; variant-level subgroup analysis was not feasible due to the absence of routine genomic surveillance data, and temporal changes in clinical severity across variant waves should be considered when interpreting COVID-19-specific findings. Additionally, inter-campus variability in physician practices or admission thresholds was not formally assessed; although both campuses share unified institutional protocols, unmeasured differences cannot be entirely excluded and should be considered as a potential source of heterogeneity. Standardized discharge criteria or formal scoring systems were not used; discharge decisions were based on clinician assessment, which may introduce variability. With respect to outcome ascertainment, hospitalizations occurring at non-affiliated community hospitals within the 7-day follow-up period may not have been captured by the institutional EHR system. Although this institution serves as the primary regional referral center, a proportion of admissions to outside facilities could have been missed, and the hospitalization rates reported in this study should therefore be interpreted as a representative assessment rather than an exhaustive enumeration of all post-discharge events.

Conclusion

Early hospitalization within 7 days following ED discharge is associated with distinct clinical patterns across viral etiologies in those children with respiratory viral infections. COVID-19-related deterioration is driven by systemic inflammation and prolonged fever, RSV by progressive respiratory distress, and influenza by persistent systemic symptoms. Incorporating etiology-informed, readily available clinical features at ED presentation into discharge planning may inform outpatient management strategies and improve post-discharge vigilance.

Ethics

Ethics Committee Approval: This study was approved by the İstanbul Medipol University Non-Interventional Clinical Research Ethics Committee (approval no.: E-10840098-202.3.02-8228, date: 04.12.2025).

Informed Consent: Retrospective cohort study.

Footnotes

Authorship Contributions

Concept: F.B., Design: F.B., Data Collection or Processing: F.B., M.A., Analysis or Interpretation: F.B., M.A., Literature Search: F.B., Writing: F.B.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors received no financial support for the research, authorship, and/or publication of this article.

Artificial Intelligence Disclosure: AI-assisted writing tools were used during the preparation of this manuscript for the purposes of language editing and grammatical refinement. All clinical content, data analysis, interpretation of results, and scientific conclusions were performed exclusively by the authors. No AI tools were used for data collection, statistical analysis, or reference generation. The authors take full responsibility for the accuracy and integrity of all content presented.

Supplementary Figure Link: https://d2v96fxpocvxx.cloudfront.net/00df66f2-5bbf-4336-8ecd-25d1bc79ad25/content-images/69dc407f-5de5-429c-a68a-827be1576b5d.pdf

References

1. Shi T, McAllister DA, O'Brien KL, et al. Global, regional, and national disease burden estimates of acute lower respiratory

infections due to respiratory syncytial virus in young children in 2015: a systematic analysis. Lancet. 2017; 390:946-58.

2. Jain S, Williams DJ, Arnold SR, et al. Community-acquired pneumonia requiring hospitalization among U.S. children. N Engl J Med. 2015; 372:835-45.
3. Kompaniyets L, Agathis NT, Nelson JM, et al. Underlying medical conditions associated with severe COVID-19 illness among children. JAMA Netw Open. 2021; 4:e2111182.
4. Meissner HC. Viral bronchiolitis in children. N Engl J Med. 2016; 374:62-72.
5. Kondrich J, Rosenthal M. Influenza in children. Curr Opin Pediatr. 2017; 29:297-302.
6. Michelson KA, Lyons TW, Bachur RG, Monuteaux MC, Finkelstein JA. Timing and location of emergency department revisits. Pediatrics. 2018; 141:e20174087.
7. Cavallaro SC, Michelson KA, D'Ambrosi G, Monuteaux MC, Li J. Critical revisits among children after emergency department discharge. Ann Emerg Med. 2023; 82:575-82.
8. Goldman RD, Ong M, Griffith E. Predictors of return visit to the emergency department in pediatric patients. Pediatr Emerg Care. 2006; 22:739-44.
9. Hacking KA, Davis PG. Emergency department revisits in children: a systematic review. Pediatr Emerg Care. 2020; 36:e349-54.
10. Ludvigsson JF. Systematic review of COVID-19 in children shows milder cases and a better prognosis than adults. Acta Paediatr. 2020; 109:1088-95.
11. Hoang A, Chorath K, Moreira A, et al. COVID-19 in 7780 pediatric patients: a systematic review. EClinicalMedicine. 2020; 24:100433.
12. Götzinger F, Santiago-García B, Noguera-Julián A, et al. COVID-19 in children and adolescents in Europe: a multinational, multicentre cohort study. Lancet Child Adolesc Health. 2020; 4:653-61.
13. Rao S, Lee GM, Razzaghi H, et al. Clinical features and burden of postacute sequelae of SARS-CoV-2 infection in children and adolescents. JAMA Pediatr. 2022; 176:1000-9.
14. Florin TA, Plint AC, Zorc JJ. Viral bronchiolitis. Lancet. 2017; 389:211-24.
15. Hasegawa K, Tsugawa Y, Brown DF, et al. Trends in bronchiolitis hospitalizations in the United States, 2000–2009. Pediatrics. 2013; 132:28-36.
16. Dawood FS, Chaves SS, Pérez A, et al. Complications and associated bacterial coinfections among children hospitalized with seasonal or pandemic influenza, United States, 2003–2010. J Infect Dis. 2014; 209:686-94.
17. Lafond KE, Nair H, Rasooly MH, et al. Global role and burden of influenza in pediatric respiratory hospitalizations, 1982–2012: a systematic analysis. PLoS Med. 2016; 13:e1001977.
18. Poehling KA, Edwards KM, Griffin MR, et al. The burden of influenza in young children, 2004-2009. Pediatrics. 2013; 131:207-17.
19. Janke AT, Mangus CW, Fung CM, et al. Emergency department care for children during the 2022 viral respiratory illness surge. JAMA Netw Open. 2023; 6:e2346769.
20. Henry BM, Benoit SW, de Oliveira MHS, et al. Laboratory abnormalities in children with mild and severe coronavirus disease 2019 (COVID-19): a pooled analysis and review. Clin Biochem. 2020; 81:1-8.
21. Qiu H, Wu J, Hong L, et al. Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study. Lancet Infect Dis. 2020; 20:689-96.
22. Mansbach JM, Clark S, Christopher NC, et al. Prospective multicenter study of bronchiolitis: predicting safe discharges from the emergency department. Pediatrics. 2008; 121:680-8.
23. Mansbach JM, Piedra PA, Teach SJ, et al. Prospective multicenter study of viral etiology and hospital length of stay in children with severe bronchiolitis. Arch Pediatr Adolesc Med. 2012; 166:700-6.
24. Chaves SS, Perez A, Beauvais L, et al. The burden of influenza hospitalizations in infants from 2003 to 2012, United States. Pediatr Infect Dis J. 2014; 33:912-9.
25. Winthrop ZA, Perez JM, Staffa SJ, McManus ML, Duvall MG. Pediatric respiratory syncytial virus hospitalizations and respiratory support after the COVID-19 pandemic. JAMA Netw Open. 2024; 7:e2416852.
26. Fitzpatrick T, Buchan SA, Mahant S, et al. Pediatric respiratory syncytial virus hospitalizations, 2017–2023. JAMA Netw Open. 2024; 7:e2416077.
27. Free RJ, Patel K, Taylor CA, et al. Hospitalization for COVID-19 and risk factors for severe disease among children: 2022–2024. Pediatrics. 2025; 156:e2025072788.

J Pediatr Res 2026;13(3):272-8
DOI: 10.4274/jpr.galenos.2026.28445

Original Article

Central Venous Catheter-Related Complications in a Paediatric Neurosurgical Population: A Retrospective Study

Francisco Barros[1], Catarina Chaves[1], Patricia Santos[2], Helder Pereira[2], Joana Garcia Azevedo[1]

[1]Unidade Local de Saúde de São João, Department of Anaesthesiology, Porto, Portugal
[2]University of Porto Faculty of Medicine, Department of Surgery and Physiology, Porto, Portugal

ABSTRACT

Aim: Central venous catheters (CVCs) are frequently used in paediatric neurosurgical procedures but are also associated with potentially serious complications. Sometimes, these complications are life-threatening or require the premature removal of the CVC, which may impair medical care. The aim of this study was to evaluate the incidence and characteristics of CVC-related complications in a paediatric neurosurgical population and to identify factors associated with their occurrence.

Materials and Methods: We conducted a retrospective observational study including children who underwent neurosurgical procedures under general anaesthesia with intraoperative CVC placement at a single tertiary hospital between January 2017 and January 2022. Demographic data, clinical characteristics, catheter-related variables, and complications were collected from the electronic medical records. Complications were classified as either immediate or delayed. Associations between patient- and procedure-related factors and the occurrence of complications were analysed using appropriate statistical tests.

Results: A total of 193 paediatric patients were included in the final analysis. The median age of the patients was 81.8 months. Overall, CVC-related complications occurred in 24 patients (12.4%). Delayed complications accounted for the majority of events (91.7%), with device dysfunction being the most frequent complication. Immediate complications were rare (8.3%). An American Society of Anesthesiologists Physical Status (ASA-PS) classification of 3 was significantly associated with a higher incidence of complications (p=0.038). No significant associations were found between complications and patient age, sex, catheter location, ultrasound guidance, or intraoperative blood product transfusion.

Conclusion: In this paediatric neurosurgical cohort, CVC-related complications occurred in 12.4% of patients and were predominantly delayed, with device dysfunction being the most common event. Higher ASA-PS was the only factor associated with increased complication risks. These findings highlight the importance of careful postoperative catheter management and the necessity of ongoing CVC assessment in this vulnerable population.

Keywords: ASA Physical Status, central venous catheter, neuroanaesthesiology, paediatric neuro-surgery, ultrasound

Corresponding Author
Catarina Chaves, MD, Unidade Local de Saúde de São João, Department of Anaesthesiology, Porto, Portugal
E-mail: catarinacchaves@gmail.com **ORCID:** orcid.org/0000-0002-6472-5196
Received: 26.11.2025 **Accepted:** 18.02.2026 **Epub:** 21.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Barros F, Chaves C, Santos P, Pereira H, Azevedo JG. Central venous catheter-related complications in a paediatric neurosurgical population: a retrospective study. J Pediatr Res. 2026;13(3):272-8

Introduction

The insertion of a central venous catheter (CVC) is frequently necessary in surgical settings, particularly in paediatric neurosurgical procedures. Approximately 25% of hospitalized children receive a central venous access device (CVAD) for medical reasons (1), as it provides durable and atraumatic venous access for blood sampling and for the safe administration of cytotoxic drugs, intravenous drugs, blood components, and parenteral nutrition (2). Approximately five million CVCs are placed annually in the United States, in the paediatric population (3). Nevertheless, the placement of a CVC is not without risk and CVAD failure prior to treatment completion remains high (20-25%) due to mechanical, infectious, and vascular complications (4,5).

Children with CVADs are already vulnerable to complications and disability due to their underlying health conditions. This vulnerability is worsened by the risk of adverse events associated with the insertion and management of CVADs (6,7). The complications associated with indwelling CVCs may be immediate or delayed (8). Immediate complications are related to the technique at the time of the insertion procedure and include vascular injuries (injury, bleeding, and/or hematoma), pulmonary complications (pneumothorax, pneumomediastinum, chylothorax, tracheal injury, recurrent laryngeal nerve injury and/or air embolism) and cardiac complications (arrhythmia and/or cardiac arrest) (8). Ultrasound has significantly reduced the incidence of immediate complications from rates previously as high as 11.8% to between 4 and 7% (9,10). Delayed complications include device dysfunction and infection (8). Device dysfunction may be caused by fibrin sheath formation, fracture, thrombosis, central venous stenosis, or infection (8). Subclavian CVCs have been reported to have the lowest rate of thrombosis (11,12), whereas femoral lines have the highest rate of thrombosis (11). Infection is a serious delayed complication associated with central venous access which can lead to sepsis, shock, and even death (8). Catheter infection sources include contamination from skin flora, contamination from infused substance, or from hematogenous spread from an unrelated site (13). In some circumstances, these complications are life-threatening or require the premature removal of the CVC, which may result in the inability to receive prescribed medications, fluids, or nutrition (14-17).

In our centre, the placement of CVCs in the paediatric neurosurgical population is relatively common and only performed by anaesthesiologists. The aim of this study was to evaluate CVC-related complications in a paediatric neurosurgical population of a single tertiary hospital centre, between 2017 and 2022, and to identify any possible risk factors related to these complications.

Materials and Methods

Ethical Considerations

Ethical approval for this study was obtained from the Ethics Committee of Centro Hospitalar Universitário de São João (CHUSJ), Porto, Portugal (approval no.: 181/2022, date: 07.09.2022). This study was conducted in accordance with the Declaration of Helsinki and all collected data were saved anonymously.

Study Design

Our initial population consisted of all children who had undergone neurosurgical procedures under general anaesthesia in our hospital's Neurosurgical Department, between January 2017 and January 2022. A total of 202 children with a record of a CVC placed in the operating room were identified.

The only exclusion criteria were incomplete records concerning the CVC (date, duration, site of placement or complications). Nine children were excluded due to missing data concerning their catheter location.

Obtaining Data and Definitions

Data were obtained by assessing and consulting the electronic medical records of all of the children in our sample. The collected data included sex, age, weight, American Society of Anesthesiologists Physical Status (ASA-PS), neurosurgical diagnosis, neurosurgical procedure, thrombotic risk (Caprini score), the presence of a previous CVC, CVC length and calibre, CVC location, the total number of insertion attempts, ultrasonographic technique usage, administration of tranexamic acid, hypertonic solutions or blood products and CVC-related complications.

Data regarding coagulation status, including platelet count and the presence of coagulopathy, were collected retrospectively when available. However, documentation was incomplete for a substantial proportion of the patients, precluding robust statistical analysis of these variables. Similarly, information regarding catheter length, calibre, and dwell time was inconsistently recorded, particularly in the earlier years of the study period.

Statistical Analysis

A descriptive analysis of the variables was performed in order to summarize the data. Dichotomous categorical variables are presented as relative or absolute frequencies.

Complications were evaluated as a dichotomous variable (the presence or absence of complications). In order to compare categorical variables, the chi-square or Fisher's exact test were performed. Differences were considered statistically significant when $p<0.05$. Statistical analysis was performed using the Statistical Package for Social Sciences version 27.0.

Results

Demographics Characteristics and CVC Locations

From a total of 202 paediatric patients enrolled, 115 were male (56.9%) and 87 were female (43.1%). Nine patients (4.4%) had missing data concerning their CVC location (and so were excluded from posterior statistical analysis), resulting in a sample of 193 children.

Table I shows the study's final population characteristics. The patients' ages ranged between 27 days and 17 years old, with a median age of 81.8 months (standard deviation ± 68.4). Of the 193 children, 62 were younger than one year, 21 were between one and three years old (11%), 49 were between three and 10 years old (25%) and 61 were older than 10 years old (32%). Most children were classified as ASA-PS 2 (n=66; 34.2%) and had oncologic neurosurgical diagnoses (n=112; 58%). Most catheters were inserted into the internal jugular vein (n=82; 42.5%).

Table I. Final population patient characteristics (n=193)

Variable	Total
Sex	
Male (%)	111 (57.5)
Female (%)	82 (42.5)
Mean age (months)	81.8±68.4
ASA Physical Status	
I (%)	51 (26.4)
II (%)	66 (34.2)
III (%)	57 (29.5)
IV (%)	14 (7.3)
V (%)	5 (2.6)
Neurosurgical diagnosis	
Oncological (%)	112 (58)
Non-oncological (%)	81 (42)
CVC location	
Internal jugular vein (%)	82 (42.5)
Femoral vein (%)	70 (36.3)
Subclavian vein (%)	41 (21.2)
CVC complication	
Yes (%)	24 (12.4)
No (%)	169 (87.6)
Delayed (%)	22 (91.7)
Immediate (%)	2 (8.3)
Device disfunction (%)	13 (54.1)
Infection/Inflammatory (%)	8 (33.3)
Thrombosis (%)	1 (4.2)
Haemorrhagic (%)	1 (4.2)
Mixed (T + I) (%)	1 (4.2)

Data presented as n (%) or mean ± standard deviation
ASA: American Society of Anaesthesiology, CVC: Central venous catheter, I: Infection, T: Thrombosis

Complications

As shown in Table I, in total, 24 complications were observed, accounting for 12.4% of all patients. Delayed complications (n=22; 91.7%) were far more frequent than immediate complications (n=2; 8.3%) and multi-cause device disfunction accounted for more than half of the complications (n=13; 54.1%). No more than a single complication episode was observed in any patient.

Factors Associated with Complications

Table II summarizes the relationship between the population characteristics and the occurrence of complications. The observed complications were identical in male and female patients (12 each), with no association between sex and the presence of complications (p=0.565). Most complications occurred in those patients aged between more than three and 10 years old. No association was found between age and the presence of complications (p=0.375).

An ASA-PS classification of 3 was significantly associated with a higher incidence of complications (p=0.038).

Table II also demonstrates that 23.5% (n=4) of those children with a vascular disease diagnosis had a complication, followed by 12.5% (n=6) of children with a congenital malformation diagnosis, 12.5% (n=1) of children with a central nervous system (CNS) infection, and 11.6% (n=13) of those with a CNS tumour diagnosis. No association was found between the neurosurgical diagnosis and complications (p=0.537).

Internal jugular vein CVCs were associated with the most complications (13 of a total of 24), followed by femoral and then subclavian veins. There was no statistically significant difference between CVC location and complications (p=0.381).

Ultrasound was used to aid CVC placement in 25 patients (13%). There was no statistically significant association between ultrasonographic technique usage and the presence of complications (p=0.799).

A total of 46 patients (23.8%) needed intraoperative blood product transfusions. We did not find any association between intraoperative blood product transfusions and the presence of CVC-associated complications (p=0.362).

Table II. Relationship between population characteristics and occurrence of complications (n=193)

Variable	Complications		Chi-square test		
	Yes	No	χ^2 (df)	Effect size (Cramér's V)	p value
Sex					
Male (%)	12 (10.8)	99 (89.2)	0.331 (1)	0.041	0.565
Female (%)	12 (14.6)	70 (85.4)			
Age					
<1 year old (%)	8 (12.9)	54 (87.1)			
1-3 years old (%)	1 (4.8)	20 (95.2)	3.11 (3)	0.127	0.375
3-10 years old (%)	9 (18.4)	40 (81.6)			
>10 years old (%)	6 (9.8)	55 (90.2)			
ASA Physical Status					
I (%)	2 (3.9)	49 (96.1)			
II (%)	6 (9.1)	60 (90.9)			
III (%)	11 (19.3)	46 (80.7)	10.15 (4)	0.23	0.038
IV (%)	4 (28.6)	10 (71.4)			
V (%)	1 (20)	4 (80)			
Neurosurgical diagnosis					
Epilepsy (%)	0 (0)	8 (100)			
CNS infection (%)	1 (12.5)	7 (87.5)			
Congenital malformation (%)	6 (12.5)	42 (87.5)	3.13 (4)	0.127	0.537
CNS tumour (%)	13 (11.6)	99 (88.4)			
Vascular (%)	4 (23.5)	13 (76.5)			
CVC location					
Femoral vein (%)	8 (11.4)	62 (88.6)			
Internal jugular vein (%)	13 (15.9)	39 (84.1)	1.931 (2)	0.10	0.381
Subclavian vein (%)	3 (7.3)	38 (92.7)			
Ultrasonographic technique					
Yes (%)	4 (16)	21 (84)	0.065 (1)	0.018	0.799
No (%)	20 (11.9)	148 (88.1)			
Transfusion					
Yes (%)	8 (17.4)	38 (82.6)	0.830 (1)	0.066	0.362
No (%)	16 (10.9)	131 (89.1)			

Data presented as n (% of given variable)
ASA: American Society of Anaesthesiology, CNS: Central nervous system, CVC: Central venous catheter, df: Degrees of freedom, χ^2: Chi-square value

Age

Table III summarizes the relationship between our defined age intervals and neurosurgical diagnosis, CVC location and ultrasonographic technique usage.

Among those patients of less than one year old, most had a malformative disease diagnosis (n=43; 69.4%). Oncologic diagnosis was the most common neurosurgical diagnosis in the other age groups. We found a statistically significant association between age and neurosurgical diagnosis (p<0.001).

In those patients of less than one year old, most CVCs (n=36; 58.1%) were placed in the femoral vein, as was also the case in patients aged between one and three years old (n=10; 47.6%). The jugular vein was the most common location in children aged between more than three years old and 10 years old (n=28; 57.1%). In children older than 10 years old, the most common CVC location was the subclavian vein (n=26; 42.6%). We found a statistically significant association between age and CVC location (p<0.001).

Ultrasound utilization for CVC placement was higher among those patients younger than one year old (n=14; 22.6%). There was a statistically significant association between age and ultrasonographic technique usage (p=0.030).

Table III. Relationship between neurosurgical diagnosis, CVC location and ultrasonographic technique usage, and age (n=193)

Variable	Age				Chi-square test		
	<1 yo	1-3 yo	>3-10 yo	>10 yo	χ^2 (df)	Effect size (Cramér's V)	p value
Neurosurgical diagnosis							
Epilepsy (%)	1 (1.6)	1 (4.8)	1 (2.0)	5 (8.2)			
CNS Infection (%)	3 (4.8)	0 (0.0)	2 (4.1)	3 (4.9)			
Congenital malformation (%)	43 (69.4)	2 (9.5)	3 (6.1)	0 (0.0)	108.28 (12)	0.432	<0.001
CNS tumour (%)	14 (22.6)	18 (85.7)	36 (73.5)	44 (72.1)			
Vascular (%)	1 (1.6)	0 (0.0)	7 (14.3)	9 (14.8)			
CVC location							
Femoral vein (%)	36 (58.1)	10 (47.6)	12 (24.5)	12 (19.7)			
Internal jugular vein (%)	24 (38.7)	7 (33.3)	28 (57.1)	23 (37.7)	41.64 (6)	0.328	<0.001
Subclavian vein (%)	2 (3.2)	4 (19.0)	9 (18.4)	26 (42.6)			
Ultrasonographic technique							
Yes (%)	14 (22.6)	3 (14.3)	5 (10.2)	3 (4.9)	8.95 (3)	0.215	0.030
No (%)	48 (77.4)	18 (85.7)	44 (89.8)	58 (95.1)			

Data presented as n (% of total in each age interval)
df: Degrees of freedom, yo: Years old, χ^2: Chi-square value, CVC: Central venous catheter, CNS: Central nervous system

Discussion

This retrospective observational study involving 193 paediatric patients provided a focused evaluation of CVC complications in a paediatric neurosurgical cohort, demonstrating a complication rate of 12.4%, which is lower than in many of the previously published paediatric populations where rates often exceeded 40% (18-20). These studies reported haematological diseases, Hickman-Broviac catheters, low neutrophil counts and haematological diseases as risk factors for CVC complications. In our study, none of the patients had these characteristics, which may explain our lower rate of complications. The predominance of delayed complications (91.7% of all complications), particularly catheter dysfunction, highlights the unique characteristics of neurosurgical patients and the perioperative pathways in which CVCs are used.

