

SoC Surgery on Children

Volume: 3

Issue: 4

Year: 2026



EDITOR-IN-CHIEF

Prof. Atilla ŞENAYLI

Pediatric Surgeon, Freelance, Kırklareli, Türkiye

ASSOCIATE EDITOR-IN-CHIEF

Assoc. Prof. Sevgi ULUSOY TANGÜL

Department of Pediatric Surgery, Faculty of Medicine, Kırklareli University, Kırklareli, Türkiye

EDITORIAL BOARD

Prof. Ahmet DEMİR

Department of Pediatric Oncology, Memorial Healthcare Group, Ankara, Türkiye

Prof. Bülent BAKAR

Department of Neurosurgery, Faculty of Medicine, Kırıkkale University, Kırıkkale, Türkiye

Prof. Haldun UMUDUM

Department of Medical Pathology, Güven Health Group, Ankara, Türkiye

Prof. Halil İbrahim UÇAR

Department of Pediatric Cardiovascular Surgery, Ankara Etlik City Hospital, Ankara, Türkiye

Prof. Haluk EMİR

Department of Pediatric Urology, Faculty of Medicine, İstanbul University, İstanbul, Türkiye

Prof. Hesham ALKADY

Department of Cardiovascular Surgery, Cairo University, Cairo, Egypt

Prof. Resul YILMAZ

Division of Pediatric Intensive Care, Department of Pediatrics, Faculty of Medicine, Selçuk University, Konya, Türkiye

Prof. Serdar ALAN

Division of Neonatology, Department of Pediatrics, Faculty of Medicine, Kırıkkale University, Kırıkkale, Türkiye

Prof. Theodossis PAPAVERAMIDIS

Department of Endocrine Surgery, Euromedica General Hospital, Rhodes, Greece

Prof. Tutku SOYER

Department of Pediatric Surgery, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Prof. Ünal BIÇAKCI

Division of Pediatric Urology, Department of Pediatric Surgery, Faculty of Medicine, Ondokuz Mayıs University, Samsun, Türkiye

Assoc. Prof. Hülya İPEK

Department of Pediatric Surgery, Faculty of Medicine, Hitit University, Çorum, Türkiye

Assoc. Prof. Mehmet CANIKLIOĞLU

Department of Urology, Faculty of Medicine, Yozgat Bozok University, Yozgat, Türkiye

Assoc. Prof. Mehmet ERDOĞAN

Department of Nuclear Medicine, Faculty of Medicine, Süleyman Demirel University, Isparta, Türkiye

Assoc. Prof. Yeşim ŞENAYLI

Anesthesiology and Reanimation Specialist, Private Practice, Kırklareli, Türkiye

Asst. Prof. Amir Hossein LADAN

Department of Pediatric Surgery, Zanjan University, Tabriz, Iran

Asst. Prof. İsmail Essam ELHALABY

Department of Pediatric Surgery, Faculty of Medicine, Tanta University, Tanta, Egypt

Asst. Prof. İsmail BÜYÜKCERAN

Department of Orthopedics and Traumatology, Faculty of Medicine, Ondokuz Mayıs University, Samsun, Türkiye

Asst. Prof. Simay ERŞAHİN, MD

Department of Plastic, Reconstructive and Aesthetic Surgery, Faculty of Medicine, Eskişehir Osmangazi University, Eskişehir, Türkiye

Asst. Prof. Yakup ACET

Department of Ophthalmology, Faculty of Medicine, Mardin Artuklu University, Mardin, Türkiye

Asst. Prof. Yunus Emre DURMUŞ

Department of Neurosurgery, Faculty of Medicine, Ondokuz Mayıs University, Samsun, Türkiye

Spec. İlhamı JAFARLI, MD

MRCs, FACS, FEBPS, Department of Pediatric Surgery, Chelsea and Westminster Hospital, London, Imperial College London, United Kingdom