There was no statistically significant association between ultrasonographic technique usage and the presence of complications. However, this finding should be interpreted with caution. Ultrasound was used in only 13% of patients, substantially limiting the statistical power of this study to detect meaningful differences and increasing the likelihood of a type II error. In addition, immediate complications were rare across the entire cohort, regardless of whether ultrasound was used or not, which may further reduce the observable effect of ultrasound on complication rates. Given the strong and consistent evidence supporting ultrasound-guided CVC placement in paediatric patients (21-25), including higher success rates and fewer mechanical complications, our results should not be interpreted as contradicting the current recommendations. Rather, they likely reflect the limited sample size of ultrasound-guided insertions and the low baseline rate of immediate complications in this neurosurgical population.

We did not identify a significant association between the location of the CVC and complications. It has been reported that internal jugular lines have the lowest complication rates (26) and femoral lines have the highest rate of thrombosis (11). Conflicting evidence exists regarding infection risks in femoral lines (27,28). Nevertheless, one observational study in paediatric surgical patients showed an association between femoral lines and bloodstream infection (29). Although CVC location was not statistically associated with complications in our cohort, age-related anatomical and procedural considerations should continue to guide site selection

Interpretation in the Context of Our Cohort

Children undergoing neurosurgical procedures represent a specific clinical group with distinct vascular access needs. In our population: most CVCs were employed for intraoperative management, where rapid fluid administration, vasoactive infusions, and reliable venous access are required; immediate complications were rare, consistent with modern ultrasound-guided techniques and experienced anaesthesia teams (21-23); and delayed complications dominated, with device dysfunction accounting for more than half of all events. This pattern aligns with the shorter perioperative dwell times in neurosurgery, where CVCs often remain in place only briefly postoperatively. Femoral access is more common in younger children, reflecting anatomical constraints and the clinical realities of urgent neurosurgical care. ASA-PS 3 was

the only factor significantly associated with complications, suggesting that physiological vulnerability rather than procedural factors played a larger role in catheter failure. To date, there is a lack of published studies on this association, but it is logical to assume that more seriously ill patients are more susceptible to CVC-associated complications.

These findings reinforce the notion that the neurosurgical perioperative environment is distinct from oncology or critical care settings, where prolonged device dwell times contribute more heavily to infectious and thrombotic complications.

Given that all vascular access devices in this cohort were temporary CVCs placed intraoperatively, our findings specifically reflect perioperative neurosurgical practice and should not be extrapolated to other device types such as peripherally inserted central catheters (PICCs) or midlines. The absence of these devices in our cohort underscores the distinct vascular access requirements of paediatric neurosurgical patients.

Clinical Implications

While CVCs are essential during neurosurgical procedures, the high proportion of delayed complications in our cohort underscores the importance of early postoperative evaluation and the consideration of alternative vascular access devices when clinically appropriate. Integrating structured, individualized vascular access planning aligned with updated paediatric guidance (30) may help to reduce complication rates and improve overall safety in this vulnerable population. Additionally, our findings underscore the need to optimise perioperative vascular access strategies in this population. As delayed complications, particularly device dysfunction, were the most frequent events, clinical teams should routinely reassess the ongoing need for a CVC after surgery and consider an early transition to alternative devices when appropriate. PICCs or midlines may provide safer medium-term options for selected patients, especially those with oncological diagnoses or anticipated prolonged postoperative therapy. Shorter CVC dwell times, stricter criteria for initial CVC placement, and routine consideration of catheter removal on postoperative day one, unless central access remains clearly required, may further reduce complication risks. These approaches are consistent with contemporary paediatric vascular access guidance and promote a more individualised, safety-focused strategy for device selection and maintenance.

Study Limitations

The retrospective design and relatively limited sample size of this single-centre study must be considered when interpreting the results. Given that all vascular access devices in this cohort were temporary CVCs placed intraoperatively, our findings specifically reflect perioperative neurosurgical practice and should not be extrapolated to other device types such as PICCs or midlines. The absence of these devices in our cohort underscores the distinct vascular access requirements of paediatric neurosurgical patients. Missing data, particularly regarding CVC dwell times, catheter lengths, calibres, the number of insertion attempts and coagulation status (e.g. platelet count and the presence of coagulopathy), reduced the statistical power and prevented the inclusion of potentially relevant variables in this study. Nine patients were excluded due to missing CVC location data, and incomplete documentation was likely affected by the transition to a new anaesthesia recording system during the study period. Second, this study's retrospective design limited control over confounding factors, and the six-year timeframe may reflect practices which have since evolved, particularly the increased adoption of ultrasound guidance and the decreasing use of femoral access in younger children. Third, this study was not powerful enough to detect small effect sizes, and the relatively low incidence of complications, especially immediate complications, may have limited our ability to identify subtler associations. Finally, we did not differentiate between complication types (immediate vs delayed) in the statistical analysis, which may partially explain the absence of an association between ultrasound guidance and complications.

Conclusion

In this paediatric neurosurgical population, CVC complications occurred in 12.4% of patients, with delayed device dysfunction being the most common. ASA-PS 3 was the only factor significantly associated with increased risk, despite strong existing literature associating CVC location and ultrasonographic technique usage to the risk of complications. These findings support the ongoing efforts to optimize catheter maintenance strategies and the adoption of updated paediatric vascular access guidance. We acknowledge that our study may have some limitations and prospective studies are needed in order to further refine risk prediction and prevention strategies.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the Ethics Committee of Centro Hospitalar Universitário de São João (CHUSJ), Porto, Portugal (approval no.: 181/2022, date: 07.09.2022).

Informed Consent: Retrospective observational study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: F.B., P.S., H.P., J.G.A., Concept: F.B., P.S., H.P., J.G.A., Design: P.S., H.P., J.G.A., Data Collection or Processing: F.B., C.C., P.S., J.G.A., Analysis or Interpretation: F.B., C.C., H.P., J.G.A., Literature Search: F.B., C.C., J.G.A., Writing: F.B., C.C., P.S.

Conflict of Interest: The authors declare no conflict of interest.

Financial Disclosure: The authors declare that no financial or material support was received for this study.

References

1. Ullman AJ, Cooke M, Kleidon T, Rickard CM. Road map for improvement: Point prevalence audit and survey of central venous access devices in paediatric acute care. J Paediatr Child Health. 2017; 53:123-30.
2. Raaf JH. Administration of chemotherapeutic agents. Supportive Care in Cancer. 1994; 2:335-46.
3. Duesing LA, Fawley JA, Wagner AJ. Central venous access in the pediatric population with emphasis on complications and prevention strategies. Nutr Clin Pract. 2016; 31:490-501.
4. Kleidon TM, Rickard CM, Schults JA, et al. Development of a paediatric central venous access device database: a retrospective cohort study of practice evolution and risk factors for device failure. J Paediatr Child Health. 2020; 56:289-97.
5. Ullman AJ, Marsh N, Mihala G, Cooke M, Rickard CM. Complications of central venous access devices: a systematic review. Pediatrics. 2015; 136:e1331-44.
6. Cesaro S, Corrò R, Pelosin A, et al. A prospective survey on incidence and outcome of Broviac/Hickman catheter-related complications in pediatric patients affected by hematological and oncological diseases. Ann Hematol. 2004; 83:183-8.
7. Perdikaris P, Petsios K, Vasilatou-Kosmidis H, Matziou V. Complications of Hickman-Broviac catheters in children with malignancies. Pediatr Hematol Oncol. 2008; 25:375-84.
8. Kornbau C, Lee KC, Hughes GD, Firstenberg MS. Central line complications. Int J Crit Illn Inj Sci. 2015; 5:170-8.
9. Bhutta ST, Culp WC. Evaluation and management of central venous access complications. Tech Vasc Interv Radiol. 2011; 14:217-24.
10. Sznajder JI, Zveibil FR, Bitterman H, Weiner P, Bursztein S. Central vein catheterization. Failure and complication rates by three percutaneous approaches. Arch Intern Med. 1986; 146:259-61.
11. Kusminsky RE. Complications of central venous catheterization. J Am Coll Surg. 2007; 204:681-96.
12. McGee DC, Gould MK. Preventing complications of central venous catheterization. N Engl J Med. 2003; 348:1123-33.
13. Early TF, Gregory RT, Wheeler JR, Snyder SO, Jr., Gayle RG. Increased infection rate in double-lumen versus single-lumen Hickman catheters in cancer patients. South Med J. 1990; 83:34-6.
14. Peng C, Monagle P, Newall F. Clinical outcomes of management of CVAD occlusions. Arch Dis Child. 2011; 96:885-7.
15. van Miert C, Hill R, Jones L. Interventions for restoring patency of occluded central venous catheter lumens. Cochrane Database Syst Rev. 2012; 2012:Cd007119.
16. Groeger JS, Lucas AB, Thaler HT, et al. Infectious morbidity associated with long-term use of venous access devices in patients with cancer. Ann Intern Med. 1993; 119:1168-74.
17. Raad I. Management of intravascular catheter-related infections. J Antimicrob Chemother. 2000; 45:267-70.
18. Athale UH, Siciliano S, Cheng J, Thabane L, Chan AK. Central venous line dysfunction is an independent predictor of poor survival in children with cancer. J Pediatr Hematol Oncol. 2012; 34:188-93.
19. Sezgin Evim M, Yörük G, Güler S, et al. The evaluation of central venous catheter-related complications in pediatric acute leukemia patients: single center experience. J Pediatr Hematol Oncol. 2023; 45:e92-6.
20. Fratino G, Molinari AC, Parodi S, et al. Central venous catheter-related complications in children with oncological/hematological diseases: an observational study of 418 devices. Annals of Oncology. 2005; 16:648-54.
21. Kehagias E, Galanakis N, Tsetis D. Central venous catheters: which, when and how. Br J Radiol. 2023; 96:20220894.
22. Vafek V, Skřívošvská T, Kosinová M, et al. Central venous catheter cannulation in pediatric anesthesia and intensive care: a prospective observational trial. Children (Basel). 2022; 9.
23. Leibowitz A, Oren-Grinberg A, Matyal R. Ultrasound guidance for central venous access: current evidence and clinical recommendations. J Intensive Care Med. 2020; 35:303-21.
24. Lau CS, Chamberlain RS. Ultrasound-guided central venous catheter placement increases success rates in pediatric patients: a meta-analysis. Pediatr Res. 2016; 80:178-84.
25. He C, Vieira R, Marin JR. Utility of ultrasound guidance for central venous access in children. Pediatr Emerg Care. 2017; 33:359-62.
26. Saugel B, Scheeren TWL, Teboul JL. Ultrasound-guided central venous catheter placement: a structured review and recommendations for clinical practice. Crit Care. 2017; 21:225.
27. Rupp SM, Apfelbaum JL, Blitt C, et al. Practice guidelines for central venous access: a report by the American Society of Anesthesiologists task force on central venous access. Anesthesiology. 2012; 116:539-73.
28. Arvaniti K, Lathyris D, Blot S, Apostolidou-Kiouti F, Koulenti D, Haidich AB. Cumulative evidence of randomized controlled and observational studies on catheter-related infection risk of central venous catheter insertion site in ICU patients: a pairwise and network meta-analysis. Crit Care Med. 2017; 45:e437-48.
29. Haldar R, Mandelia A, Mishra P, Mishra A, Siddiqui Y. Central venous catheter-related infectious complications in pediatric surgical patients: a single-center experience. J Pediatr Intensive Care. 2022; 11:240-6.
30. National Infusion and Vascular Access Society (NIVAS). Guidance for vascular access device flushing and locking in paediatric patients. Version 3. London: National Infusion and Vascular Access Society; 2025.

J Pediatr Res 2026;13(3):279-84
DOI: 10.4274/jpr.galenos.2026.58708

Original Article

Etiology-Based Clinical Outcomes after Pediatric Tracheostomy in Intensive Care: A Retrospective Observational Study

Fatih Durak[1], Emine Pınar Küllüoğlu[1], Gökçen Özçifçi[2], Deniz Koçak Ter[2]

[1]İzmir Katip Çelebi University Faculty of Medicine, Department of Pediatric Intensive Care, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir City Hospital, Department of Pediatric Intensive Care, İzmir, Türkiye

ABSTRACT

Aim: To evaluate etiology-based differences in clinical outcomes among pediatric patients undergoing tracheostomy in a tertiary pediatric intensive care unit.

Materials and Methods: In this retrospective observational study, the medical records of patients aged 1 month to 18 years who had undergone tracheostomy between October 2023 and December 2025 were reviewed. The patients were classified into three etiological groups: neuromuscular diseases, chronic respiratory failure, and upper airway obstruction/malformation. Their demographic characteristics, mechanical ventilation durations, intensive care and hospital length of stay, complications, and clinical outcomes were analyzed.

Results: A total of 59 patients were included. The most common tracheostomy indications were neuromuscular diseases (52.5%) and chronic respiratory failure (33.9%). Significant differences were observed among the groups regarding their age, anthropometric measurements, and mechanical ventilation durations ($p<0.05$). Those patients with chronic respiratory failure had longer pre- and post-tracheostomy mechanical ventilation durations and the lowest successful ventilator weaning rate. Early and late complication rates were 10.2% and 25.4%, respectively. The decannulation rate was 5.1%, and overall mortality was 22%.

Conclusion: Tracheostomy indication was significantly associated with clinical outcomes in pediatric intensive care patients. Those patients with chronic respiratory failure demonstrated prolonged mechanical ventilation and lower ventilator weaning success. Etiology-based individualized management strategies warrant further investigation in this population.

Keywords: Tracheostomy, intensive care units, pediatric, respiration, artificial, respiratory insufficiency, neuromuscular diseases

Introduction

Tracheostomy is a commonly utilized surgical intervention in pediatric intensive care units (PICUs) for those children who require prolonged respiratory or airway support. The procedure has several potential benefits, including improved comfort, decreased sedative and analgesic requirements, the facilitation of ventilator weaning, and a reduction in intensive care burden in selected patients (1). Previous studies have reported that nearly 1-3% of children admitted to PICUs undergo tracheostomy during their clinical course (2).

Corresponding Author
Asst. Prof. Emine Pınar Küllüoğlu, İzmir Katip Çelebi University Faculty of Medicine, Department of Pediatric Intensive Care, İzmir, Türkiye
E-mail: kulluoglupinar@gmail.com **ORCID:** orcid.org/0000-0003-1011-6106
Received: 15.05.2026 **Accepted:** 17.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Durak F, Küllüoğlu EP, Özçifçi G, Koçak Ter D. Etiology-based clinical outcomes after pediatric tracheostomy in intensive care: a retrospective observational study. J Pediatr Res. 2026;13(3):279-84

Despite the increasing use of pediatric tracheostomy, standardized recommendations regarding optimal timing, indications, and long-term outcomes remain limited (3-5). The timing of tracheostomy in pediatric populations varies substantially according to the underlying disease, expected duration of ventilatory support, and institutional preferences, generally ranging between several weeks and months after the initiation of mechanical ventilation (2). In addition to prolonged mechanical ventilation, common indications include upper airway obstruction, impaired airway protection, swallowing dysfunction, ineffective cough reflex, and inadequate secretion clearance (6).

Several reports have suggested that tracheostomy may contribute to reduced ventilator-associated complications, shorter mechanical ventilation duration, and improved long-term respiratory management (7). Nevertheless, procedure-related complications such as bleeding, infection, accidental decannulation, and tracheal stenosis may still occur despite advances in intensive care practice (8).

Recent updates have further refined our understanding of contemporary indications, timing considerations, and long-term care pathways for pediatric tracheostomy, emphasizing prolonged mechanical ventilation as the predominant modern indication and highlighting the evolving practices in decannulation planning and multidisciplinary follow-up (9,10). Population-level data further indicate that pediatric tracheostomy rates and associated health care resource usage continue to evolve over time, underscoring the ongoing need for etiology-specific outcome data in order to guide clinical decision-making (11,12).

Although pediatric tracheostomy is widely performed, studies specifically evaluating the impact of different etiological indications on clinical outcomes remain relatively limited. Therefore, this study aimed to evaluate tracheostomy indications, complications, clinical course, and etiology-based outcome differences among pediatric patients managed in a tertiary PICU.

Materials and Methods

Study Design and Population

A retrospective single-center observational study was carried out in a tertiary PICU in accordance with the ethical principles of the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board of University of Health Sciences Türkiye, İzmir City Hospital (approval no.: 2026/111; date: 18.02.2026). This study was conducted in a 54-bed tertiary PICU, which admits approximately 1,200 patients annually. This study was reported in accordance with the *Strengthening the Reporting of Observational Studies in Epidemiology*-STROBE statement for observational cohort studies.

The medical records of pediatric patients who had undergone tracheostomy between October, 2023 and December, 2025 were retrospectively reviewed following ethics approval. Follow-up data for decannulation and mortality outcomes were collected up until December 31st, 2025.

Inclusion and Exclusion Criteria

Children aged between 1 month and 18 years who had undergone tracheostomy and remained in the PICU for at least 24 hours were included in this analysis. Those patients with insufficient clinical data or PICU stays shorter than 24 hours were excluded.

The patients were grouped according to their underlying etiology into three groups: neuromuscular diseases (NMs), chronic respiratory failure (CRF) secondary to pulmonary disease, and upper airway obstruction/malformation (UAO/M). Tracheostomy was considered in those patients requiring prolonged invasive mechanical ventilation, persistent upper airway obstruction, impaired airway protection, or chronic respiratory insufficiency, following a multidisciplinary evaluation involving pediatric intensivists, pulmonologists, and otolaryngologists as appropriate. The patients were classified according to their primary/predominant clinical indication for tracheostomy, as determined by the treating multidisciplinary team; those patients with NM were classified in the NMs group when NM was considered the primary driver of respiratory failure, even if chronic respiratory insufficiency was also present as a consequence of the underlying neuromuscular disorder, whereas other patients were classified in the CRF group only when a chronic pulmonary process independent of NM was the primary/dominant indication. "Chronic lung disease" as a baseline comorbidity is conceptually distinct from the CRF etiological group and could be present in patients classified into any of the three groups.

Data Collection

The data retrieved from electronic hospital records included demographic information, tracheostomy indications, duration of mechanical ventilation, duration of hospitalization, complications, and outcome measures. Early complications were defined as those occurring within the first 7 days after tracheostomy, and late complications were defined as those occurring after day 7.

Statistical Analysis

All statistical analyses were performed using SPSS software version 23.0 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as median and interquartile range, while categorical variables are expressed as numbers and percentages. Distribution normality was assessed using the Shapiro-Wilk test.

Comparisons between groups were conducted using the Mann-Whitney U or Kruskal-Wallis tests for continuous variables. Pairwise comparisons following a significant Kruskal-Wallis test were performed using Dunn's test with Bonferroni correction. Categorical variables were analyzed using chi-square or Fisher's exact tests where appropriate; given the small expected cell count in the UAO/M subgroup, the Fisher-Freeman-Halton exact test was additionally applied to the 3×2 comparison of ventilator weaning success by etiological group. Correlations between continuous variables were evaluated using Spearman correlation analysis. Statistical significance was accepted at $p<0.05$.

Results

A total of 59 pediatric patients were included in the study cohort. NMs and CRF represented the most frequent indications for tracheostomy. Demographic characteristics and accompanying comorbidities are summarized in Table I.

A comparison of the etiological groups demonstrated significant differences regarding age, body measurements, and mechanical ventilation duration parameters. The CRF group exhibited markedly prolonged pre- and post-tracheostomy ventilatory support durations compared with the remaining groups. Detailed comparisons of the ventilatory and clinical parameters are shown in Table II. Pairwise post hoc comparisons (Dunn's test with Bonferroni correction) are presented in Table II.

No statistically significant differences were identified between the groups in terms of PICU or overall hospital stay.

Early complications occurred in 10.2% of the patients (n=6), most commonly cannula-related problems/displacement (n=4) and air leakage (n=3); no bleeding was recorded. Late complications were observed in 25.4% of the patients (n=15), most commonly cannula-related problems/displacement (n=13) and granuloma (n=4); no tracheal stenosis, tracheomalacia or airway obstruction was recorded (some patients had more than one complication). Cannula displacement/malposition was the most frequently encountered complication during both the early and late periods. The overall clinical outcomes and ventilation-related findings are presented in Table III.

The overall decannulation rate was 5.1% (n=3), and the mortality rate was 22% (n=13). The mortality and decannulation rates according to their etiological group are summarized in Table IV; briefly, mortality was 29.0% in the NMs group, 20.0% in the CRF group, and 0% in the UAO/M

Table I. Demographic characteristics and comorbidities of patients

Variable	n (%) or median (IQR)
Total number of patients	59
Age (months)	11 (4-108)
Male sex	37 (62.7%)
Comorbidities	
Genetic syndrome	19 (32.2%)
Neurometabolic disease	17 (28.8%)
Chronic lung disease	10 (16.9%)
Others*	13 (22%)

*Permanent organ dysfunction secondary to various underlying conditions
IQR: Interquartile range

Table II. Clinical and ventilator parameters according to etiology groups

Variable	UAO/M	CRF	NM	p value	Post-hoc
Age (months)	5 (2.25-54.25)	4 (3-7.63)	62 (12-180)	<0.001	NM-UAO/M*, NM-CRF*
Weight (kg)	3.95 (3.43-10.88)	3.5 (3.23-3.98)	16 (7-44)	<0.001	NM-CRF*
Height (cm)	59 (51-95.75)	55 (50.25-57.5)	102 (65-148)	0.001	NM-CRF*
Pre-tracheostomy MV (days)	19 (1.75-34.5)	70.5 (42-96.75)	31 (24-36)	<0.001	CRF-UAO/M*, CRF-NM*
Post-tracheostomy MV (days)	9 (6-16)	130 (79-394)	16 (9-82)	0.002	CRF-UAO/M*, CRF-NM*
PICU length of stay (days)	71 (33.5-154.5)	133 (65.25-199)	67 (54-127)	0.199	-
Hospital length of stay (days)	140 (75.25-228)	202 (96.5-342)	105 (58-200)	0.059	-

*Statistically significant pairwise comparison
CRF: Chronic respiratory failure, NM: Neuromuscular diseases, UAO/M: Upper airway obstruction/malformation, MV: Mechanical ventilation, PICU: Pediatric intensive care unit, IQR: Interquartile range

Table III. Clinical course, mechanical ventilation, and outcomes

Clinical parameter	Median (IQR) or n (%)
Pre-tracheostomy MV (days)	34 (24-65)
Post-tracheostomy MV (days)	25 (11-129)
PICU length of stay (days)	87 (54-177)
Hospital length of stay (days)	124 (65-252)
Complications	
Early complications	6 (10.2%)
Late complications	15 (25.4%)
Cannula-related problem/displacement (early)	4
Air leakage (early)	3
Infection (early)	1
Cannula-related problem/displacement (late)	13
Granuloma (late)	4
Outcomes	
Decannulation	3 (5.1%)
Mortality	13 (22%)

MV: Mechanical ventilation, PICU: Pediatric intensive care unit, IQR: Interquartile range

Table IV. Mortality and decannulation rates according to etiology groups

Etiology group	n	Mortality, n (%)	Decannulation, n (%)
NM	31	9 (29.0%)	2 (6.5%)
CRF	20	4 (20.0%)	0 (0%)
UAO/M	8	0 (0%)	1 (12.5%)

CRF: Chronic respiratory failure; NM: Neuromuscular diseases; UAO/M: Upper airway obstruction/malformation

group, while decannulation was 6.5% in the NMs group, 0% in the CRF group, and 12.5% in the UAO/M group. Of the three patients who achieved successful decannulation, one was a young infant with an underlying genetic syndrome and upper airway obstruction/malformation who required prolonged pre-tracheostomy ventilation; one was an infant with an underlying neurometabolic disorder, classified in the NM group; and one was an adolescent patient, also classified in the NM group.

Ventilator weaning success differed significantly according to the etiological category (χ^2=9.401, p=0.009). The highest successful weaning rate was observed in the UAO/M group, whereas patients with CRF demonstrated the poorest weaning outcomes. Weaning results according to the etiological groups are summarized in Table V. This finding was confirmed using the Fisher-Freeman-Halton exact test (p=0.011), given the small expected cell count in the UAO/M subgroup.

Table V. Mechanical ventilation weaning rates according to etiology groups

Etiology group	Weaned n (%)	Continued MV n (%)
UAO/M	6 (75%)	2 (25%)
CRF	3 (15%)	17 (85%)
NM	13 (41.9%)	18 (58.1%)

Statistical test: chi-square (χ^2=9.401, p=0.009)
CRF: Chronic respiratory failure; NM: Neuromuscular diseases; UAO/M: Upper airway obstruction/malformation; MV: Mechanical ventilation

Discussion

The present study demonstrated that underlying tracheostomy indication was significantly associated with clinical outcomes in pediatric intensive care patients. In particular, children with CRF experienced prolonged mechanical ventilation and lower successful ventilator weaning rates compared with the other etiological groups.

Our findings are consistent with previous studies emphasizing the importance of underlying disease characteristics in determining post-tracheostomy outcomes (13,14). Can et al. (13) conducted a retrospective study including 63 pediatric patients in a Turkish tertiary PICU and reported that upper airway obstruction was associated with a higher decannulation rate (12.6% overall) and more favorable weaning outcomes than prolonged mechanical ventilation for chronic systemic disease, with no deaths directly attributable to tracheostomy complications. Similarly, Tsuboi et al. evaluated long-term survival and decannulation outcomes in a pediatric tracheostomy cohort and reported that children with underlying neurological and neuromuscular disorders had lower decannulation rates and more prolonged ventilator dependence than those with airway-based indications (14).

In the present cohort, the CRF group consistently exhibited less favorable outcomes across multiple clinical parameters. Persistent pulmonary pathology, impaired respiratory mechanics, and accompanying comorbid conditions may explain the prolonged dependence on mechanical ventilation observed in these patients. In addition, chronic respiratory disorders frequently require long-term respiratory support, thereby complicating ventilator liberation. Recent evidence has shown that tracheostomized children with chronic respiratory disease are more likely to require home mechanical ventilation and rehospitalization following discharge (15).

The prolonged pre- and post-tracheostomy ventilation durations identified in the CRF group are also consistent with previous reports demonstrating that prolonged ventilation and severe comorbidity burden are associated with poorer prognosis. A systematic review evaluating pediatric tracheostomy complications reported overall complication rates approaching 40%, particularly among younger patients and those with significant underlying disease (16).

The overall complication rate in our study was comparable with previously published pediatric tracheostomy series. Cannula displacement represented the most common complication and remains one of the major concerns in pediatric tracheostomy management.

The relatively low decannulation rate observed in this study may be explained by the large proportion of patients with chronic neuromuscular and respiratory disorders. Such patients frequently fail to achieve complete respiratory independence, thereby reducing the likelihood of successful decannulation. Previous reviews have identified effective cough reflex, preserved neurological function, and adequate secretion management as major predictors of successful decannulation, whereas chronic neurological disease is associated with lower decannulation success (17,18).

The overall mortality rate was 22%. Mortality in this population is likely more closely related to the severity of the underlying disease rather than the tracheostomy procedure itself. Previous studies similarly demonstrated that long-term survival in pediatric tracheostomy patients is strongly influenced by the primary diagnosis and associated comorbidities (14). Noy et al. (19) reported an overall mortality rate of 26% in a retrospective cohort of children undergoing tracheostomy, identifying prematurity, young age, congenital heart disease, and ventilatory dependency as factors associated with early mortality, while Kılıç and Başak Kılıç (20) reported a lower mortality rate (14.3%) in a smaller single-center Turkish PICU cohort in which most deaths were attributed to the underlying disease rather than tracheostomy-related complications, broadly consistent with our findings. Beyond mortality, recent population-level data have also indicated that health care resource usage and long-term care pathways after pediatric tracheostomy continue to evolve, further supporting the need for etiology-specific outcome data such as those presented here (11,12).

Taken together, these findings suggest that pediatric tracheostomy should not be approached as a uniform intervention, but rather as an etiology-specific management strategy. Careful patient selection and individualized follow-up planning which is tailored to the underlying etiology warrant further investigation in prospective studies as potential strategies to optimize both short- and long-term outcomes.

Study Limitations

This study had several limitations. First, the retrospective single-center design may limit generalizability. Second, the duration of mechanical ventilation prior to tracheostomy was available and analyzed; however, more granular information regarding the timing of the tracheostomy relative to the onset of the underlying illness or hospital admission, which would have allowed for the evaluation of "early" versus "late" tracheostomy strategies as discussed in the literature, was not systematically available, preventing additional timing-based analyses. The optimal timing of pediatric tracheostomy remains controversial and no universally accepted recommendations have yet been established for this patient population (21). Third, baseline differences in age and anthropometric characteristics among etiological groups may have partially confounded the observed associations between etiology and ventilatory outcomes. Due to the limited sample size in this study, multivariable adjustment was not performed, and these associations should therefore be interpreted with caution. Fourth, the UAO/M subgroup was small (n=8), which limits the statistical power of pairwise comparisons involving this group. Finally, patients enrolled toward the end of the study period had a shorter potential follow-up interval, which may have limited the ascertainment of later complications, decannulation, or mortality events in this subgroup.

Conclusion

The underlying tracheostomy indication appears to be significantly associated with clinical outcomes in pediatric intensive care patients. Children with CRF demonstrated prolonged ventilatory dependence and lower rates of successful ventilator weaning. Etiology-based differences should therefore be considered during both pre-tracheostomy evaluation and post-procedural management. Those patients with upper airway obstruction tended to exhibit more favorable respiratory outcomes, whereas chronic respiratory and neuromuscular disorders were associated with prolonged ventilatory support requirements. An individualized multidisciplinary management strategy tailored to the underlying disease process warrants further investigation as a means to help optimize outcomes in pediatric tracheostomy patients.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Institutional Review Board of University of Health Sciences Türkiye, İzmir City Hospital (approval no.: 2026/111; date: 18.02.2026).

Informed Consent: The requirement for informed consent was waived due to the retrospective nature of this study.

Footnotes

Authorship Contributions

Concept: F.D., E.P.K., G.Ö., D.K.T., Design: F.D., E.P.K., G.Ö., D.K.T., Data Collection or Processing: F.D., E.P.K., G.Ö., D.K.T., Analysis or Interpretation: F.D., E.P.K., G.Ö., D.K.T., Literature Search: F.D., E.P.K., G.Ö., D.K.T., Writing: F.D., E.P.K., G.Ö., D.K.T.

Conflict of Interest: The authors declare no conflict of interest.

Financial Disclosure: The authors declare that no financial or material support was received for this study.

Declaration Regarding the Use of AI and AI-Assisted Technologies: During the preparation of this work, the authors utilized OpenAI ChatGPT for language editing and academic writing support. All generated content was carefully reviewed, revised, and verified by the authors. The authors take full responsibility for the final content of the publication.

References

1. Freeman BD. Tracheostomy update: when and how. Crit Care Clin. 2017; 33:311-22.
2. Wood D, McShane P, Davis P. Tracheostomy in children admitted to paediatric intensive care. Arch Dis Child. 2012; 97:866-9.
3. Mahida JB, Asti L, Boss EF, et al. Tracheostomy placement in children younger than 2 years: 30-day outcomes using the Kids' Inpatient Database. JAMA Otolaryngol Head Neck Surg. 2016; 142:241-6.
4. Watters KF. Tracheostomy in infants and children. Respir Care. 2017; 62:799-825.
5. Trouillet J, Collange O, Belafia F, et al. Tracheotomy in the intensive care unit: guidelines from a French expert panel. Anaesth Crit Care Pain Med. 2018; 37:281-94.
6. Mitchell RB, Hussey HM, Setzen G, et al. Clinical consensus statement: tracheostomy care. Otolaryngol Head Neck Surg. 2013; 148:6-20.
7. Szakmany T, Russell P, Wilkes AR, Hall JE. Effect of early tracheostomy on resource utilization and clinical outcomes in critically ill patients. Br J Anaesth. 2015; 114:396-405.
8. Özmen S, Özmen ÖA, Ünal ÖF. Pediatric tracheotomies: a 37-year experience in 282 children. Int J Pediatr Otorhinolaryngol. 2009; 73:959-61.
9. Komori M. Update on pediatric tracheostomy. Auris Nasus Larynx. 2024; 51:429-32.
10. Fuller C, Wineland AM, Richter GT. Update on pediatric tracheostomy: indications, technique, education, and decannulation. Curr Otorhinolaryngol Rep. 2021; 9:188-99.
11. Ofosu D, Round J, Soril L, Kam K, Castro-Codesal M. Trends, outcomes, and resource use of pediatric tracheostomy in Alberta: a cohort study. Hosp Pediatr. 2026; 16:642-51.
12. Powell J, Keltie K, Sims A, Richardson H, Brodlie M, Powell S. National cohort study of health care resource use after pediatric tracheostomy. JAMA Pediatr. 2022; 176:817-9.
13. Can FK, Anıl AB, Anıl M, et al. The outcomes of children with tracheostomy in a tertiary care pediatric intensive care unit in Turkey. Turk Pediatri Ars. 2018; 53:177-84.
14. Tsuboi N, Ide K, Nishimura N, Nakagawa S, Morimoto N. Pediatric tracheostomy: survival and long-term outcomes. Int J Pediatr Otorhinolaryngol. 2016; 89:81-5.
15. Morrison JM, Kono N, Rush M, et al. Factors leading to rehospitalization for tracheostomized and ventilator-dependent infants through 2 years of age. Pediatr Pulmonol. 2024; 59:2761-71.
16. Lubianca Neto JF, Castagno OC, Schuster AK. Complications of tracheostomy in children: a systematic review. Braz J Otorhinolaryngol. 2022; 88:882-90.
17. Calderone A, Filoni S, De Luca R, et al. Predictive factors of successful decannulation in tracheostomy patients: a scoping review. J Clin Med. 2025; 14:3798.
18. Gallice T, Cugy E, Branchard O, et al. Predictive factors for successful decannulation in patients with tracheostomies and brain injuries: a systematic review. Dysphagia. 2024; 39:552-72.
19. Noy R, Eytan D, Cohen JT, Ostrovsky D, Shkedy Y, Gordin A. Factors associated with early mortality in pediatric tracheostomy placement: a retrospective cohort study. Eur J Pediatr. 2024; 184:5.
20. Kılıç S, Başak Kılıç A. Outcomes of tracheostomy in a pediatric intensive care unit: a single-center experience. Turk J Pediatr Surg. 2025; 39:47-52.
21. Güngör S, Çiğdem YK, Karaoğlanoğlu N, et al. Bedside tracheostomy for critically ill pediatric patients in the PICU: clinical experience in a single center. Children (Basel). 2025; 12:1558.

J Pediatr Res 2026;13(3):285-93
DOI: 10.4274/jpr.galenos.2026.19616

Original Article

Association between Breastfeeding Self-Efficacy and Infant Feeding Attitudes among Mothers of Infants Aged 0-6 Months

Dilek Zengin[1], Feyza Koç[2], Selinnur Durmaz[3], Merve Tosyalı[4]

[1]Ege University Faculty of Nursing, Department of Pediatric Nursing, İzmir, Türkiye
[2]Ege University Faculty of Medicine, Department of Pediatrics, İzmir, Türkiye
[3]Ege University Faculty of Nursing, İzmir, Türkiye
[4]Ege University Faculty of Medicine, Department of Pediatrics, İzmir, Türkiye

ABSTRACT

Aim: The purpose of this study was to determine the association between mothers' breastfeeding self-efficacy for infants aged 0-6 months and their attitudes toward infant feeding.

Materials and Methods: This descriptive, cross-sectional study was conducted with mothers of infants aged 0-6 months. Data were collected using a questionnaire assessing sociodemographic characteristics and breastfeeding status, as well as the Breastfeeding Self-Efficacy Scale and the Infant Feeding Attitude Scale. For the statistical analysis of the data, mean, standard deviation, percentage distributions, and linear regression analysis were employed.

Results: This study found that the mothers had high levels of breastfeeding self-efficacy (57.66±8.49) and positive infant feeding attitudes (56.99±7.09). According to the results of the regression analysis, it was determined that the mothers' breastfeeding self-efficacy scores positively and statistically significantly predicted their infant feeding attitude scores (β=0.208; p=0.001).

Conclusion: This study showed that mothers' perceptions of their breastfeeding self-efficacy positively influence their attitudes toward infant feeding. Accordingly, the development of clinical interventions focused on breastfeeding self-efficacy, particularly during the critical first six months after birth, may play an important role in increasing breastfeeding success.

Keywords: Breastfeeding, self-efficacy, attitude, mothers

Introduction

Breastfeeding is recognized as the gold standard for infant nutrition, offering significant health benefits for both the mother and the infant. Consistent evidence from diverse global settings indicates that breastfeeding reduces the incidence and severity of various conditions in both populations (1). Guidelines from the World Health Organization and United Nations Children's Fund recommend exclusive breastfeeding for the first 6 months, followed by continued breastfeeding alongside appropriate and safe complementary foods from 6 months of age until 2

Corresponding Author
Asst. Prof. Dilek Zengin, Ege University Faculty of Nursing, Department of Pediatric Nursing, İzmir, Türkiye
E-mail: zngn.dilek@gmail.com **ORCID:** orcid.org/0000-0002-4240-7090
Received: 08.07.2026 **Accepted:** 27.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Zengin D, Koç F, Durmaz S, Tosyalı M. Association between breastfeeding self-efficacy and infant feeding attitudes among mothers of infants aged 0-6 months. J Pediatr Res. 2026;13(3):285-93

years of age or beyond (2). Breastfeeding is one of the most effective strategies to protect children's health and ensure their survival. However, contrary to recommendations, less than half (44%) of infants aged 0-6 months are exclusively breastfed. If all children aged 0-23 months were breastfed optimally, the lives of more than 820,000 children under the age of 5 could be saved each year (3). Breast milk offers optimal and balanced nutrition. Breastfeeding provides numerous benefits for both the mother and the baby and it is fundamental for infants' health and their long-term well-being. Evidence indicates that breastfeeding significantly promotes healthy growth and immune system development. Those infants who are fed breast milk have a lower risk of asthma, obesity, diabetes, and severe lower respiratory tract infections. Concurrently, it has been reported that breastfeeding mothers have a lower risk of breast cancer, hypertension, diabetes, endometrial cancer, and ovarian cancer (4).

Breastfeeding self-efficacy is widely utilized in breastfeeding research, as defined by Dennis (7) in 1999, who adapted Bandura's self-efficacy theory to breastfeeding as being a mother's perceived confidence and belief in her ability to breastfeed (5,6). Breastfeeding self-efficacy theory posits that a mother's breastfeeding self-efficacy influences her response (effort and thought patterns) to breastfeeding, which subsequently affects the initiation and continuation of breastfeeding behaviours (7). Breastfeeding self-efficacy is a woman's belief and confidence in her perceived ability to breastfeed. Knowledge, attitudes, and intentions regarding breastfeeding also provide an indication of breastfeeding confidence and they are associated with improved breastfeeding outcomes (5).

The feeding method for infants aged 0-6 months is of critical importance for both the infant's and the mother's health. The literature frequently emphasizes the importance of breastfeeding self-efficacy in initiating and maintaining breastfeeding behaviour. However, there have been few studies examining the relationship between mothers' perceptions of their breastfeeding self-efficacy and their attitudes toward infant feeding, particularly during the first six months, which is the most critical period of development. In this context, this study aimed to determine the association between mothers' breastfeeding self-efficacy and their attitudes toward infant feeding for infants aged 0-6 months. In order to achieve this primary objective, this study examined mothers' levels of breastfeeding self-efficacy and their attitudes toward infant feeding. In addition, this study examined associations between the sociodemographic and the clinical characteristics of the mother and infant (e.g., the mother's age, infant's age, gestational age, etc.) and the mothers' breastfeeding self-efficacy and their attitudes toward infant feeding. This study hypothesized that the mothers' levels of breastfeeding self-efficacy would positively and significantly predict their infant feeding attitudes (H_1) and that the mothers' educational levels would show a significant association with their infant feeding attitude scores (H_2).

Materials and Methods

Study Design

This descriptive and cross-sectional study was conducted in order to determine the association between mothers' breastfeeding self-efficacy for infants aged 0-6 months and their attitudes toward infant feeding.

Setting and Sample

This descriptive/cross-sectional study was conducted between June and December 2025 at the outpatient clinics of Ege University Children's Hospital in the province of İzmir, Türkiye. With its numerous outpatient clinics, this hospital is one of the region's major centres, providing comprehensive pediatric diagnosis, treatment, and care services.

In the study's main hypotheses, it was planned to examine the relationships between the independent variables using correlation coefficients and regression models. The calculation yielding the highest number according to the main hypothesis was taken into account. Using the G*Power (version 3.1.9.2) program (8), at a 95% confidence level (α=0.05), the standardized effect size was taken as 0.15 (medium level for the Linear Regression Model and number of independent variables: 10), as proposed by Cohen (9) (2013) due to the originality of the study, and the minimum sample size was calculated as 172 with 95% theoretical power (8). Considering potential data losses and the risks of incomplete form filling, this study was completed with 180 participants. Those mothers who visited the Social Pediatrics (Healthy Child) Outpatient Clinic and had infants aged 0-6 months, who were literate in Turkish and had no communication barriers, and who volunteered to participate in this study were included in the research. Those mothers who had acute or chronic health problems themselves or in their infants, as well as those who did not complete the data collection forms in full, were excluded from this study.

Data Collection Tool

The data for this study were collected using the "Case Report Form", the "Breastfeeding Self-Efficacy Scale," and the "Infant Feeding Attitude Scale".

Case Report Form: This form was developed by the researchers in order to collect demographic information regarding the mothers and infants. It includes questions concerning the mother's age, educational background, occupation, parity, whether she has received breastfeeding education, and the duration of breastfeeding as well as the infant's postnatal age, gestational age at birth, and birth weight (1,4-6).

Breast-feeding Self-Efficacy Scale-Short Form: Developed by Dennis in 2003 and adapted into Turkish by Aluş Tokat et al. (10) (2008), this scale consists of 14 items. The "Breast-feeding Self-Efficacy Scale-Short Form" is a 5-point Likert-type scale ranging from 1 (Not at all confident) to 5 (Always confident). The Cronbach's alpha value was found to be 0.94. The total score range for this scale is between 14 and 70. A high score indicates high breastfeeding self-efficacy (10).

Infant Feeding Attitude Scale (Iowa): This scale, developed by De La Mora and Russel in 1999, was adapted into Turkish by Ekşioğlu et al. (11) (2016). This scale aims to assess women's attitudes toward breastfeeding. It consists of 17 items and uses a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Eight items on the scale contain positive statements regarding breastfeeding, while nine items contain positive statements regarding formula feeding. The total scale score ranges from 17 (indicating a positive attitude toward bottle-feeding) to 85 (reflecting a positive attitude toward breastfeeding). The scale has no cut-off value; a high score on this scale indicates a positive attitude toward breastfeeding. The scale's Cronbach's alpha internal consistency coefficient is 0.71 (11).

Data Collection

The data were collected by the researchers between June and December 2025 at the Social Pediatrics (Healthy Child) Outpatient Clinic using face-to-face interviews. The interviews were conducted in the outpatient clinic setting in a quiet area where the mothers could feel at ease and where confidentiality was ensured. The data collection instruments were completed by the mothers under the supervision of the researchers. The questionnaires were handed to the mothers, who were allowed to complete them independently in a quiet area. The researchers remained on-site throughout the process in order to explain the questions as needed and to check the returned forms for any missing items. The administration of the data collection tools took an average of 15 minutes. The questionnaire forms were completed either during the infants' pre-examination waiting period or at an appropriate time following the examination in order to avoid any external influences.

Ethical Considerations

Written institutional approval was obtained from the Medical Research Ethics Committee of Ege University (approval no.: 25-6T/25, date 20.06.2025) and from the hospital where this study was conducted (date/ref. no.: 03.06.2025-E.2470168) before carrying out this study. During the data collection phase, in accordance with the principles of the Declaration of Helsinki, the mothers were informed about the purpose and scope of the study; an "Informed Consent Form" was obtained from those who agreed to participate. Participants were informed that they had the right to withdraw from this study at any time, that the data collected would be used solely for scientific purposes, and that their personal identifying information would not be used.

Statistical Analysis

In this study, descriptive statistics for the data (frequency, percentage, mean, standard deviation, median, minimum, maximum, and the 25th and 75th percentiles) are presented. The assumption of normal distribution was tested using the Shapiro-Wilk test. When the normality assumption was not met, the Mann-Whitney U test was used to compare two independent groups, and the Kruskal-Wallis test was used to compare three groups. Post-hoc Bonferroni-corrected tests were used to identify the group or groups responsible for the difference. Spearman correlations were performed in order to examine relationships between continuous measurements which did not follow a normal distribution. Linear regression analysis using the stepwise approach was employed to mathematically model relationships between dependent continuous variables and independent variables. Analyses were conducted using the IBM Statistical Package for the Social Sciences (SPSS) Windows version 27 (SPSS Inc., Chicago, IL, USA).

Results

The distributions of the demographic and clinical characteristics of the mothers and infants participating in this study were examined and the data obtained are summarized in Table I. When the educational backgrounds

of the participating mothers were examined, it was found that the vast majority of the participants held a bachelor's degree (33.3%) or high school diploma (30.0%). Regarding the mothers' occupational distribution, 40.6% (n=73) were not employed in any job, while the total proportion of mothers working in healthcare (doctors, nurses, and other healthcare personnel) was determined to be 14.4%.

It was determined that 57.8% of the mothers had only one child and 97.8% had had a singleton pregnancy. In terms of the mode of delivery, 73.3% of the mothers (n=132) had had a caesarean section, while 26.7% (n=48) had given birth vaginally. When evaluating the mothers' educational status regarding breastfeeding, 58.9% (n=106) of the participants had received prior education. Among those who had received education, 66.0% (n=70) had been educated by nurses and 34.0% (n=36) by midwives. The mean age of the mothers included in this study was found to be 30.51±5.04 years. The average postnatal age of the infants was 3.79±2.12 months, and the average gestational age at birth was recorded as 37.59±2.12 weeks. The mothers' average daily breastfeeding duration was 185.83±101.49 minutes, and the frequency of breastfeeding was determined to be 9.94±3.2 times per day (Table I).

The descriptive statistics and internal consistency coefficients for the scales used in this study are presented in Table II. Analysis of the data from the Breastfeeding Self-Efficacy Scale revealed that the mothers' self-efficacy score was 57.66±8.49. This result indicates that the mothers included in this study had high self-efficacy regarding breastfeeding. The reliability analysis of the scale showed that the Cronbach's alpha coefficient was 0.829, indicating a high level of reliability. Upon examining the scores obtained from the Infant Feeding Attitude Scale, it was determined that the mothers' mean score was 56.99±7.09 and that their attitudes toward infant feeding were at a moderate level. The scale's internal consistency coefficient in this study fell below the thresholds generally accepted in the literature, indicating limited consistency. It is noted that while the Cronbach's alpha value should be higher than 0.70, a level of 0.60 is also considered acceptable in exploratory studies (12). Since the reliability of the scale's original form (Cronbach's alpha=0.71) has been established in the literature, this scale was included in the analyses despite the low reliability observed in our study.

The distributions of scores on the Infant Feeding Attitude Scale and the Breast-feeding Self-Efficacy Scale according to mothers' demographic characteristics are presented in Table III. When examining the distribution of infant

Table I. Distributions of descriptive and demographic characteristics of mothers and infants

		n	**%**
The mother's educational background	Elementary school	11	6.1
	Middle school	20	11.1
	High school	54	30.0
	Associate's degree	29	16.1
	Bachelor's degree	60	33.3
	Postgraduate degree	6	3.3
Mother's occupation	Not working	73	40.6
	Nurse	6	3.3
	Doctor	7	3.9
	Other healthcare worker	13	7.2
	Teacher	16	8.9
	Civil servant	5	2.8
	Private sector	27	15.0
	Other	33	18.3
Number of children in the family	1	104	57.8
	2	59	32.8
	3	14	7.8
	4	3	1.7
Multiple pregnancy	Yes	4	2.2
	No	176	97.8
Mode of delivery	Vaginal	48	26.7
	Caesarean	132	73.3
The mother's status regarding breastfeeding education	Yes	106	58.9
	No	74	41.1
Where/from whom was the training received?	Nurse	70	66.0
	Midwife	36	34.0
	X±SD (Median)	**Min-Max.**	
Maternal age (years)	30.51±5.04 (30)	19-43	
Infant's postnatal age (months)	3.79±2.12 (4)	0-6	
Gestational age at birth (weeks)	37.59±2.12 (38)	29-42	
Birth weight (g)	3062.24±584.18 (3100)	800-4,300	
Duration of breastfeeding (minutes/day)	185.83±101.49 (160)	20-600	
Breastfeeding frequency (times/day)	9.94±3.2 (10)	2-24	

SD: Standard deviation, Min: Minimum, Max: Maximum

feeding attitude scale scores according to the mothers' demographic characteristics, a statistically significant difference was found between the mothers' educational status and their attitude scores (p=0.008). A numerical downward trend in attitude scores was observed as the mothers' educational level increased. It was found that the number of children the mothers had (p=0.717) and whether they had received education on breastfeeding (p=0.543) did not create statistically significant differences in the total scores of the infant feeding attitude scale (p>0.05).

When the distribution of breast-feeding self-efficacy scale scores was examined according to the mothers' demographic characteristics, no statistically significant difference was found between the mothers' educational status (p=0.102), the total number of children in the family (p=0.362), and whether they had received education on breast-feeding (p=0.907) in relation to their breast-feeding self-efficacy scores (p>0.05) (Table III).

Spearman's correlation coefficients were used to examine the relationships between the participants' demographic characteristics and their total scores on the Infant Feeding Attitude Scale and the Breastfeeding Self-Efficacy Scale, with these findings being presented in Table IV. The analysis revealed a positive, weak, and statistically significant relationship between the total scores of the two scales (r=0.242; p=0.001). Accordingly, it was determined that as the mothers' infant feeding attitudes increased in a positive direction, their perceptions of breastfeeding self-efficacy also increased. When other variables were examined, a negative, weak, and statistically significant relationship was found between the infant's postnatal age and the scores on the Infant Feeding Attitude Scale (r=-0.181; p=0.015). However, no statistically significant relationship was found between the mothers' age, the gestational age at birth, the birth weight, and the duration of breastfeeding and the total scores on either scale (p>0.05).

Table II. Descriptive statistics and internal consistency coefficients of the scales used

	X±SD (Median)	Min-Max.	Cronbach's alpha
Breast-feeding self-efficacy scale	57.66±8.49 (59)	33-70	0.829
Infant feeding attitude scale	56.99±7.09 (57)	37-77	0.637

SD: Standard deviation, Min: Minimum, Max: Maximum

Table III. Comparison of scores on the infant feeding attitude scale and the breastfeeding self-efficacy scale according to the mothers' demographic characteristics

		Infant feeding attitude scale				Breast-feeding self-efficacy scale			
		X±SD	Median (IQR)	Test stat.	p value	X±SD	Median (IQR)	Test stat.	p value
Maternal education level	Elementary school	62.45±9.34	65 (60-69)	15.670[a]	0.008*	63.18±5.4	63 (59-69)	9.194[a]	0.102
	Middle school	59.7±9.03	62 (54.5-64.5)			60.4±6.24	61.5 (55.5-64.5)		
	High school	56.54±6.15	56.5 (53-61)			56.41±8.67	57.5 (50-64)		
	Associate's degree	55.9±7.44	57 (51-60)			58.72±8.28	60 (54-67)		
	Bachelor's degree	56.27±5.91	57 (53-60)			56.42±9.24	59 (52-62.5)		
	Graduate degree	54.67±8.71	56 (53-59)			57±6.84	59.5 (53-62)		
Number of children	1	56.51±6.96	57 (52-60.5)	0.451[a]	0.717	57.05±8.46	58.5 (52-63)	3.199[a]	0.362
	2	57.47±6.91	57 (54-63)			58.19±8.82	60 (54-65)		
	3	58.14±8.97	59.5 (53-62)			58.64±7.94	58 (54-65)		
	4	59±7.81	55 (54-68)			64±3.61	63 (61-68)		
Breastfeeding education	Yes	57.26±6.8	57 (53-61)	0.610[b]	0.543	57.82±8.06	59 (54-63)	-0.116[b]	0.907
	No	56.61±7.52	57.5 (52-62)			57.43±9.13	60 (52-65)		

*p<0.05, [a]: Kruskal-Wallis test, [b]: Mann-Whitney U test
IQR: Interquartile range, SD: Standard deviation

In order to include categorical variables in the regression models, dummy variables were created, and certain groups were selected as references. For educational status, the elementary school group was taken as the reference. Linear regression analysis using the stepwise approach was applied in order to investigate the effect of participants' Breastfeeding Self-Efficacy Scale scores, as well as their age and educational status, on the total scores of the Infant Feeding Attitude Scale. An ANOVA test was conducted in order to test the significance of the model, and the established regression model was found to be significant (F=11.776 and p<0.001). The proportion of the dependent variable explained by the independent variables in the model was 5.7% (R^2=0.057). Upon examining the VIF values of the variables in the model, it was determined that all values were less than 5, indicating no multicollinearity issues, and there was no autocorrelation problem (Durbin-Watson statistic: 1.897). The Shapiro-Wilk test was used to check the standardized residuals, and it was found that the assumption of normal distribution was met (p>0.05). The standardized regression coefficient (β) for the Breastfeeding Self-Efficacy Scale was calculated as 0.208 and this was found to be statistically significant (p=0.001) (Table V).

Table IV. Relationships between mothers' demographic characteristics and total scores on the infant feeding attitude and breastfeeding self-efficacy scale

		Infant feeding attitude scale	Breast-feeding self-efficacy scale
Maternal age	r	-0.091	-0.046
	p	0.225	0.543
Infant's postnatal age (months)	r	-0.181	-0.128
	p	0.015*	0.086
Gestational age at birth	r	-0.105	-0.015
	P	0.159	0.846
Birth weight (g)	r	-0.135	0.031
	p	0.071	0.678
Duration of breastfeeding (minutes/day)	r	0.124	-0.073
	p	0.096	0.331
Infant feeding attitude scale	r	-	0.242
	p	-	0.001*

*p<0.05, R: Spearman's correlation coefficient

Table V. Linear regression analysis for the infant feeding attitude scale

Predictors	**β**	**S.E.**	**Std. β**	**t**	**p**	**95% CI lower limit**	**95% CI upper limit**	**VIF**
Constant	45.004	3.532		12.743	<0.001*	38.035	51.973	
Breast-feeding self-efficacy scale	0.208	0.061	0.249	3.432	0.001*	0.088	0.328	1.000
F:11.776 and p<0.001*				R^2=0.062 and adjusted R^2=0.057				
Variables not in the model						**β**	**t**	**p**
Maternal age						-0.084	-1.155	0.250
Maternal education level (Middle school)						0.108	1.486	0.139
Maternal education level (High school)						-0.018	-0.251	0.802
Maternal education level (Associate degree)						-0.082	-1.129	0.260
Maternal education level (Bachelor's degree)						-0.047	-0.649	0.517
Maternal education level (Postgraduate)						-0.058	-0.792	0.430

*p<0.05, S.E.: Standard error, Std.: Standard, Std. β.: Standardized beta coefficient, CI: Confidence interval, VIF: Variance inflation factor

Discussion

This study was conducted in order to determine the association between mothers' breastfeeding self-efficacy for infants aged 0-6 months and their attitudes toward infant feeding. The main findings of this study indicated that mothers' breastfeeding self-efficacy scores were positive and statistically significant predictors of their infant feeding attitude scores (β=0.208; p<0.05). This result demonstrates that the mothers' perceived sense of competence regarding their breastfeeding abilities serves as a determinant in shaping their attitudes toward infant feeding.

Discussion of Findings Regarding Mothers' Breastfeeding Self-Efficacy and Attitudes Toward Infant Feeding

In this study, the mean breastfeeding self-efficacy score for the mothers was found to be 57.66±8.49. This finding indicates that the mothers' perceptions of their self-efficacy regarding the breastfeeding process are high. A review of the literature revealed that a study conducted by Akça and Yıldız (13) similarly found that mothers' breastfeeding self-efficacy scores were 54.0±7.1, indicating a high level of self-efficacy. In this regard, our research findings are consistent with the literature. However, Akça and Yıldız (13) reported that those mothers who received prenatal breastfeeding education (p=0.008) and those with higher education levels (p<0,001) had higher breastfeeding self-efficacy scores, which contradicts our findings. The findings of this study are consistent with those reported by Konukoğlu and Pasinlioğlu (14), who found that pregnant women's perceptions of breastfeeding self-efficacy were above average. They also noted that participants with undergraduate and graduate education levels, as well as those experiencing their second pregnancy, had significantly higher levels of breastfeeding self-efficacy compared to the other groups (14). However, these findings differ from our results, as we found no statistically significant relationship between breastfeeding self-efficacy and the mothers' educational status (p=0.102) or their number of children (p=0.362). The literature highlights the lasting effects of breastfeeding education on self-efficacy. For instance, a study conducted by Golnam et al. (15) demonstrated that self-efficacy scores, which were at an intermediate level prior to the intervention, increased significantly at the 3- and 6-month follow-ups, whereas the control group experienced a decline over time. Additionally, other studies have indicated that breastfeeding education enhances mothers' breastfeeding self-efficacy (16,17). These findings suggest that the high average self-efficacy scores observed in our study may be related to the fact that more than half (58.9%) of the mothers in the sample had received breastfeeding education.

In this study, the mean score on the infant feeding attitude scale for mothers was found to be 56.99±7.09. This finding suggests that mothers possess a positive attitude toward infant feeding, which supports breastfeeding and prioritizes breast milk. A positive maternal attitude toward breastfeeding during the 0-6-month period, which is crucial for the protection and promotion of infant health, is consistent with the prior literature (18-22). Similarly, Kurtbeyoglu and Caferoğlu (23) determined that mothers exhibited a positive attitude toward infant feeding, reporting a mean score of 65.08±6.8. Our study found an inverse relationship between the mothers' education levels and their infant feeding attitude scores; specifically, as the mothers' education levels increased, their attitude scores decreased. Additionally, we observed that the mothers' number of children (p=0.717) and their prenatal breastfeeding education status (p=0.543) had no statistically significant effects on their infant feeding attitudes. However, contrary to our findings, several studies in the literature have indicated that infant feeding attitudes improve as the mothers' education levels rise (23,24). Factors such as highly educated mothers returning to work early during the postpartum period due to employment and career goals, working conditions, the amount of time spent at work, and greater financial resources may lead to the use of formula. This suggests that there may be an indirect relationship between higher levels of education and a tendency to use formula.

The Determining Role of Breastfeeding Self-Efficacy on Infant Feeding Attitudes

The most striking finding of this study was that the mothers' breastfeeding self-efficacy scores positively and statistically significantly predicted their infant feeding attitude scores (β=0.208; p<0.05). This finding supports the notion that a mother's belief in her ability to breastfeed directly influences her cognitive attitude toward breastfeeding. This finding aligns with Bandura's Self-Efficacy Theory. While previous studies in the literature have sought to explain mothers' attitudes toward infant feeding using demographic variables, our study contributes a unique perspective to the literature by highlighting the impact of a psychosocial factor, namely self-efficacy, on these attitudes. Specifically, Selvi et al. (25) reported that breastfeeding education significantly increased self-efficacy perceptions during both the antenatal and

postnatal periods. However, they noted that high levels of breastfeeding self-efficacy among mothers did not influence breastfeeding attitudes during the postnatal period. This particular finding contrasts with both Bandura's Self-Efficacy Theory and the current study (25). The literature indicates that high levels of breastfeeding self-efficacy among mothers are directly associated with the tendency to feed their infants exclusively with breast milk and with prolonged breastfeeding durations (6,26,27). Furthermore, this phenomenon is not limited to healthy pregnancies; even among mothers who had experienced high-risk pregnancies or delivered premature infants, high self-efficacy perceptions were positively associated with breastfeeding success, prolonged breastfeeding, and increased milk production (28,29). By contrast, Ergezen et al. (30) found no association between breastfeeding self-efficacy and breastfeeding success, reporting instead that as the number of children increased, both mothers' breastfeeding self-efficacy and breastfeeding success increased. Additionally, Awaliyah et al. (31) demonstrated that breastfeeding self-efficacy is positively correlated with breastfeeding satisfaction (31).

The findings of this study indicate that postpartum mothers' levels of breastfeeding self-efficacy have a direct and decisive effect on their infant feeding attitudes. This clinical picture redefines the role of nurses who provide breastfeeding counselling, maternal education, and postpartum care services. In particular, nurses working in primary care settings and in areas such as neonatal clinics and well-baby clinics should assess breastfeeding self-efficacy using validated scales during routine infant follow-ups. Interventions which strengthen the mother's breastfeeding self-efficacy (such as peer support, positive feedback, hands-on demonstrations of breastfeeding techniques, and anxiety management) should be incorporated into the focus of care. It is anticipated that as the mother's perception of self-efficacy increases, her attitudes toward infant feeding will also improve positively, thereby ensuring the sustainability of breastfeeding and thus support mother-infant health.

Study Limitations

The strengths of this study include its focus on mothers during the first six months which is a critical postpartum period, the use of validated measurement tools, and the examination of breastfeeding self-efficacy as a psychosocial predictor of infant feeding attitudes. This study had some limitations. This study was conducted at a single centre; this may limit the generalizability of our findings. Additionally, in this study, the internal consistency of the Infant Feeding Attitudes Scale was lower than the value reported in the original validity study (Cronbach's α=0.637 versus 0.71). This relatively low reliability may be related to the homogeneity of the study sample, cultural differences, or the narrow age range of the infants (0-6 months).

Conclusion

In conclusion, this study has demonstrated that the mothers' perceptions of breastfeeding self-efficacy are a significant psychosocial factor in determining their infant feeding attitudes. It is believed that self-efficacy-focused clinical interventions and nursing counselling aimed at enhancing mothers' confidence in their own breastfeeding skills, particularly during the first six months postpartum, may contribute to an increase in breastfeeding success rates. In this context, it is recommended that healthcare professionals working in primary care, obstetrics, and neonatology clinics integrate psychosocial approaches grounded in theory (such as Bandura's Self-Efficacy Theory) into routine breastfeeding education in order to strengthen mothers' perceptions of self-efficacy.

Ethics

Ethics Committee Approval: Written institutional approval was obtained from the Medical Research Ethics Committee of Ege University (approval no.: 25-6T/25, date 20.06.2025) and from the hospital where this study was conducted (date/ref. no.: 03.06.2025-E.2470168) before carrying out this study.

Informed Consent: During the data collection phase, in accordance with the principles of the Declaration of Helsinki, the mothers were informed about the purpose and scope of the study; an "Informed Consent Form" was obtained from those who agreed to participate.

Footnotes

Authorship Contributions

Concept: D.Z., F.K., Design: D.Z., F.K., Data Collection or Processing: D.Z., S.D., M.T., Analysis or Interpretation: D.Z., Literature Search: D.Z., S.D., M.T., Writing: D.Z., F.K.

Conflict of Interest: The authors declare no conflict of interest.

Financial Disclosure: The authors declare that no financial or material support was received for this study.

References

1. Otsuka K, Taguri M, Dennis CL, et al. Effectiveness of a breastfeeding self-efficacy intervention: do hospital practices make a difference?. Matern Child Health J. 2014; 18:296-306.

2. World Health Organization, UNICEF. Indicators for assessing infant and young child feeding practices: definitions and measurement methods [Internet]. Geneva: World Health Organization; 2021 [cited 2026 Apr 7]. Available from: https://apps.who.int/iris/handle/10665/340706

3. World Health Organization. Infant and young child feeding [Internet]. Geneva: World Health Organization; 2023 [cited 2026 Apr 7]. Available from: https://www.who.int/news-room/fact-sheets/detail/infant-and-young-child-feeding

4. Değer MS, Sezerol MA, Altaş ZM. Breastfeeding self-efficacy, personal well-being and related factors in pregnant women living in a district of Istanbul. Nutrients. 2023; 15:4541

5. McGovern LM, O'Toole L, Laws RA, Skinner TC, McAuliffe FM, O'Reilly SL. An exploration of prenatal breastfeeding self-efficacy: a scoping review. Int J Behav Nutr Phys Act. 2024; 21:9.

6. Brockway M, Benzies K, Hayden KA. Interventions to improve breastfeeding self-efficacy and resultant breastfeeding rates: a systematic review and meta-analysis. J Hum Lact. 2017; 33:486-99.

7. Dennis CL. Theoretical underpinnings of breastfeeding confidence: a self-efficacy framework. J Hum Lact. 1999; 15:195-201.

8. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007; 39:175-91.

9. Cohen J. Statistical power analysis for the behavioral sciences. New York: Routledge; 2013.

10. Aluş Tokat M, Okumuş H, Dennis CL. Translation and psychometric assessment of the breast-feeding self-efficacy scale-short form among pregnant and postnatal women in Turkey. Midwifery. 2010; 26:101-8.

11. Ekşioğlu A, Yeşil Y, Turfan EÇ. Bebek beslenmesi tutum ölçeğinin (Iowa) Türkçeye uyarlanması: geçerlik ve güvenirlik çalışması. Hemşirelikte Eğitim ve Araştırma Dergisi. 2016; 13.

12. Hair JF Jr, Black WC, Babin BJ, Anderson RE. Multivariate data analysis. 8th ed. Cengage Learning; 2019.

13. Akça G, Yıldız E. Mothers' breastfeeding self-efficacy level and factors affecting breastfeeding. Turk J Fam Pract. 2024; 28:18-24.

14. Konukoğlu T, Pasinlioğlu T. Gebelerde emzirme öz-yeterliliği ve etkileyen faktörlerin belirlenmesi. J Midwifery Health Sci. 2021; 4:12-22.

15. Golnam M, Hassani L, Goodarzi R, Ghanbarnejad A. The effectiveness of a theory based health education program on self-efficacy and breastfeeding behaviors continuity of working mothers in Iran. Sci Rep. 2025; 15.

16. Demir R, Taşpınar A. Doğum sonrası farklı eğitim yöntemleriyle verilen taburculuk eğitiminin taburculuğa hazır oluşluk, doğum sonu uyum süreci ve emzirme öz-yeterliliğine etkisi. Adnan Menderes Üniversitesi Sağlık Bilimleri Fakültesi Dergisi. 2022; 6:97-115.

17. Duman FN, Gölbaşı Z. Investigation of the effect of breastfeeding education on mothers' breastfeeding self-efficacy levels in Turkey. Turk J Fam Med Prim Care. 2022; 16:140-50.

18. Bednarek A, Bodys-Cupak I, Serwin A, Cipora E. Mothers' attitudes towards breastfeeding in terms of health safety and professional lactation education: a national survey of women. J Multidiscip Healthc. 2023; 16:3273-86.

19. Bień A, Kulesza-Brończyk B, Przestrzelska M, Iwanowicz-Palus G, Ćwiek D. The attitudes of Polish women towards breastfeeding based on the Iowa Infant Feeding Attitude Scale (IIFAS). Nutrients. 2021; 13:4338.

20. Han FL, Ho YJ, McGrath JM. The influence of breastfeeding attitudes on breastfeeding behavior of postpartum women and their spouses. Heliyon. 2023; 9:e1398.

21. Suppiger D, Natalucci G, Reinelt T. Reliability and validity of the German version of the Iowa Infant Feeding Attitude Scale (IIFAS-G) and relations to breastfeeding duration and feeding method. Int Breastfeed J. 2024; 19:58.

22. Twells LK, Midodzi WK, Ludlow V, et al. Assessing infant feeding attitudes of expectant women in a provincial population in Canada: validation of the Iowa Infant Feeding Attitude Scale. J Hum Lact. 2016; 32:NP9-18.

23. Kurtbeyoglu E, Caferoğlu Z. Annelerin bebek beslenmesi tutumları ve ilişkili faktörler. İzmir Katip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi. 2023; 8:925-32.

24. Chaabna N, Mahfoud ZR, Letourneau N, Forgrave D, White D. Muslim women's attitudes toward infant feeding in Qatar: an exploration using the Iowa Infant Feeding Attitude Scale. Midwifery. 2022; 114:103470.

25. Selvi Y, Desdicioğlu R, Eray İK. Antenatal dönemde verilen emzirme eğitiminin emzirme bilgi düzeyi ve emzirme öz yeterlilik ölçeğine etkisinin değerlendirilmesi. Ankara Med J. 2021; 21.

26. Blyth R, Creedy DK, Dennis CL, Moyle W, Pratt J, De Vries SM. Effect of maternal confidence on breastfeeding duration: an application of breastfeeding self-efficacy theory. Birth. 2002; 29:278-84.

27. Saavedra Sanchez S, Rodríguez-Gallego I, Leon-Larios F, Andina-Diaz E, Perez-Contreras R, Gonzalez-Sanz JD. Influence of perceived maternal self-efficacy on exclusive breastfeeding initiation and consolidation: a systematic review. Healthcare. 2024; 12:2347.

28. Brani P, Mrvoljak-Theodoropoulou I, Pechlivani F, et al. Mothers' breastfeeding self-efficacy after a high-risk or normal pregnancy: a Greek longitudinal cohort study. Eur J Investig Health Psychol Educ. 2024; 14:1803-20.

29. Liu Y, Shan H, Chen Y, Yuan Z, Zhou J, Wang L. Breastfeeding self-efficacy, milk output, and influencing factors in mothers of preterm infants: a prospective cohort study in the NICU. BMC Pregnancy Childbirth. 2025; 25.

30. Ergezen Y, Efe E, Çalışkan F, Dikmen Ş. Doğum sonu dönemde annelerin emzirme öz-yeterlilik algıları ile emzirme başarıları arasındaki ilişki. Dokuz Eylül Üniversitesi Hemşirelik Fakültesi Elektronik Dergisi. 2021; 14:217-23.

31. Awaliyah SN, Rachmawati IN, Rahmah H. Breastfeeding self-efficacy as a dominant factor affecting maternal breastfeeding satisfaction. BMC Nurs. 2019; 18(Suppl 1):30.

J Pediatr Res 2026;13(3):294-302
DOI: 10.4274/jpr.galenos.2026.37039

Original Article

Large Language Models in Pediatric Penile Conditions: Guideline Concordance, Readability, and Language-Based Performance

Bilge Türedi[1], Ali Sezer[2]

[1]University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Pediatric Urology, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Kocaeli City Hospital, Clinic of Pediatric Urology, Kocaeli, Türkiye

ABSTRACT

Aim: Large language models (LLMs) are increasingly being used by patients and caregivers in order to obtain medical information. This study evaluated the guideline concordance and public readability of two contemporary LLMs in answering frequently asked questions regarding common pediatric penile conditions.

Materials and Methods: Five common pediatric penile conditions were evaluated: phimosis, hypospadias, congenital penile curvature, paraphimosis, and buried penis. Ten standardized questions were developed for each condition, covering diagnosis, treatment, surgical considerations, complications, and long-term outcomes. A total of 50 questions were submitted to two widely used LLMs in both English and Turkish. The responses were assessed according to the European Association of Urology Paediatric Urology Guidelines and graded as fully concordant (Grade 1), partially concordant (Grade 2), partially concordant with inaccuracies (Grade 3), or non-concordant (Grade 4). Public readability was evaluated using a five-point scale.

Results: ChatGPT (OpenAI, GPT-5.5 version; accessed June 2026) achieved Grade 1 concordance in 40 out of 50 responses (80%), with 49 responses (98%) considered acceptable (Grade 1 or 2). Gemini (Google DeepMind, Gemini 3.5 Pro version; accessed June 2026) achieved Grade 1 concordance in 41 of 50 responses (82%), and all responses (100%) were classified as acceptable. No Grade 4 responses were identified, and only one Grade 3 response was observed overall. Concordance rates did not differ significantly between the 2 models ($p>0.05$). Public readability was high for both systems (4.60±0.58 vs. 4.67±0.49; $p>0.05$). No significant differences were observed between the English and Turkish responses. Reduced concordance occurred in different questions, with only two overlapping areas identified.

Conclusion: Both LLMs demonstrated high guideline concordance and excellent public readability. No clinically significant misinformation was identified, and most deficiencies reflected omissions of specific guideline details. LLMs may serve as reliable educational resources but should complement, rather than replace, professional medical consultation.

Keywords: Artificial intelligence, clinical practice guidelines, hypospadias, large language models, patient education

Corresponding Author
Bilge Türedi, MD, University of Health Sciences Türkiye, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Clinic of Pediatric Urology, İstanbul, Türkiye
E-mail: blgtrd@gmail.com **ORCID:** orcid.org/0000-0003-3532-0912
Received: 13.06.2026 **Accepted:** 09.07.2026 **Epub:** 21.08.2026 **Publication Date:** 18.09.2026

Cite this article as: Türedi B, Sezer A. Large language models in pediatric penile conditions: guideline concordance, readability, and language-based performance. J Pediatr Res. 2026;13(3):294-302

Introduction

Pediatric penile conditions constitute a common source of concern for parents and caregivers. Conditions such as phimosis, hypospadias, congenital penile curvature (CPC), paraphimosis, and buried penis frequently prompt families to seek medical information long before specialist consultation. Although many of these conditions are benign or treatable, misconceptions and misinformation may lead to unnecessary anxiety, delayed presentation, inappropriate treatment attempts, or unrealistic expectations regarding outcomes (1).

In recent years, large language models (LLMs) have emerged as increasingly popular sources of health information. Millions of users now consult conversational artificial intelligence platforms for medical advice, explanations of diagnoses, and treatment recommendations. Consequently, the quality, accuracy, and readability of the information generated by these systems have become important topics of investigation (2). Early studies evaluating LLM-generated medical information reported variable performance, with concerns regarding factual inaccuracies, hallucinations, and incomplete recommendations. However, rapid model development and continuous training have resulted in substantial improvements in both medical accuracy and user-oriented communication (3).

Several studies have evaluated the quality of online health information using content obtained from search engines, social media platforms, video-sharing websites, or frequently asked questions (FAQs) collected from public sources (4). However, standardized evaluations of LLM performance within narrowly defined pediatric urological domains remain limited. Furthermore, comparisons between different languages and between contemporary LLM platforms are scarce (5).

Pediatric penile conditions represent an ideal model for such an evaluation as they are frequently encountered in clinical practice, generate substantial parental concern, and are associated with numerous recurring questions regarding diagnosis, treatment, surgery, long-term outcomes, and fertility (6).

The aim of this study was to assess the guideline concordance and public readability of LLM-generated responses regarding common pediatric penile conditions and to compare the performance of two contemporary LLM platforms across different languages.

Materials and Methods

Study Design

This cross-sectional comparative study evaluated the quality and guideline concordance of responses generated by two LLMs regarding common pediatric penile conditions. This assessment focused on both the medical accuracy and the public readability of the responses provided in both English and Turkish.

Selection of Clinical Topics and Questions

Five commonly encountered pediatric penile conditions were selected based on their clinical relevance and frequency in pediatric urology practice:

1. Phimosis,
2. Hypospadias,
3. CPC,
4. Paraphimosis,
5. Buried penis.

In order to ensure a systematic and comprehensive assessment across all conditions, questions were developed according to predefined thematic domains commonly addressed by patients and caregivers seeking health information. These domains included:

- Definition,
- Epidemiology,
- Etiology,
- Symptoms,
- Diagnosis,
- Classification (when applicable),
- Treatment,
- Surgical considerations,
- Complications,
- Long-term outcomes.

For each disease, ten FAQs were generated, resulting in a total of 50 questions. Equivalent question sets were prepared in both English and Turkish.

LLM Response Generation

Each question was submitted separately to the evaluated LLMs in both English and Turkish. All responses were generated during the same study period using the default model settings without additional prompting, refinement, or follow-up interactions. The responses were collected and archived for subsequent analysis.

Guideline Concordance Assessment

The medical accuracy of each response was evaluated using the current European Association of Urology (EAU) Guidelines on Paediatric Urology-2026 as the reference standard (7). Guideline concordance assessments were performed independently by two pediatric urologists certified by the European Board of Paediatric Urology. Any discrepancies in grading were resolved through discussion and consensus.

Each response was independently reviewed and assigned one of four concordance grades:

- **Grade 1:** Fully concordant with the guideline recommendations,
- **Grade 2:** Generally concordant but containing minor omissions or incomplete details,
- **Grade 3:** Partially concordant but including inaccurate or potentially misleading information,
- **Grade 4:** Non-concordant with the guideline recommendations.

For secondary analyses, responses were categorized as:

- **Successful:** Grade 1.
- **Partially successful:** Grade 2.
- **Failed:** Grade 3 or Grade 4.

Additionally, acceptable responses were defined as those being Grade 1 or Grade 2.

Public Readability Assessment

In order to evaluate the usefulness of LLM-generated information for non-medical users, all responses were assessed for public readability. Readability assessments were performed independently by two non-medical reviewers representing the target lay audience. One reviewer was a native English speaker who evaluated the English responses, whereas the second reviewer was a bilingual Turkish-English speaker who evaluated the Turkish responses. Neither reviewer had a formal medical education or healthcare-related professional experience, thereby reflecting the perspective of the typical information-seeking patients and caregivers.

Readability was evaluated using a five-point Likert-type scale based on the overall comprehensibility, the clarity of language, the use of medical terminology, sentence complexity, and the suitability of the information given for a lay audience.

Scores were assigned as follows:

- **5 points:** Excellent readability; easily understood by the general public.
- **4 points:** Good readability with minor use of medical terminology.
- **3 points:** Moderate readability requiring some health literacy.
- **2 points:** Poor readability with substantial technical content.
- **1 point:** Very poor readability; unlikely to be understood by a non-medical audience.

Disease-specific and overall readability scores were calculated separately for the English and Turkish responses.

Statistical Analysis

Categorical variables are presented as frequencies and percentages, whereas continuous variables are reported as mean±standard deviation. Fisher's exact test was used for pairwise comparisons between the 2 models and the 2 languages, while chi-square tests were used to compare concordance outcomes among the disease categories. Readability scores are summarized descriptively and compared between the 2 languages and the 2 models. A two-sided p value <0.05 was considered statistically significant.

Outcomes

The primary outcome was guideline concordance of LLM-generated responses.

The secondary outcomes included:

- The public readability scores,
- A comparison of the English and Turkish responses,
- A comparison between the LLMs,
- The identification of disease-specific and model-specific knowledge gaps,
- The distribution of successful, partially successful, and failed responses.

Importantly, beyond overall concordance rates, the specific questions associated with reduced concordance were compared between the 2 models in order to identify model-specific knowledge-gap profiles.

Results

A total of 50 questions covering five common pediatric penile conditions (phimosis, hypospadias, CPC, paraphimosis, and buried penis) were evaluated for each LLM in both English and Turkish. The responses were assessed for guideline concordance according to the EAU Paediatric Urology Guidelines and for public readability.

Guideline Concordance

ChatGPT

Among the 50 evaluated responses, 40 (80%) were classified as Grade 1 (fully concordant), 9 (18%) as Grade 2 (partially concordant), and 1 (2%) as Grade 3. No Grade 4 responses were identified. Consequently, 49 of 50 responses (98%) were considered acceptable (either Grade 1 or Grade 2) (Table I).

Disease-specific Grade 1 concordance rates were 70% for phimosis, 90% for hypospadias, 80% for CPC, 100% for paraphimosis, and 60% for buried penis. When acceptable responses (Grade 1 or 2) were considered, concordance rates reached 90% for phimosis and 100% for all of the remaining conditions (Figure 1).

No significant differences were observed among the disease categories regarding successful responses (p=0.37) or for the acceptable responses (p=0.91).

Gemini

Among the 50 evaluated responses, 41 (82%) were classified as Grade 1 and 9 (18%) as Grade 2. No Grade 3 or Grade 4 responses were identified. Therefore, all responses (50/50, 100%) were considered acceptable according to the predefined criteria (Table II).

Disease-specific Grade 1 concordance rates were 80% for phimosis, 80% for hypospadias, 80% for CPC, 90% for paraphimosis, and 80% for buried penis. Acceptable response rates (Grade 1 or 2) reached 100% across all disease categories (Figure 2).

No significant differences were detected among the disease categories regarding successful responses (p=0.989) or for the acceptable responses (p=1.000).

Comparison Between Models

Overall guideline concordance was comparable between the two LLMs. Grade 1 concordance rates were 80% for ChatGPT and 82% for the Gemini (p=1.000). Acceptable response rates (Grade 1 or 2) were 98% and 100%, respectively (p=1.000). The frequency of failed responses (Grade 3-4) did not differ significantly between the two models (2% vs. 0%, p=1.000).

Although overall performance was similar, the specific questions associated with reduced concordance differed substantially between the two models. Only a limited overlap was observed in non-Grade 1 responses, suggesting model-specific knowledge-gap profiles rather than disease-specific weaknesses.

Table I. Guideline concordance outcomes according to disease category and language

Disease	Language	Successful (Grade 1) n (%)	Partially successful (Grade 1, 2) n (%)	Failed (Grade 3, 4) n (%)	p value*
Phimosis	English	7/10 (70%)	9/10 (90%)	1/10 (10%)	
	Turkish	7/10 (70%)	9/10 (90%)	1/10 (10%)	
Hypospadias	English	9/10 (90%)	10/10 (100%)	0/10 (0%)	
	Turkish	9/10 (90%)	10/10 (100%)	0/10 (0%)	
CPC	English	8/10 (80%)	10/10 (100%)	0/10 (0%)	
	Turkish	8/10 (80%)	10/10 (100%)	0/10 (0%)	
Paraphimosis	English	10/10 (100%)	10/10 (100%)	0/10 (0%)	
	Turkish	10/10 (100%)	10/10 (100%)	0/10 (0%)	
Buried penis	English	6/10 (60%)	10/10 (100%)	0/10 (0%)	
	Turkish	6/10 (60%)	10/10 (100%)	0/10 (0%)	
Overall	**English**	**40/50 (80%)**	**49/50 (98%)**	**1/50 (2%)**	**1.000†**
	Turkish	**40/50 (80%)**	**49/50 (98%)**	**1/50 (2%)**	
p value (among disease categories)	**English**	**0.37‡**	**0.91§**	-	
	Turkish	**0.37‡**	**0.91§**	-	

*p values compare distributions of guideline concordance outcomes.
†Overall English versus Turkish comparison (Fisher's exact test).
‡Comparison among the disease categories for successful responses (Grade 1) using Fisher-Freeman-Halton exact test.
§Comparison among the disease categories for partially successful responses (Grade 1, 2) using Fisher-Freeman-Halton exact test.
CPC: Congenital penile curvature

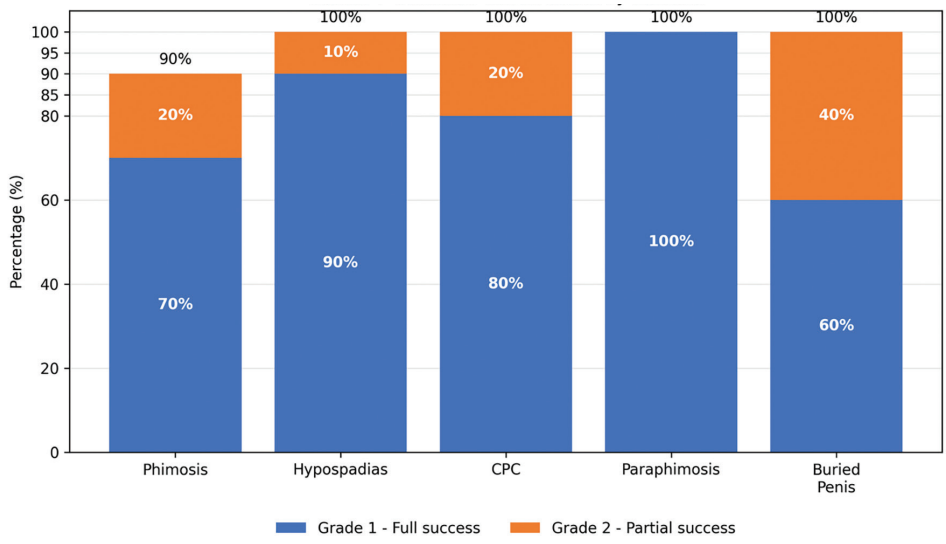

Figure 1. ChatGPT guideline concordance by disease

Table II. Guideline concordance outcomes according to disease category and language for Gemini

Disease	Language	Grade 1 n (%)	Grade 2 n (%)	Grade 3, 4 n (%)	Acceptable responses (Grade 1, 2) n (%)
Phimosis	English	8 (80%)	2 (20%)	0 (0)	10 (100%)
	Turkish	8 (80%)	2 (20%)	0 (0)	10 (100%)
Hypospadias	English	8 (80%)	2 (20%)	0 (0)	10 (100%)
	Turkish	8 (80%)	2 (20%)	0 (0)	10 (100%)
CPC	English	8 (80%)	2 (20%)	0 (0)	10 (100%)
	Turkish	8 (80%)	2 (20%)	0 (0)	10 (100%)
Paraphimosis	English	9 (90%)	1 (10%)	0 (0)	10 (100%)
	Turkish	9 (90%)	1 (10%)	0 (0)	10 (100%)
Buried penis	English	8 (80)	2 (20)	0 (0)	10 (100)
	Turkish	8 (80)	2 (20)	0 (0)	10 (100)
Overall	**English**	**41 (82)**	**9 (18)**	**0 (0)**	**50 (100)**
	Turkish	**41 (82)**	**9 (18)**	**0 (0)**	**50 (100)**
p value (among disease categories)	**English**	**0.989**	-	-	**1.000**
	Turkish	**0.989**	-	-	**1.000**
p value (English vs. Turkish)	-	**1.000**	**1.000**	**1.000**	**1.000**

CPC: Congenital penile curvature

Public Readability

ChatGPT

The mean readability scores were high in both languages. English responses achieved an overall score of 236/250 (4.72±0.48), while Turkish responses achieved 224/250 (4.48±0.66). Although English responses demonstrated numerically higher readability scores, the difference did not reach statistical significance (p=0.081).

Disease-specific readability scores for English responses ranged from 4.6 to 4.8, whereas Turkish scores ranged from 4.2 to 4.7. No significant differences were observed among the disease categories for the English (p=0.812) or the Turkish responses (p=0.091).

Gemini

Gemini also demonstrated high readability in both languages. English responses achieved an overall readability

score of 236/250 (4.72±0.45), while Turkish responses achieved 231/250 (4.62±0.53). No significant difference was observed between the two languages (p=0.365) (Table III)

Disease-specific readability scores ranged from 4.6 to 4.8 for the English responses and from 4.4 to 4.8 for the Turkish responses. No significant variations among the disease categories were detected for the English (p=0.851) or the Turkish responses (p=0.634).

Comparison Between Models

English readability scores were identical between ChatGPT and Gemini (4.72 vs. 4.72, p=1.000). Turkish readability scores were numerically higher for Gemini (4.62 vs. 4.48), although this difference was not statistically significant (p=0.31). Overall readability scores did not differ significantly between the two models (4.60±0.58 vs. 4.67±0.49, p=0.54).

Analysis of Non-Grade 1 Responses

Although overall guideline concordance rates were comparable between ChatGPT and Gemini, the specific questions associated with reduced concordance differed considerably between the two models. ChatGPT deficiencies were mainly related to epidemiological, etiological, and counseling-related topics, whereas Gemini more frequently demonstrated omissions involving diagnostic evaluation, surgical indications, treatment details, and the timing of interventions.

Only two overlapping areas of reduced concordance were identified across the 50 evaluated questions (Phimosis Q4 and Buried Penis Q7). Overall, 15 of 17 non-Grade 1 responses (88.2%) were model-specific, suggesting distinct knowledge-gap profiles despite similar overall performance metrics (Table IV).

Overall Findings

Both LLMs demonstrated high levels of guideline concordance and public readability across all of the evaluated pediatric penile conditions. No statistically significant differences were identified between the English and Turkish responses or between the two models evaluated.

Gemini achieved a slightly higher rate of fully concordant responses and demonstrated no guideline-discordant answers. However, these numerical differences did not reach statistical significance. Importantly, despite similar overall performance metrics, the specific domains associated with reduced guideline concordance differed between the two models, indicating distinct patterns of knowledge limitations.

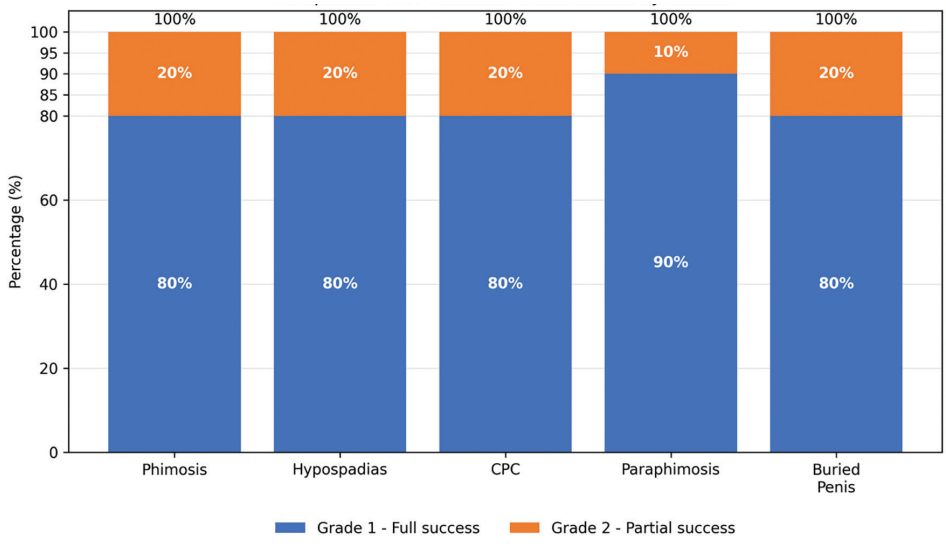

Figure 2. Gemini guideline concordance by disease

Table III. Public readability scores according to disease category and language

Disease	ChatGPT English	ChatGPT Turkish	p value	Gemini English	Gemini Turkish	p value
Phimosis	4.8±0.4	4.7±0.5	0.640	4.8±0.4	4.7±0.5	0.651
Hypospadias	4.8±0.4	4.5±0.7	0.322	4.7±0.5	4.6±0.5	0.681
CPC	4.6±0.5	4.2±0.8	0.170	4.6±0.5	4.4±0.7	0.576
Paraphimosis	4.7±0.5	4.7±0.5	1.000	4.8±0.4	4.8±0.4	1.000
Buried Penis	4.7±0.5	4.3±0.7	0.146	4.7±0.5	4.6±0.5	0.681
Overall	**4.72±0.48**	**4.48±0.66**	**0.081**	**4.72±0.45**	**4.62±0.53**	**0.365**
p value (among disease categories)	**0.812**	**0.091**	-	**0.851**	**0.634**	-

CPC: Congenital penile curvature

Table IV. Comparison of non-Grade 1 responses between ChatGPT and Gemini

Disease	ChatGPT (Grade 2-3 responses)	Gemini (Grade 2 responses)	Overlap
Phimosis	Q2 (Etiology) Q3 (Epidemiology) Q4 (Symptoms)	Q4 (Symptoms) Q7 (Non-surgical treatment)	Q4
Hypospadias	Q7 (Need for surgery)	Q5 (Diagnosis) Q8 (Ideal age for repair)	None
CPC	Q2 (Etiology) Q8 (Sexual function)	Q6 (Surgical indication) Q7 (Timing of treatment)	None
Paraphimosis	None	Q6 (Treatment)	None
Buried penis	Q2 (Etiology) Q3 (Epidemiology) Q7 (Surgical indication) Q8 (Complications)	Q7 (Surgical indication) Q9 (Surgical outcomes)	Q7
Overall	**10 non-Grade responses**	**9 non-Grade 1 responses**	**2 overlapping questions**

CPC: Congenital penile curvature

Discussion

The present study demonstrated that both of the LLMs evaluated provided highly accurate and readable information regarding common pediatric penile conditions. Guideline concordance rates exceeded 98% for ChatGPT and reached 100% for Gemini when acceptable responses (Grade 1 or 2) were considered. Importantly, no Grade 4 responses were identified and only a single Grade 3 response was observed across all of the ChatGPT responses evaluated. These findings suggest that current-generation LLMs are capable of delivering information which is largely consistent with contemporary pediatric urology guidelines.

The performance observed in the present study appears superior to that reported in many early investigations evaluating previous generations of LLMs. Initial studies frequently highlighted concerns regarding hallucinated information, factual inaccuracies, and inconsistent recommendations. As model architectures and training datasets have evolved, however, recent evidence has demonstrated progressive improvements in medical accuracy and clinical reasoning (8). Our findings support this trend and suggest that current LLM versions may provide useful supplementary educational information for both healthcare professionals and the general public. Nevertheless, clinical decisions should continue to rely on professional expertise, guideline-based recommendations, and individualized patient assessments.

Interestingly, despite similar overall concordance rates, the specific domains associated with reduced concordance differed considerably between the two models. Only two overlapping areas of reduced concordance were identified among the fifty evaluated questions. Most non-Grade 1 responses were model-specific, indicating distinct knowledge-gap profiles despite comparable overall performances. This observation suggests that future evaluations should not focus solely on overall accuracy rates but should also investigate the nature and distribution of model-specific deficiencies.

Another notable finding was the consistently high readability of the responses. Both LLMs achieved mean readability scores above 4.5/5 in both English and Turkish, indicating that the information generated was generally understandable to individuals without formal medical training. Although the English responses tended to achieve slightly higher scores, no statistically significant differences were observed between the two languages.

Similarly, no significant language-related differences were identified in guideline concordance. Taken together, these findings suggest that contemporary LLMs can provide relatively consistent health information across different linguistic settings. Such multilingual consistency may be particularly important for those patients and caregivers in non-English-speaking regions where access to specialist pediatric urology resources may be more limited. The ability of modern LLMs to maintain a comparable performance across languages may therefore contribute to reducing language-related disparities in access to health information (9).

The observed variation among the disease categories may partly reflect differences in the amount of the available training data. Conditions such as hypospadias have been extensively studied and discussed within both scientific and public domains for decades. A simple search of the biomedical literature demonstrates that hypospadias has generated substantially more publications than many other pediatric penile conditions (10). Consequently, it is not surprising that the LLMs demonstrated particularly strong performances in this area. In contrast, less frequently discussed conditions may provide fewer opportunities for model training and refinement, potentially contributing to occasional omissions or incomplete responses.

Compared with previous investigations evaluating online health information, our study employed a more standardized methodology. Earlier studies commonly relied on questions collected from search engines, YouTube, Facebook, Google Trends, or publicly available FAQ databases (5,11). While these approaches reflect real-world information-seeking behavior, they may introduce substantial variability in topic selection and question complexity. In order to minimize this variability, we developed a structured question set based on predefined clinical domains, ensuring consistent coverage of definitions, epidemiology, etiology, symptoms, diagnosis, treatment, complications, surgical issues, and long-term outcomes across all of the disease categories.

Study Limitations

Several limitations should be acknowledged. First, LLMs frequently generate information beyond the specific question asked. While this may enhance educational value, it complicates objective scoring because responses often contain content not directly addressed by the evaluation criteria. Second, only those pediatric penile conditions specifically addressed within the EAU Paediatric Urology Guidelines were evaluated. Therefore, our findings may not be generalizable to conditions not covered by these guideline recommendations. Third, readability assessments were performed using a limited number of evaluators and may not fully represent the perceptions of all patient populations. In addition, readability assessment inherently involves a degree of subjectivity. Fourth, only two contemporary LLM platforms were evaluated, and other models may demonstrate different performance characteristics. Finally, this study focused on general educational questions commonly asked by families. Complex clinical scenarios, individualized treatment decisions, and unusual presentations were not assessed. Consequently, these findings support the reliability of LLMs for general patient education but should not be extrapolated to complex clinical decision-making (12). Visual information was not evaluated because corresponding illustrations were not included within the reference guideline used for concordance assessment. Another important limitation is the dynamic nature of LLM development. Model architectures, training datasets, and system updates evolve continuously over time. Therefore, the present findings should be interpreted as reflecting model performance during the study period rather than the permanent characteristics of the evaluated systems.

Taken together, contemporary LLMs appear capable of providing accurate and understandable information regarding common pediatric penile conditions. Although they cannot replace professional medical consultations, they may serve as useful educational resources. Continued reassessment remains necessary as both clinical guidelines and LLM technologies evolve over time (13,14).

Conclusion

Contemporary LLMs demonstrated high guideline concordance and excellent public readability when answering FAQs regarding common pediatric penile conditions. Importantly, no clinically significant misinformation was identified in either of the models evaluated, and most observed deficiencies reflected minor omissions of specific guideline details rather than inaccurate recommendations.

No significant differences were observed between the two models or between the English and Turkish responses. Although overall performance was similar, the specific areas of reduced concordance differed between the two models, suggesting distinct knowledge-gap profiles.

These findings suggest that the current-generation of LLMs may serve as reliable educational resources for families and healthcare professionals. However, they should complement rather than replace professional medical consultations, particularly in complex clinical situations requiring individualized decision-making. Further studies evaluating additional pediatric urological conditions, visual content, and future versions of LLMs are warranted as both clinical guidelines and artificial intelligence technologies continue to evolve over time.

Ethics

Ethics Committee Approval: Ethics committee approval was not required because the study did not involve human participants, animal subjects, or identifiable personal data.

Informed Consent: Not applicable. This study did not involve human participants or patient data. The analyses were based on responses generated by large language models to standardized guideline-based questions; therefore, informed consent was not required.

Acknowledgements

The authors thank Dr. Emre Bülbül for performing the statistical analyses of this study. The authors also thank Ivan Pointon, a native English-speaking non-medical reviewer, and Onur Can Türedi, a bilingual Turkish-English non-medical reviewer, for their valuable contributions to the public readability assessment of the LLM-generated responses.

Footnotes

Authorship Contributions

Concept: B.T., A.S., Design: B.T., A.S., Data Collection or Processing: B.T., A.S., Analysis or Interpretation: B.T., A.S., Literature Search: B.T., A.S., Writing: B.T., A.S.

Conflict of Interest: The authors declare no conflicts of interest.

Financial Disclosure: The authors received no financial support for the conduct, authorship, or publication of this study.

References

1. Ratnagandhi JA, Godavarthy P, Gnaneswaran M, Lim B, Vittalraj R. Enhancing anesthetic patient education through the utilization of large language models for improved communication and understanding. Anesth Res. 2025; 2:4.
2. Tulgar S, Aksu C, Selvi O, et al. A comparative evaluation of the quality of responses provided by different large language model chatbots to frequently asked questions regarding nerve blocks. BMC Anesthesiol. 2026; 26:98.
3. Green Z, Ashton JJ, Beattie RM. Learning from the past, structuring the future: using large language models to unlock a century of paediatric research in Archives of Disease in Childhood. Arch Dis Child. 2026; 111:418-22.
4. Kocaoğlu MH, Demirel A, Kaya İ. Accuracy, quality, and readability analyses of responses from large language models to questions on pediatric dental sedation. BMC Oral Health. 2026; 26:492.
5. Caglar U, Yildiz O, Meric A, et al. Evaluating the performance of ChatGPT in answering questions related to pediatric urology. J Pediatr Urol. 2024; 20:26.e1-5.
6. Spinoit AF, Waterschoot M, Sinatti C, et al. Fertility and sexuality issues in congenital lifelong urology patients: male aspects. World J Urol. 2021; 39:1013-9.
7. Radmayr C, van Uitert A, Kennedy U, et al. EAU Guidelines on Paediatric Urology. In: EAU Guidelines. Arnhem, The Netherlands, European Association of Urology, 2026.
8. Cheng Y, Zhu L. A review of ChatGPT in medical education: exploring advantages and limitations. Int J Surg. 2025; 111:4586-602.
9. Tan J, Wang L, Wang G, et al. Safety and user perception of general-purpose large language models in pediatric healthcare: evaluations of ChatGPT by doctors and parents. Digit Health. 2026; 12:20552076261427505.
10. Faraj S, Clermidi P, Irtan S, Madec FX. Artificial intelligence chatbots vs. YPUC pediatric urologists: performance on a Campbell Walsh urology hypospadiology questionnaire. World J Urol. 2025; 43:719.
11. Gumus K, Yilmaz AB, Cinbek AF, Ozdemir HB, Kutman KG. Evaluating AI chatbots in enuresis nocturna information: a comparative analysis of readability, reliability and quality. Fr J Urol. 2026; 36:103062.
12. Adriansyah IA, Wahyudi I, Vallasciani S, et al. Evaluating the utility of ChatGPT in enhancing parental education and clinical support in hypospadias care. J Pediatr Urol. 2025; 21:639-43.
13. MacNevin W, Dawe N, Harkness L, Salman B, Keefe DT. Evaluation of ChatGPT's performance on answering pediatric urology questions based on association guidelines. Can Urol Assoc J. 2025; 19:E362-7.
14. Kucuker K, Akinci A, Duran MB, et al. Awareness, use, and perceived barriers to artificial intelligence in pediatric urology: a multicenter survey. Ther Adv Urol. 2026; 18:17562872261422939.

J Pediatr Res 2026;13(3):303-13
DOI: 10.4274/jpr.galenos.2026.78871

Original Article

Can a Large Language Model Effectively Answer Parents' Questions About Children's Oral Health?

Irem Nur Bitisik[1], Selcuk Savas[1], Ilgin Timarci[2], Mehmet Izgi[3], Ebru Kucukyilmaz Izgi[1]

[1]Izmir Katip Celebi University Faculty of Dentistry, Department of Pediatric Dentistry, Izmir, Türkiye
[2]Izmir Katip Celebi University Faculty of Medicine, Department of Public Health, Izmir, Türkiye
[3]Private Practice, Izmir, Türkiye

ABSTRACT

Aim: This study aimed to evaluate the readability, quality, and temporal consistency of Turkish responses generated by ChatGPT to frequently asked questions in pediatric dentistry.

Materials and Methods: Twenty open-ended questions were adapted from the frequently asked questions published on the American Academy of Pediatric Dentistry parent information page. Each question was submitted to ChatGPT (GPT-4, OpenAI, USA) in Turkish over seven consecutive days, three times per day, using a new chat session for each interaction, standardized prompts and default system settings. A total of 420 responses were analyzed. Readability was assessed using both the Ateşman and Çetinkaya-Uzun readability formulas, while response quality was evaluated using the Global Quality Score (GQS). Inter-day and intra-day comparisons were performed using repeated-measures ANOVA or Friedman tests according to the data distribution. Intraclass Correlation Coefficient (ICC) analysis was used to evaluate temporal consistency (α=0.05).

Results: A significant intra-day variability in the readability scores was observed during the first four days for the Ateşman and Çetinkaya-Uzun formulas ($p<0.05$), whereas no significant differences were identified during Days 5-7. Between-day comparisons revealed statistically significant variations in readability scores for both formulations ($p<0.001$). Ateşman's scores were found to correspond to "moderate difficulty" and "easy" readability levels, while Çetinkaya-Uzun's scores indicated an "instructional reading level". By contrast, no significant intra-day or between-day differences were observed in the GQS scores ($p=0.650$), and overall quality scores remained consistently high throughout the study period. The ICC analyses demonstrated good-to-excellent consistency for readability and GQS measurements, with overall ICC values ranging between 0.906 and 0.961.

Conclusion: The Turkish responses generated by ChatGPT to frequently asked questions in pediatric dentistry demonstrated moderate readability, high information quality, and good-to-excellent temporal consistency. However, due to the variability in readability scores across repeated assessments and over time, Large Language Models (LLMS) may serve as supportive tools for patient and parent education, however, human oversight remains necessary for the interpretation and clinical use of LLM-generated health information.

Keywords: Large language model, pediatric dentistry, readability, patient education as topic

Introduction

Large Language Models (LLM) represent a significant innovation in technology, with the objective of emulating cognitive processes which are characteristic of human intelligence through the utilization of computer systems. This development has precipitated substantial

Corresponding Author
Prof. Ebru Kucukyilmaz Izgi, Izmir Katip Celebi University Faculty of Dentistry, Department of Pediatric Dentistry, Izmir, Türkiye
E-mail: ebrukucukyilmaz@hotmail.com **ORCID:** orcid.org/0000-0002-6086-7410
Received: 08.06.2026 **Accepted:** 28.07.2026 **Epub:** 03.09.2026 **Publication Date:** 18.09.2026

Cite this article as: Bitisik IN, Savas S, Timarci I, Izgi M, Kucukyilmaz Izgi E. Can a large language model effectively answer parents' questions about children's oral health?. J Pediatr Res. 2026;13(3):303-13

transformations across numerous domains, with particular relevance to the field of healthcare in recent years (1). The increasing data capacity, more advanced algorithms, and developments in LLMs have resulted in the wider use of AI applications in the early diagnosis of diseases, clinical decision support systems, treatment planning, and patient information processes (2). These developments provide rapid and straightforward access to information and influence the decision-making processes of patients and their relatives regarding health-related issues (2).

Pediatric dentistry is a specialty in which preventive approaches and parental education play a decisive role in treatment success (3). In this context, parents' access to accurate, understandable, and accessible information is of great importance for clinical outcomes. Today, internet-based information-seeking behaviors have changed with the widespread use of LLMs, and users have begun to prefer platforms which provide direct answers to their questions rather than relying solely on traditional search engines (4,5). However, the readability and consistency of the information provided through these platforms have become critical evaluation criteria, particularly in the field of healthcare.

In recent years, studies investigating the capacity of ChatGPT to generate health-related information have increased in the literature (6-10). Nevertheless, when the current literature is examined, it is evident that most studies have focused largely on English-language content, while research addressing Turkish-language health-related content, particularly in the field of pediatric dentistry, remains limited (11,12). Although recent Turkish studies have evaluated LLMs' performance for general parental inquiries as well as specific pediatric dental contexts such as fluoride education, sedation-related questions, pulpotomy, and molar incisor hypomineralization, these studies have primarily focused on response accuracy or single-timepoint performance (13-17). Consequently, important aspects of LLM-generated patient education, including temporal consistency across repeated interactions, changes in readability over repeated queries, and the combined evaluation of readability, information quality, and consistency, remain insufficiently investigated in Turkish-language pediatric dentistry. Considering the effect of language structure on readability, the differences in Turkish in terms of word and syllable structure necessitate a separate evaluation of the comprehensibility of texts produced in this language. Therefore, evaluating not only readability but also the informational quality and temporal consistency of LLM-generated Turkish health information has become an important research priority.

A comprehensive evaluation of LLM-generated health information requires the use of validated and language-appropriate assessment tools. For Turkish texts, the Ateşman and Çetinkaya-Uzun readability formulas are among the most widely used readability indices (18,19). While the Ateşman formula classifies texts according to their overall reading difficulty, the Çetinkaya-Uzun formula estimates the educational level required for comprehension, thereby providing complementary perspectives on textual readability. In addition to readability, the Global Quality Score (GQS) is a validated 5-point Likert-type scale widely used to evaluate the overall quality, educational value, and usefulness of patient-oriented health information, including LLM-generated content (20). Together, these complementary instruments enable a multidimensional assessment of LLM-generated health information by evaluating not only its linguistic accessibility, but also its educational quality and consistency. In order to address the aforementioned knowledge gaps, the present study adopted a repeated-measures design in order to comprehensively evaluate the readability, informational quality, and temporal consistency of ChatGPT-generated responses using validated Turkish readability formulas, the GQS, and Intraclass Correlation Coefficient (ICC) analyses. Accordingly, the present study aimed to evaluate the readability, overall quality, and consistency of ChatGPT's responses to frequently asked questions by patients and parents in the field of pediatric dentistry. The null hypotheses tested in this study were determined as follows: (i) The Ateşman readability scores of the responses generated by ChatGPT do not show statistically significant differences between different days and repetitions. (ii) The Çetinkaya-Uzun readability scores of the responses generated by ChatGPT do not show statistically significant differences between different days and repetitions. (iii) The GQS scores of the responses generated by ChatGPT do not show statistically significant differences between different days and repetitions. (iv) The responses generated by ChatGPT do not demonstrate a high level of consistency across different time periods.

Materials and Methods

Study Design

The present study was designed as an observational and descriptive investigation with the objective of evaluating the overall quality, consistency, and readability of responses generated by a LLM (ChatGPT Plus, GPT-4o, OpenAI, USA) to frequently asked questions by patients in the field of pediatric dentistry. The LLM was accessed on 10th of June,

2025 using the default system settings, with no additional prompting or parameter modification.

Question Preparation

In this study, 20 open-ended questions were prepared regarding topics frequently asked by patients and parents in the field of pediatric dentistry. Rather than constituting a novel questionnaire or psychometric instrument, the question set was designed to reflect common parental concerns regarding children's oral health. The questions employed in this study were adapted to the scope of the research based on the frequently asked questions published on the parent information page of the American Academy of Pediatric Dentistry (21). The final set of 20 questions was subsequently reviewed by two experienced pediatric dentists in order to ensure clinical relevance, clarity, and appropriateness for Turkish-speaking parents before data collection. Any differences in wording or content were resolved through discussion until consensus was reached before the data collection (Supplementary File 1). As no accepted methodology exists for determining the number of questions in LLM evaluation studies, the present sample size of 20 questions was selected in accordance with comparable pediatric dentistry studies using parent-oriented, question-based ChatGPT evaluation designs (13,14,17).

Data Collection

The finalized 20 questions were posed to ChatGPT in Turkish on three occasions per day for a duration of seven days. The timing of these data collections was as follows: in the morning between 08:00-10:00, in the afternoon between 12:00-14:00, and in the evening between 18:00-20:00. All questions were entered into the system manually in a standardized format. Responses were generated using the ChatGPT web interface under default system settings, with no custom instructions, system prompts, or parameter modifications. Each interaction was conducted in a new chat session. Before every measurement, a new conversation within the same user account was initiated in order to prevent carry-over effects from previous interactions and to ensure the independence of each generated response. The responses were evaluated by an assistant professor (10 years of clinical experience) and a professor (15 years of clinical experience), both specialized in pediatric dentistry. Any discrepancies in the assessments were resolved through discussion until a consensus was reached.

Ateşman Readability Formula

The Ateşman Readability Formula, developed by Ateşman in 1997 as an adaptation of the Flesch Reading Ease Formula into Turkish, was used to evaluate the readability level of the texts (18). According to the Ateşman readability formula, readability levels were classified as very difficult, difficult, moderately difficult, easy or very easy (18).

Çetinkaya-Uzun Readability Formula

This formula, developed to evaluate the readability of Turkish texts, is based on average word length and sentence length (19). The readability formula was calculated and readability levels were classified as frustration reading level (10^{th}-12^{th} grade), instructional reading level (8^{th}-9^{th} grade) or independent reading level (5^{th}-7^{th} grade) (19).

GQS Scale

The overall quality level of the responses was evaluated using the GQS, a 5-point Likert-type scale (20). The GQS scale levels were classified as excellent quality, low quality, moderate quality or high/excellent quality (20).

Statistical Analysis

The statistical analyses were conducted utilizing IBM SPSS Statistics Version 23.0 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean±standard deviation or median (Q1-Q3), while the categorical variables are presented as n (%). In order to assess differences between repeated measurements, repeated measures ANOVA was applied when the normal distribution assumptions were met. In instances where the assumption of normality was not met, the Friedman test was employed, and if statistical significance was detected, pairwise comparisons with Bonferroni correction were conducted. For GQS scores, the assumptions for repeated-measures ANOVA were not satisfied; therefore, all within-day comparisons were performed using the Friedman test.

In order to evaluate the consistency of responses over time, the ICC was calculated. The analysis was conducted in accordance with the two-way mixed effects and absolute agreement model [ICC (3,k)] (22). Inter-rater agreement was determined by calculating Cohen's Kappa coefficient (23). In all analyses, the level of statistical significance was accepted as $p<0.05$.

Results

The results are presented according to the predefined outcome measures, including readability, overall quality, and temporal consistency of the ChatGPT-generated responses.

The findings obtained using the Ateşman readability formula are presented in Table I. A significant difference was observed among repeated assessments within Days

1-4 (p=0.017, p=0.016, p=0.003, and p=0.018, respectively), whereas no significant intra-day differences were detected for Days 5-7 (p=0.937, p=0.754, and p=0.222, respectively). Furthermore, between-day comparisons revealed statistically significant differences in Ateşman readability scores (p<0.001). The highest mean readability score was observed on Day 2, Repetition 1 (79.2±15.1), whereas the lowest mean score was recorded on Day 7, Repetition 2 (60.3 ±15.6). Despite these statistically significant differences, the readability scores remained within the same readability category throughout the study period, indicating that the observed variations were not associated with a meaningful change in the practical readability of the generated responses.

The results obtained using the Çetinkaya-Uzun readability formula are presented in Table II. A significant intra-day variability was observed among repeated assessments on Days 1-4 (p=0.022, p=0.020, p=0.003, and p=0.009, respectively), while no significant differences were detected on Days 5-7 (p=0.887, p=0.525, and p=0.399, respectively). Furthermore, between-day comparisons revealed statistically significant differences in readability scores (p<0.001). Similarly, although statistically significant differences were observed, these variations did not result in meaningful changes in the corresponding educational reading levels, suggesting that the practical comprehensibility of the responses remained largely stable over time.

Table I. Comparison of Ateşman readability scores between days and repeated assessments within each day

Day	Mean ± SD	Median (Q1-Q3)	Repetition	Mean ± SD	Median (Q1-Q3)	p value
Day 1	69.4±16.2^{AB}	72.3 (58.4-81.6)	Repetition 1	64.4±14.9^{a}	62.8 (56.3-78.3)	p=0.017$^{£}$
			Repetition 2	75.7±14.8^{b}	72.1 (68.7-81.7)	
			Repetition 3	68.1±17.2^{ab}	73.6 (64.8-79.2)	
Day 2	73.7±18.2^{BC}	75.1 (61.7-86.8)	Repetition 1	79.2±15.1^{a}	77.8 (71.3-85)	p=0.016$^{¥}$
			Repetition 2	69.4±20^{b}	69.9 (57.2-83.2)	
			Repetition 3	72.4±18.5^{ab}	72.4 (64.6-79.7)	
Day 3	68.9±16.8^{AB}	70.4 (57.1-80.3)	Repetition 1	71.3±12.2^{a}	71.5 (61.4-81.4)	p=0.003$^{¥}$
			Repetition 2	61.1±15.5^{b}	62.7 (49.1-70.4)	
			Repetition 3	74.3±19.6^{ab}	77.9 (61.9-84.9)	
Day 4	72.3±14.6^{BC}	73.5 (61.2-82.4)	Repetition 1	69.8±15^{a}	68.2 (59-81.4)	p=0.018$^{¥}$
			Repetition 2	76.2±16.9^{b}	73.8 (63.8-94.3)	
			Repetition 3	70.7±11.1^{ab}	70.1 (63.2-76.2)	
Day 5	71.1±14.2^{B}	72.1 (60.8-80.1)	Repetition 1	71.6±16.3^{a}	76.5 (62.7-81.5)	p=0.937$^{¥}$
			Repetition 2	70.8±10.9^{a}	71.6 (62.7-79.8)	
			Repetition 3	70.6±15.4^{a}	69.9 (59-78.6)	
Day 6	75.3±14.2^{C}	77.2 (65.1-87.9)	Repetition 1	76.1±12^{a}	77.6 (69.6-83.6)	p=0.754$^{¥}$
			Repetition 2	75.8±17.7^{a}	75.9 (60.5-88.1)	
			Repetition 3	73.9±12.7^{a}	74 (65.2-82.3)	
Day 7	63.3±15.4^{A}	64.8 (51.6-73.9)	Repetition 1	65.9±14.4^{a}	64.8 (55.6-74)	p=0.222$^{¥}$
			Repetition 2	60.3±15.6^{a}	57.3 (47.4-70.2)	
			Repetition 3	63.6±16.2^{a}	65.1 (56-72.9)	
Between-day comparison p value	<0.001*$^{£}$		Between 21 repeated assessments p value	<0.001*$^{£}$		

Data are presented as mean ± SD and median (Q1-Q3)
$^{£}$: Friedman test, $^{¥}$: Repeated measures ANOVA
*Comparisons between repetitions within each day were performed using repeated-measures ANOVA for normally distributed data and the Friedman test for non-normally distributed data
Different lowercase superscript letters ($^{a, b}$) indicate statistically significant differences between repeated assessments within the same day (p<0.05)
Different uppercase superscript letters ($^{A, B, C}$) indicate statistically significant differences between days (p<0.05)
SD: Standard deviation; Q1: First quartile; Q3: Third quartile

The results of the GQS evaluation are presented in Table III. No statistically significant differences were found between repetitions within the same day for any of the days analyzed (respectively: p=0.497; p=0.595; p=0.368; p=0.692; p=0.311; p=0.174; p=0.054). In the analysis in which all repetitions were evaluated together (1st-21st repetitions), no significant difference was detected (p=0.650).

The ICC analysis results for the Ateşman readability scores are presented in Table IV. Reliability ranging from moderate to good was observed in the within-day repeated measurements (ICC=0.588-0.883). In the analysis evaluating all repetitions together, the ICC value was 0.958 [95% confidence interval (CI): 0.925-0.980]. All ICC values were found to be statistically significant (p<0.001).

The ICC analysis results for the Çetinkaya-Uzun readability scores are presented in Table V. Moderate, good, and excellent levels of reliability were obtained in the within-day repeated measurements (ICC=0.650-0.919). In the analysis evaluating all measurements together, the ICC value was 0.961 (95% CI: 0.932-0.982). All ICC values obtained in the analyses were statistically significant (p<0.001).

The ICC analysis results for the GQS scores are presented in Table VI. In the analysis evaluating all measurements together, the ICC value was 0.919 (95% CI: 0.919-0.979). All ICC analyses were found to be statistically significant (p<0.001).

Table II. Comparison of Çetinkaya-Uzun readability scores between days and repeated assessments within each day

Day	Mean ± SD	Median (Q1-Q3)	Repetition	Mean ± SD	Median (Q1-Q3)	p value
Day 1	41.6±9.4^{AB}	43.1 (35.2-47.9)	Repetition 1	39.2±8.1^{a}	39.7 (32.5-46.6)	p=0.022$^{£}$
			Repetition 2	45±8.5^{b}	42.4 (40.1-48.4)	
			Repetition 3	40.6±10.7^{ab}	43.9 (38.6-47.2)	
Day 2	44.1±11.1^{BC}	45.2 (37.4-50.8)	Repetition 1	47.9±9.8^{a}	45.4 (42.5-50.7)	p=0.020$^{¥}$
			Repetition 2	47.1±12.3^{b}	42.8 (32.8-48.1)	
			Repetition 3	43.6±10.8^{ab}	43.9 (37.4-47.8)	
Day 3	41.1±9.8^{AB}	41.8 (34.7-47.1)	Repetition 1	42.1±7.1^{a}	41.5 (38.9-47.7)	p=0.003$^{¥}$
			Repetition 2	37.3±9.4^{b}	36.2 (31.3-42.8)	
			Repetition 3	43.9±11.6^{ab}	45 (35.8-49.6)	
Day 4	43.1±9.2^{BC}	43.9 (36.9-48.5)	Repetition 1	42.2±9.2^{a}	40.5 (34.3-50.4)	p=0.009$^{¥}$
			Repetition 2	45.4±10.6^{b}	42.7 (38.1-55.3)	
			Repetition 3	41.6±7.4^{a}	39.8 (35.6-46.6)	
Day 5	42.6±8.2^{B}	42.7 (36.4-47.6)	Repetition 1	42.9±8.7^{a}	43.7 (37.2-48.6)	p=0.887$^{¥}$
			Repetition 2	42.1±6^{a}	41.4 (37.2-48.4)	
			Repetition 3	42.8±9.7^{a}	40.7 (36.3-46.4)	
Day 6	44.8±8.8^{C}	46.3 (39.1-51.4)	Repetition 1	45.3±7.6^{a}	45.3 (41.3-49.3)	p=0.525$^{¥}$
			Repetition 2	45.6±10.8^{a}	44.6 (38.4-53.6)	
			Repetition 3	43.4±7.9^{a}	42.7 (37.7-48.1)	
Day 7	38.1±9.4^{A}	38.5 (31.7-43.8)	Repetition 1	39±9.1^{a}	37.2 (31-44.3)	p=0.399$^{¥}$
			Repetition 2	36.6±9.5^{a}	36 (31.2-40.8)	
			Repetition 3	38.5±9.7^{a}	39.2 (34.7-43.8)	
Between-day comparison p value	<0.001*$^{£}$		Between 21 repeated assessments p value	<0.001*$^{£}$		

Data are presented as mean ± SD and median (Q1-Q3)
$^{£}$: Friedman test, $^{¥}$: Repeated measures ANOVA
*Comparisons between repetitions within each day were performed using repeated-measures ANOVA for normally distributed data and the Friedman test for non-normally distributed data
Different lowercase superscript letters ($^{a, b}$) indicate statistically significant differences between repeated assessments within the same day (p<0.05)
Different uppercase superscript letters ($^{A, B, C}$) indicate statistically significant differences between days (p<0.05)
SD: Standard deviation; Q1: First quartile; Q3: Third quartile

Table III. Comparison of Global Quality Score (GQS) scores according to days and repeated assessments within each day

Day	Repetition	Mean ± SD	Median (Q1-Q3)	p value
Day 1	Repetition 1	4.4±0.59	4 (4-5)	p=0.497
	Repetition 2	4.2±0.77	4 (4-5)	
	Repetition 3	4.3±0.55	4 (4-5)	
Day 2	Repetition 1	4.3±0.64	4 (4-5)	p=0.595
	Repetition 2	4.2±0.89	4 (4-5)	
	Repetition 3	4.1±0.97	4 (4-5)	
Day 3	Repetition 1	4.2±0.67	4 (4-5)	p=0.368
	Repetition 2	4.3±0.57	4 (4-5)	
	Repetition 3	4.3±0.57	4 (4-5)	
Day 4	Repetition 1	4.2±0.99	4 (4-5)	p=0.692
	Repetition 2	4.3±0.57	4 (4-5)	
	Repetition 3	4.2±0.70	4 (4-5)	
Day 5	Repetition 1	4.5±0.51	4 (4-5)	p=0.311
	Repetition 2	4.3±0.66	4 (4-5)	
	Repetition 3	4.3±0.57	4 (4-5)	
Day 6	Repetition 1	4.3±0.47	4 (4-5)	p=0.174
	Repetition 2	4.2±0.88	4 (4-5)	
	Repetition 3	4.4±0.50	4 (4-5)	
Day 7	Repetition 1	4.0±1.05	4 (4-5)	p=0.054
	Repetition 2	4.3±0.79	4 (4-5)	
	Repetition 3	4.2±0.75	4 (4-5)	
Between-day comparison	-	-	-	p=0.650

Data are presented as mean ± SD and median (Q1-Q3)
Within-day comparisons were performed using the Friedman test
SD: Standard deviation, Q1: First quartile, Q3: Third quartile

Table IV. ICC reliability analysis of Ateşman readability scores

Repetitions	ICC	95% CI	Interpretation	p value
A1-A2-A3	0.736	0.448-0.887	Moderate	p<0.001
A4-A5-A6	0.836	0.649-0.930	Good	p<0.001
A7-A8-A9	0.733	0.421-0.887	Moderate	p<0.001
A10-A11-A12	0.883	0.746-0.950	Good	p<0.001
A13-A14-A15	0.812	0.600-0.920	Good	p<0.001
A16-A17-A18	0.588	0.117-0.826	Moderate	p<0.001
A19-A20-A21	0.802	0.589-0.915	Good	p<0.001
A1-A21	0.958	0.925-0.980	Excellent	p<0.001

ICC values were calculated using a two-way mixed-effects model with absolute agreement [ICC (3,k)]
A1-A21 represent sequential Ateşman readability measurements across all repetitions and days
ICC: Intraclass Correlation Coefficient, CI: Confidence interval

Table V. ICC reliability analysis of Çetinkaya-Uzun readability scores

Repetitions	ICC	95% CI	Interpretation	p value
C1-C2-C3	0.699	0.386-0.869	Moderate	p<0.001
C4-C5-C6	0.870	0.722-0.945	Good	p<0.001
C7-C8-C9	0.797	0.555-0.915	Good	p<0.001
C10-C11-C12	0.919	0.818-0.966	Excellent	p<0.001
C13-C14-C15	0.799	0.574-0.915	Good	p<0.001
C16-C17-C18	0.650	0.261-0.851	Moderate	p<0.001
C19-C20-C21	0.822	0.629-0.924	Good	p<0.001
C1-C21	0.961	0.932-0.982	Excellent	p<0.001

ICC values were calculated using a two-way mixed-effects model with absolute agreement [ICC (3,k)]
C1-C21 represent sequential Çetinkaya-Uzun readability score measurements across all days and repetitions
ICC: Intraclass Correlation Coefficient, CI: Confidence interval

Table VI. ICC reliability analysis of GQS scores

Repetitions	ICC	95% CI	Interpretation	p value
G1-G2-G3	0.839	0.663-0.931	Good	p<0.001
G4-G5-G6	0.596	0.137-0.829	Moderate	p<0.001
G7-G8-G9	0.867	0.722-0.943	Good	p<0.001
G10-G11-G12	0.783	0.543-0.908	Good	p<0.001
G13-G14-G15	0.791	0.564-0.911	Good	p<0.001
G16-G17-G18	0.767	0.517-0.900	Good	p<0.001
G19-G20-G21	0.906	0.803-0.960	Excellent	p<0.001
G1-G21	0.919	0.919-0.979	Excellent	p<0.001

ICC values were calculated using a two-way mixed-effects model with absolute agreement [ICC (3,k)]
G1-G21 represent sequential GQS measurements across all days and repetitions
ICC: Intraclass Correlation Coefficient, CI: Confidence interval, GQS: Global Quality Scale

Discussion

The findings of the present study demonstrated that ChatGPT-generated responses to frequently asked question in pediatric dentistry generally exhibited a moderate level of readability, high quality scores, and overall good consistency between repetitions. However, variability in readability scores was observed both within the same day and across different days. Accordingly, the first and second null hypotheses, proposing that the Ateşman and Çetinkaya-Uzun readability scores did not differ significantly across different days and repetitions, were rejected. The third null hypothesis, which proposed that GQS scores did not demonstrate significant differences over time, was accepted, whereas the fourth null hypothesis, which proposed that ChatGPT responses did not demonstrate a high level of consistency over time, was rejected. Collectively, these results suggest that although ChatGPT maintains a relatively stable level of informational quality over time, the linguistic complexity of its responses may fluctuate depending on when the same prompt is submitted.

According to the findings obtained in the present study, the readability results demonstrated that the Ateşman readability scores were generally within the range of 60-80, indicating that the responses were between the "moderately difficult" and "easy" reading levels. Similarly, the Çetinkaya-Uzun scores ranging between 35 and 47 suggested that the contents could be classified within the "instructional reading level" category. These results indicate that the responses are understandable for a certain user population. However, the "instructional reading level" identified by the Çetinkaya-Uzun formula approximately corresponds to an eighth- to ninth-grade educational level, suggesting that individuals with limited educational attainment or lower health literacy may still experience difficulties in fully comprehending or applying the information provided. Consequently, readability should be considered as an important determinant of the practical usefulness of LLM-

generated patient education materials, in addition to their factual accuracy and informational quality.

Previous studies have similarly reported that the readability levels of health-related content generated by LLMs is generally acceptable, although substantial variation exists depending on the clinical topic, language, and assessment method employed (24-27). In a study conducted by Aypar Akbağ (26), the mean Ateşman readability score of content related to gestational diabetes generated by ChatGPT and Gemini was reported as 77.8. The researchers stated that these contents had a high level of readability and were considered understandable for users (26). Similarly, in a study conducted by Boztaş Demir and Görgülü (27), the Turkish responses generated by ChatGPT and Gemini to patient questions in the field of orthodontics were reported to be generally adequate in terms of accuracy, comprehensiveness, and readability. Erdat et al. (25) also reported that ChatGPT's responses to questions related to colorectal cancer demonstrated moderate readability and good quality characteristics. Overall, the findings of the present study are consistent with this growing body of evidence indicating that LLM-generated health information is generally understandable and of acceptable educational quality. Nevertheless, direct comparison across studies should be interpreted cautiously as the evaluated clinical topics, response languages, and readability formulas differ considerably.

When the literature specific to pediatric dentistry is considered, an interesting pattern emerges. Although direct numerical comparison is limited by differences in language and readability indices, Kocaoğlu et al. (15) and Aydın Varol et al. (16) similarly reported that ChatGPT-generated responses addressing procedural or guideline-based pediatric dental topics in English generally required university-level reading ability. By contrast, the parental frequently asked questions evaluated in the present study demonstrated moderate-to-easy readability. This discrepancy is likely attributable to the inherently greater linguistic complexity of procedure-oriented subjects, such as pediatric dental sedation and pulpotomy, compared with preventive oral health information intended for parents, rather than reflecting a true language-related difference. These observations further emphasize that the readability of LLM-generated responses depends not only on the characteristics of the language model itself, but also on the complexity and intended audience of the clinical topic being addressed.

One of the notable findings of the present study was that readability scores varied significantly despite the use of identical prompts. Significant differences were observed both between repeated assessments performed on the same day and between different days, indicating that the linguistic characteristics of ChatGPT-generated responses are not entirely stable over time. This situation may be explained by the probabilistic nature of LLMs. LLMs do not always generate the same output for the same input and may demonstrate variability in characteristics such as sentence structure, word choice, and text length. As readability indices are inherently sensitive to these textual characteristics, fluctuations in readability scores are expected even when the underlying informational content remains largely unchanged. These findings suggest that readability assessments performed at a single time point may not adequately represent the linguistic variability of LLM-generated health information. Importantly, the temporal fluctuations observed in the present study provide support for concerns regarding potential model drift and response variability which have previously been discussed in the pediatric dentistry literature. Kocaoğlu et al. (15), for example, identified their cross-sectional design as an important limitation and recommended repeated longitudinal evaluations in order to determine whether chatbot performance changes over time. Similarly, Aydın Varol et al. (16) acknowledged that the temporal consistency of individual chatbots could not be assessed within their study design, while Uçar Gündoğar and Sarıoğlu (17) limited repeated evaluations to a single day, precluding an assessment of between-day variability. By employing a seven-day, three-times-daily repeated-measures protocol, the present study directly addresses this methodological gap and demonstrated that although the overall informational quality of ChatGPT responses remains remarkably stable, their linguistic complexity may fluctuate across repeated interactions. These findings indicate that temporal reproducibility constitutes an independent dimension of LLM performance which cannot be adequately captured by conventional single-timepoint evaluations.

When the quality evaluation was examined, ChatGPT responses demonstrated a consistent performance with high GQS scores across all time periods, and no significant differences were observed between repetitions. Similarly, previous studies have reported that the responses of LLMs in the healthcare field are generally of moderate-to-high quality (28,29). In a systematic review published in 2024, Li et al. (29) reported that ChatGPT mostly demonstrated

"passing" or '"moderate" performance across different healthcare applications, although it was not yet considered fully reliable for clinical use. In another systematic review, ChatGPT was reported to have a broad range of applications in healthcare education, research, and clinical practice, offering significant advantages in areas such as improving writing quality, data analysis, and patient education (28). That same study also emphasized that despite high efficiency and explanatory capacity in content generation, risks such as inaccuracies, bias, and "hallucinations" remain important concerns. The phenomenon of "hallucinations", defined as the ability of LLMs to generate different or inaccurate responses to the same input, has been highlighted as a potential risk for users, particularly in the healthcare field (28,29). Therefore, despite the high-quality scores observed, such systems should not be regarded as definitive standalone sources of information. Recent discussions regarding the design philosophy of LLMs further reinforce these concerns. Contemporary LLMs are generally optimized to generate a response to user prompts rather than explicitly indicating uncertainty or declining to answer. Consequently, inaccurate or unsupported information may occasionally be presented in a fluent and confident manner, making it difficult for users to distinguish between reliable and erroneous content. In the context of pediatric dentistry, where parents may rely on online information to make decisions regarding their children's oral health, this behavior highlights the importance of professional verification and reinforces that LLM-generated information should complement, but not replace, clinical judgment.

The stability of GQS scores observed in the present study parallels findings reported in recent pediatric dentistry investigations. Both Karamüftüoğlu et al. (14) and Aydın Varol et al. (16) similarly observed no significant differences in GQS values, although their comparisons were performed between different chatbot platforms rather than across time. Considered together with the present findings, these studies suggest that GQS may represent a relatively robust measure of the overall educational quality of LLM-generated content, remaining comparatively stable regardless of whether the source of variation is model selection or repeated prompting over time. Furthermore, Kocaoğlu et al. (15) demonstrated no significant association between DISCERN quality scores and Flesch-Kincaid Grade Level readability scores across multiple chatbots, supporting the concept that readability and informational quality constitute distinct dimensions of LLM-generated health information. The present study further strengthens this interpretation by demonstrating that readability may fluctuate substantially over time while overall quality remains stable, indicating that linguistic accessibility and informational quality should be evaluated independently rather than being considered interchangeable indicators of content performance.

Nevertheless, consistently high GQS scores should not be interpreted as evidence of complete factual reliability. Recent pediatric dentistry studies have demonstrated that even LLM-generated responses rated highly for quality may contain fabricated or inaccurate references. Uçar Gündoğar and Sarıoğlu (17) reported that although DeepSeek R1 demonstrated greater overall consistency and accuracy than ChatGPT-4o when generating information related to molar-incisor hypomineralization, both models exhibited substantial rates of fabricated citations. Similarly, Sağlam et al. (13) found that the accuracy of cited references was consistently lower than the accuracy of the explanatory content itself across all evaluated LLMs. Taken together with the present findings, these observations indicate that fluent language, stable quality scores, and coherent explanations should not be interpreted as guarantees of source reliability. Consequently, LLM-generated health information, particularly when intended for patient education, should continue to undergo professional verification before being incorporated into clinical communication or educational resources.

Upon examination of the consistency analyses evaluated in the present study, it was seen that the responses generated by ChatGPT generally exhibited a satisfactory level of consistency, as evidenced by the high ICC values obtained. In particular, the high ICC values observed in the analyses evaluating all repetitions collectively suggest that the model is generally capable of producing stable and predictable responses to similar prompts. However, the presence of significant differences in readability scores both within the same day and across different days, despite the absence of significant differences in GQS scores over time, suggests that LLMs may exhibit variability in linguistic complexity while maintaining a more stable level of overall information quality. The findings, when considered as a whole, suggest that the evaluation of health-related content generated by LLMs should consider not only a single response, but also temporal variability and response consistency over time.

From a clinical perspective, when the findings obtained from this study are evaluated collectively, LLMs appear to be potential tools which may support rapid access to information for parents in the field of pediatric dentistry. However, the

results also demonstrate that these systems should not be considered completely reliable and standardized sources of information. In particular, the fact that readability levels may not be equally suitable for all user groups and that responses may exhibit a certain degree of variability over time suggests that LLMs should be regarded as supportive tools in clinical practice. Although the Çetinkaya-Uzun readability scores corresponded to the instructional reading level (approximately the 8th-9th grade), this level may still exceed the reading and health literacy skills of a substantial proportion of caregivers. Consequently, despite the consistently high quality of the information provided, some parents may experience difficulty in fully understanding or applying the generated recommendations. Furthermore, the possibility that semi-technical or clinical expressions may be misinterpreted by individuals with low health literacy should also be taken into consideration. These findings highlight the importance of optimizing the readability of LLM-generated patient education materials in addition to ensuring their informational quality particularly as such tools are intended for use by diverse populations with varying levels of health literacy. Therefore, it is recommended that health-related information provided by LLMs should, when necessary, be supported by expert evaluation and interpreted with caution.

From a methodological perspective, the present study possesses several important strengths. Unlike previous investigations which primarily adopted cross-sectional or single-timepoint designs, the present study incorporated a longitudinal repeated-measures protocol consisting of three independent assessments per day over seven consecutive days. This design enabled simultaneous evaluations of readability, information quality, and temporal reproducibility while minimizing the likelihood that the findings reflected a single stochastic response. Furthermore, the combined use of two validated Turkish readability formulas, the GQS, and intraclass correlation analyses provided a multidimensional assessment framework, allowing for the complementary evaluation of linguistic accessibility, educational quality, and response stability. Such an approach offers a more comprehensive characterization of LLM-generated patient education materials in comparison to studies relying on a single evaluation metric.

Study Limitations

Nevertheless, several limitations should be considered when interpreting the findings. First, the present study evaluated only ChatGPT; therefore, the results cannot be generalized to other LLMs, which may demonstrate different linguistic characteristics, quality profiles, or temporal stability. This deliberate methodological choice allowed for a rigorous repeated-measures design focused on the temporal reproducibility of a single model but precluded direct comparisons between competing LLMs. Future studies incorporating multiple LLMs within longitudinal repeated-measures protocols would help determine whether the temporal variability observed in the present study is model-specific or represents a more general characteristic of contemporary LLMs. Second, the evaluation period was limited to seven consecutive days. Given that commercial LLMs undergo continuous algorithmic refinement and periodic updates, longer follow-up periods may reveal additional patterns of temporal variation which could not be captured within the present study. Finally, although readability, information quality, and temporal consistency were comprehensively assessed, this study did not evaluate the factual accuracy, reference validity, or clinical correctness of the generated responses. These dimensions remain essential components of the performance of LLMs and should be incorporated into future investigations in order to provide a more holistic assessment of LLM-assisted patient education.

Conclusion

This study demonstrated that the Turkish responses generated by ChatGPT to frequently asked questions in the field of pediatric dentistry generally exhibited moderate readability, high quality, and good consistency. However, variability in readability scores was observed both within the same day and across different days. The findings obtained indicate that LLMs may be used as supportive tools in patient and parent information processes; however, as the generated content may vary in terms of readability and consistency, these systems should not be regarded as completely reliable sources of information in clinical settings without human supervision and expert evaluation.

Ethics

Ethics Committee Approval: This study was conducted using responses generated by ChatGPT and did not involve human participants, patient data, personal information, or biological materials. Therefore, ethics committee approval was not required for this study.

Informed Consent: This study was conducted using responses generated by ChatGPT and did not involve human participants, patient data, personal information, or biological materials. Therefore, informed consent was not required for this study.

Footnotes

Authorship Contributions

Concept: Irem Nur Bitisik, Ebru Kucukyilmaz Izgi, Design: Ebru Kucukyilmaz Izgi, Irem Nur Bitisik, Data Collection: Selcuk Savas, Mehmet Izgi, Analysis or Interpretation: Ilgin Timarci, Literature Search: Mehmet Izgi, Selcuk Savas, Writing: Irem Nur Bitisik, Selcuk Savas.

Conflict of Interest: The authors declare no conflicts of interest.

Financial Disclosure: The authors received no financial support for the conduct, authorship, or publication of this study.

Supplementary Link: https://d2v96fxpocvxx.cloudfront.net/beb8919b-f013-4ea1-b1c8-40332e840fe1/content-images/14902003-19f3-4e74-8533-4cfb14d01fc2.pdf

References

1. Michalski RS, Carbonell JG, Mitchell TM, editors. Machine learning: an artificial intelligence approach. 1st ed. Berlin, Germany: Springer-Verlag; 1983.
2. Agrawal P, Nikhade P. Artificial intelligence in dentistry: past, present, and future. Cureus. 2022; 14:e27405.
3. Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. J Dent Res. 2020; 99:769-74.
4. Sezer B, Okutan AE. Evaluation of ChatGPT-4's performance on pediatric dentistry questions: accuracy and completeness analysis. BMC Oral Health. 2025; 25:1427.
5. Eggmann F, Weiger R, Zitzmann NU, Blatz MB. Implications of large language models such as ChatGPT for dental medicine. J Esthet Restor Dent. 2023; 35:1098-102.
6. Shahsavar Y, Choudhury A. User intentions to use ChatGPT for self-diagnosis and health-related purposes: cross-sectional survey study. JMIR Hum Factors. 2023; 10:e47564.
7. Issa J, Sohrabniya F, Brinz J, Dyszkiewińska-Konwińska M, Chaurasia A. The role of large language models in dental diagnosis, treatment planning, and prognosis. J Stomatol. 2025; 78:298-303.
8. Islam MR, Urmi TJ, Mosharrafa RA, Rahman MS, Kadir MF. Role of ChatGPT in health science and research: a correspondence addressing potential application. Health Sci Rep. 2023; 6:e1625
9. Alhaidry HM, Fatani B, Alrayes JO, Almana AM, Alfhaed NK. ChatGPT in dentistry: a comprehensive review. Cureus. 2023; 15:e38317
10. Fatani B. ChatGPT for future medical and dental research. Cureus. 2023; 15:e37285.
11. Gülçin Çetin S, Karadağ GG, Çetin SG. Comparative evaluation of artificial intelligence models in answering pediatric dentistry questions of the Turkish Dental Specialization Exam (DUS). Dicle Dent J. 2025; 26:31-7.
12. Aşık A, Kuru E. Analysis of ChatGPT's answers to pedodontics questions asked in the dentistry specialization training entrance exam: cross-sectional study. Turkiye Klinikleri J Dent Sci. 2025; 31:401-6.
13. Sağlam C, Aşık A, Kuru E, et al. Evaluating the efficacy of large language models in providing information for parental inquiries regarding primary care of pediatric oral and dental health. J Clin Pediatr Dent. 2026; 50:132-41.
14. Karamüftüoğlu N, Varol EA, Bal C. Exploring artificial intelligence chatbots in pediatric fluoride education: a cross-sectional study. Sci Rep. 2025; 16:182.
15. Kocaoğlu MH, Demirel A, Kaya İ. Accuracy, quality, and readability analyses of responses from large language models to questions on pediatric dental sedation. BMC Oral Health. 2026; 26:492.
16. Aydın Varol E, Öztürk Z, Bal C, Karamüftüoğlu N. Comparison of two different artificial intelligence chatbots that provide information to patients and parents about primary tooth pulpotomy treatments. BMC Oral Health. 2026; 26:685.
17. Uçar Gündoğar Z, Sarıoğlu D. Comparative assessment of quality, consistency, and reference accuracy of MIH-related clinical information generated by ChatGPT-4o and DeepSeek R1. BMC Oral Health. 2026; 26.
18. Ateşman E. Türkçede okunabilirliğin ölçülmesi. AÜ Tömer Dil Dergisi. 1997; 58:171-4.
19. Çetinkaya G, Uzun L. Türkçe ders kitaplarındaki metinlerin okunabilirlik özellikleri. In: Ülper H, editor. Türkçe Ders Kitabı Çözümlemeleri. Ankara: Pegem Akademi; 2010. p. 141-55.
20. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol. 2007; 102:2070-7.
21. American Academy of Pediatric Dentistry. Frequently asked questions for parents. Available from: https://www.aapd.org/resources/parent/faq/. Accessed: June 2026.
22. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: Wiley; 1981. p. 212-25.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159-74.
24. Koç U, Güneş YC, Çolakoğlu MN, et al. ChatGPT-generated informed consent forms in interventional radiology: a randomized controlled evaluation of comprehension, attitudinal responses, and readability. Eur J Radiol. 2026; 195:112624.
25. Erdat EC, Yalçıner M, Utkan G. Readability and quality analysis of ChatGPT o1's responses on colorectal cancer: a study of an AI language model. Acta Haematol Oncol Turc. 2025; 58:74-80.
26. Aypar Akbağ NN. Assessing artificial intelligence-generated patient educational material on gestational diabetes mellitus. J Perinat Neonatal Nurs. 2025; 39:210-7.
27. Boztaş Demir G, Görgülü S. The Turkish proficiency of ChatGPT and Gemini in orthodontics: an evaluation of the accuracy, completeness, and readability of responses to patient questions. Yeditepe J Dent. 2025; 21:151-8.
28. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel). 2023; 11.
29. Li J, Dada A, Puladi B, Kleesiek J, Egger J. ChatGPT in healthcare: a taxonomy and systematic review. Comput Methods Programs Biomed. 2024; 245:108013.

J Pediatr Res 2026;13(3):314-7
DOI: 10.4274/jpr.galenos.2026.72368

Case Report

Ochoa Syndrome Associated with a Novel HPSE2 Deletion: A Case Report

Günay Ekberli[1], Sevgin Taner[2], Esra Işık[3], Türkan Turkut Tan[3]

[1]University of Health Sciences Türkiye, Ankara Bilkent City Hospital, Clinic of Pediatric Urology, Ankara, Türkiye
[2]Ege University Faculty of Medicine, Department of Pediatric Nephrology, İzmir, Türkiye
[3]Ege University Faculty of Medicine, Department of Pediatric Genetic Diseases, İzmir, Türkiye

ABSTRACT

Ochoa syndrome is a rare disorder characterized by abnormal facial expression and severe lower urinary tract dysfunction. We report a two-year-old girl with recurrent urinary tract infections and a novel homozygous *HPSE2* exon 5-11 deletion. Early recognition and timely intervention are essential in order to prevent renal damage and improve clinical outcomes.

Keywords: Ochoa syndrome, urofacial syndrome, *HPSE2*, recurrent urinary tract infection, vesicostomy

Introduction

Ochoa syndrome, also referred to as urofacial syndrome (UFS), is a rare disorder characterized by the coexistence of abnormal facial expression and lower urinary tract dysfunction in the absence of an identifiable neurological or anatomical obstruction. Affected individuals typically demonstrate an inverted facial expression during smiling or laughing (1). UFS follows an autosomal recessive inheritance pattern, and pathogenic variants involving the *inactive heparanase-2* (*HPSE2*) gene have been identified in the majority of reported cases (2). As the characteristic facial findings may be subtle during early childhood, recognition of this syndrome can be delayed, leading to recurrent urinary tract infections (UTIs) and progressive upper urinary tract damage. We present this case in order to emphasize the importance of early recognition of Ochoa syndrome as a rare but significant cause of severe lower urinary tract dysfunction, recurrent UTI, and chronic kidney disease.

Case Report

A two-year-old girl was referred with recurrent febrile UTIs. Physical examination demonstrated poor weight gain and failure to thrive associated with repeated infectious episodes. Perineal and lumbosacral examinations were unremarkable. However, an abnormal inverted facial expression while smiling was noticed during clinical evaluation (Figure 1). Antibiotic prophylaxis was initiated.

There was no known family history of Ochoa syndrome or other hereditary urinary tract disorders, and no known parental consanguinity. Urinary ultrasonography demonstrated bilateral grade 3 hydroureteronephrosis, severe bladder wall thickening, and thinning of the left

Corresponding Author
Günay Ekberli, MD, University of Health Sciences Türkiye, Ankara Bilkent City Hospital, Clinic of Pediatric Urology, Ankara, Türkiye
E-mail: gnyekbrl@yahoo.com **ORCID:** orcid.org/0000-0002-0021-5998
Received: 22.05.2026 **Accepted:** 21.07.2026 **Epub:** 04.09.2026 **Publication Date:** 18.09.2026

Cite this article as: Ekberli G, Taner S, Işık E, Turkut Tan T. Ochoa syndrome associated with a novel HPSE2 deletion: a case report. J Pediatr Res. 2026;13(3):314-7

renal parenchyma. Dimercaptosuccinic acid scintigraphy revealed cortical irregularities, upper pole parenchymal thinning, and mild dilatation of the left collecting system. Voiding cystourethrography demonstrated vesicoureteral reflux together with trabeculation and diverticular changes of the bladder wall.

Based on the coexistence of characteristic facial findings and severe lower urinary tract dysfunction, Ochoa syndrome was suspected. Molecular confirmation was obtained through whole exome sequencing analysis. Sequence evaluation did not reveal any pathogenic or likely pathogenic single nucleotide variants associated with the phenotype. However, copy number variation (CNV) analysis identified a homozygous deletion within the 10q26.3 region (Hg19: chr10:100242132-100481844), involving exons 5-11 of the *HPSE2* (*NM_021828.4*) gene. No additional morbid OMIM genes were identified within the deleted region.

Figure 1. An abnormal inverted facial expression while smiling

This deletion had not previously been reported in databases including DGV or gnomAD. According to American College of Medical Genetics and Genomics criteria, the CNV was classified as likely pathogenic. Integrative Genomics Viewer analysis demonstrated a complete absence of read depth across exons 5-11 of the *HPSE2* gene, whereas adequate read depth was present in the control samples processed in the same sequencing run (Figure 2). Confirmation analysis was planned; however, multiplex ligation-dependent probe amplification (MLPA) confirmation could not be performed because no specific probe for *HPSE2* was available. In addition, the parents declined segregation analysis.

Despite continuous antibiotic prophylaxis, the patient experienced two episodes of urosepsis requiring hospitalization during follow-up. Considering the severity of the lower urinary tract dysfunction and the patient's clinical condition, vesicostomy was performed. During the first postoperative year, she remained infection-free and demonstrated improvement in growth parameters appropriate for her age. Follow-up radiological evaluation showed improvement of the upper urinary tract. Based on her favorable clinical course, radiological improvement, and satisfactory cystoscopic bladder mucosal appearance, vesicostomy closure with anti-reflux surgery was performed in the 14th postoperative month.

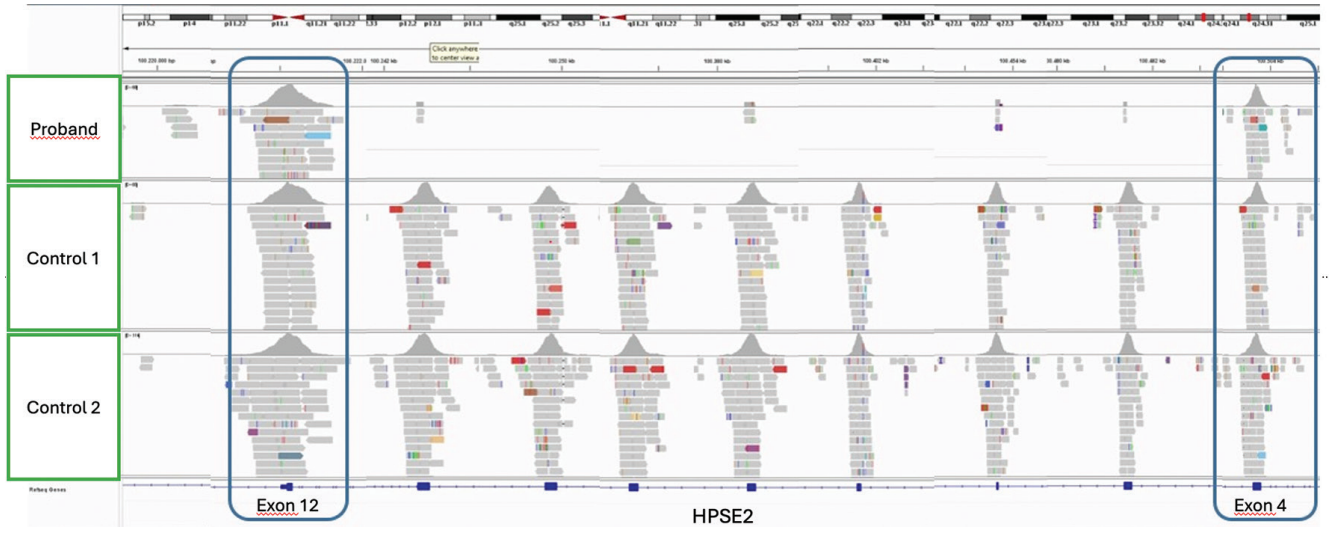

Figure 2. Integrative Genomics Viewer analysis demonstrating the homozygous *HPSE2* exon 5-11 deletion in the patient.

Discussion

Ochoa syndrome, also known as UFS, was initially described clinically in the early 1960s. In 1979, the term 'Ochoa syndrome' was introduced by Elejalde (4) following genetic investigations of the affected families (3). The diagnosis of Ochoa syndrome is established in the presence of one or both of the following findings: 1) inverted facial expression accompanied by clinical and radiological evidence of lower urinary tract dysfunction, and 2) identification of a homozygous pathogenic variant involving the *HPSE2* or *leucine-rich repeats and immunoglobulin-like domains protein 2* (*LRIG2*) gene.

Lower urinary tract dysfunction in these patients may manifest with recurrent UTIs, dysuria, urinary incontinence, frequency, urgency, or enuresis, and may eventually progress to end-stage renal disease if left untreated (5). Although the characteristic paradoxical facial expression during smiling is considered highly suggestive of the syndrome, it may remain unrecognized as bladder-related symptoms usually dominate the clinical presentation.

From a genetic perspective, UFS is an autosomal recessive disorder caused by pathogenic variants affecting both alleles of either the *HPSE2* or *LRIG2* gene (6). Advances in next-generation sequencing (NGS) technologies over the last decade have significantly improved the identification of inherited Mendelian disorders. NGS allows for the simultaneous evaluation of multiple genomic regions in a single analysis and has become a practical and cost-effective diagnostic approach for those patients with suspected genetic diseases (7). Nevertheless, conventional NGS analysis may be insufficient for the detection of CNVs, including exon-level deletions, which represent an important mechanism underlying many genetic disorders. The incorporation of CNV detection algorithms into NGS-based diagnostic workflows has therefore substantially increased diagnostic accuracy in genetically heterogeneous conditions (8,9).

In the present case, whole exome sequencing combined with CNV analysis identified a homozygous deletion involving exons 5-11 of the *HPSE2* gene in the 10q26.3 region, which was classified as likely pathogenic. Previous studies have highlighted the functional importance of exon deletions affecting *HPSE2*. Beaman et al. (10) demonstrated four possible heparanase-2 isoforms generated through differential splicing of exons 3 and 4. In addition, Daly et al. (2) described a family with UFS carrying a whole exon 3 deletion in *HPSE2*. Unlike previously reported variants, our patient demonstrated a homozygous deletion spanning exons 5-11, a finding that has not previously been described in association with UFS. This observation further expands the known mutational spectrum of *HPSE2*. The absence of isoforms involving exons 5-11 may interfere with the pathways required for normal bladder innervation and facial muscle function, thereby contributing to the typical clinical phenotype observed in UFS.

Although the molecular findings strongly supported the diagnosis, several limitations should be acknowledged. Confirmatory MLPA analysis could not be performed because a specific probe for *HPSE2* was unavailable, and the family declined segregation analysis. Despite these limitations, our findings support the growing evidence regarding the central role of *HPSE2* in the pathogenesis of Ochoa syndrome and underline the importance of comprehensive genomic analysis in rare inherited disorders. However, extensive genetic confirmation should not delay the initiation of treatment, since timely management remains essential for the preservation of renal function and the improvement of long-term outcomes, particularly considering that approximately 16% of patients may not harbor currently recognized mutations (1).

Conclusion

In children presenting with severe lower urinary tract dysfunction, recognition of the characteristic inverted facial expression may facilitate an earlier diagnosis of Ochoa syndrome. Timely intervention is essential in order to reduce the risk of progressive renal impairment.

Ethics

Informed Consent: Written informed consent was obtained from the patient's legal guardian for publication of this case report and the accompanying clinical images.

Footnotes

Authorship Contributions

Surgical and Medical Practices: G.E., Design: S.T., Analysis or Interpretation: S.T., Literature Search: E.I., Writing: G.E., T.T.T.

Conflict of Interest: The authors declare no conflict of interest.

Financial Disclosure: The authors declare that no financial or material support was received for this study.

References

1. Osorio S, Rivillas ND, Martinez JA. Urofacial (Ochoa) syndrome: a literature review. J Pediatr Urol. 2021; 17:246-54.

2. Daly SB, Urquhart JE, Hilton E, et al. Mutations in HPSE2 cause urofacial syndrome. Am J Hum Genet. 2010; 86:963-9.

3. Ochoa B. Can a congenital dysfunctional bladder be diagnosed from a smile? The Ochoa syndrome updated. Pediatr Nephrol. 2004; 19:6-12.

4. Elejalde BR. Genetic and diagnostic considerations in three families with abnormalities of facial expression and congenital urinary obstruction: "The Ochoa syndrome". Am J Med Genet. 1979; 3:97-108.

5. Tu Y, Yang P, Yang J, et al. Clinical and genetic characteristics for the urofacial syndrome (UFS). Int J Clin Exp Pathol. 2014; 7:1842-8.

6. Newman WG, Woolf AS. Urofacial syndrome. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2021.

7. Jamuar SS, Tan EC. Clinical application of next-generation sequencing for Mendelian diseases. Hum Genomics. 2015; 9:1-6.

8. Ellingford JM, Campbell C, Barton S, et al. Validation of copy number variation analysis for next-generation sequencing diagnostics. Eur J Hum Genet. 2017; 25:719-24.

9. Atik T, Avci Durmusalioglu E, Isik E, et al. Diagnostic yield of exome sequencing-based copy number variation analysis in Mendelian disorders: a clinical application. BMC Med Genomics. 2024; 17:239.

10. Beaman GM, Lopes FM, Hofmann A, et al. Expanding the HPSE2 Genotypic Spectrum in Urofacial Syndrome, A Disease Featuring a Peripheral Neuropathy of the Urinary Bladder. Front Genet. 2022;13:896125

Erratum

J Pediatr Res 2026;13(1):51-57
DOI: 10.4274/jpr.galenos.2025.58908

Following the publication of "Erbaş S, Erbasan Zİ, Özbaran NB. Transdiagnostic Impact of Temperament on Symptom Severity and Quality of Life in Preschoolers with Neurodevelopmental Disorders. J Pediatr Res. 2026 Mar 23;13(1):51-57. doi: 10.4274/jpr.galenos.2025.58908" article, it has been requested that the surname of one of the authors, Zeynep İrem Erbasan, be updated to "Şener" due to a change in marital status. This correction has been implemented. The sections where corrections were made are detailed below:

Published on page 51;

Seda Erbaş[1], Zeynep İrem Erbasan[2], Nazlı Burcu Özbaran[3]

Corrected on page 51;

(The corrected parts are given in bold)

Seda Erbaş[1], Zeynep İrem **Şener**[2], Nazlı Burcu Özbaran[3]

Published on page 51;

Cite this article as: Erbaş S, Erbasan Zİ, Özbaran NB. Transdiagnostic impact of temperament on symptom severity and quality of life in preschoolers with neurodevelopmental disorders. J Pediatr Res. 2026;13(1):51-57

Corrected on page 51;

(The corrected parts are given in bold)

Cite this article as: Erbaş S, **Şener** Zİ, Özbaran NB. Transdiagnostic impact of temperament on symptom severity and quality of life in preschoolers with neurodevelopmental disorders. J Pediatr Res. 2026;13(1):51-57

Published on page 56;

Authorship Contributions

Concept: S.E., Design: N.B.Ö., Data Collection or Processing: Z.İ.E., Analysis or Interpretation: S.E., Z.İ.E., Literature Search: S.E., N.B.Ö., Writing: S.E., Z.İ.E., N.B.Ö.

Published on page 56;

(The corrected parts are given in bold)

Authorship Contributions

Concept: S.E., Design: N.B.Ö., Data Collection or Processing: Z.İ.**Ş**., Analysis or Interpretation: S.E., Z.İ.**Ş**., Literature Search: S.E., N.B.Ö., Writing: S.E., Z.İ.**Ş**., N.B.Ö.