Spec. Nilüfer BİÇAKCI, MD

Department of Nuclear Medicine, Samsun City Hospital, Samsun, Türkiye



ENGLISH LANGUAGE EDITOR

Ahmet Baki ŞENAYLI

International Affairs and Global Politics, METU, Ankara, Türkiye

STATISTICS EDITOR

Prof. Ünal ERKORKMAZ

Department of Biostatistics, Faculty of Medicine, Sakarya University, Sakarya, Türkiye

LAYOUT EDITOR

Hatice AKYIL

Biologist, MediHealth Academy Publishing, Ankara, Türkiye

ORIGINAL ARTICLES

- Comparison between topical timolol and oral propranolol for the treatment of superficial infantile hemangioma..... 107-112
Tarannum T, Al Faisal A, Islam K.
- Variability of histologic pattern and transition zone length in Hirschsprung disease and associated factors.....113-120
Tamba RP, Kurnia N.
- Interdisciplinary management of children with cerebral palsy: neurological perspective on neurosurgical interfaces 121-125
Majidova N, Madjidova YN.
- Evaluation of the pan-immune-inflammation value (PIV) as a prognostic and surgical decision-making marker in necrotising enterocolitis..... 126-132
Yıldırım İ, Ünlü Ballı SE, Bahadır GB, Sürer İ.
- The potential of AI applications to provide clinical counseling to parents: a pilot study.....133-138
Kuas N, Aktaş Ö, Karkın EB, Tekkanat B.
- Iodine nutrition, bamboo shoot consumption, thiocyanate exposure, thyroid function, weight and height, and associations among them in the tribal children of Tripura, Northeast India 139-148
Ray I, Chandra AK, Debbarma SK, Mookerjee SK, Datta A, Singh LH.
- Beyond fecal diversion: stoma-related morbidity in a cohort of 318 children 149-154
Uzun AU, Öztörün Cİ.

REVIEW

- Pediatric supernumerary testis: preservation, orchiopexy, or orchiectomy? A contemporary review of diagnosis and surgical decision-making 155-160
Baltogiannis N, Fili F, Tsangos A, Baltogianni ME.

Comparison between topical timolol and oral propranolol for the treatment of superficial infantile hemangioma

©Tasniya Tarannum^{*1}, ©Abdullah Al Faisal², ©Kabirul Islam³

¹Department of Pediatric Surgery, Dhaka Medical College & Hospital, Dhaka, Bangladesh

²Department of Pediatric Surgery, Shaheed Suhrawardy Medical College & Hospital, Dhaka, Bangladesh

³Department of Pediatric Surgery, Bangladesh Shishu Hospital and Institute, Dhaka, Bangladesh

Received: 03/04/2026

Accepted: 17/05/2026

Published: 20/10/2026

Cite this article: Tarannum T, Al Faisal A, Islam K. Comparison between topical timolol and oral propranolol for the treatment of superficial infantile hemangioma. *Surg Child.* 2026;3(4):107-112. doi:10.51271/SOC-0082

*Corresponding Author: Tasniya Tarannum, drtasniya@gmail.com

ABSTRACT

Aims: Infantile hemangiomas are the most common benign paediatric vascular tumors. In spite of good results in the management of infantile hemangioma with oral propranolol, it may cause systemic side effects in the long run. The use of topical timolol can minimize systemic side effects, and it is easy to apply in the paediatric age group. This study aimed to evaluate a safer, more effective, and more easily applicable drug for the treatment of infantile hemangioma.

Methods: It was a prospective comparative interventional study that was conducted in Dhaka Shishu (Children) Hospital from May 2018 to September 2020. A total of 89 patients with superficial infantile hemangioma were selected purposively for this study who completed 6 months of follow-up after the initial treatment. They were randomly assigned to the topical timolol group (group A=44) and oral propranolol group (group B=45). Regression of size, change of color, adverse effects, and parent satisfaction regarding the mode of administration of drugs were statistically compared between the two groups.

Results: The age range of the patients was 1 to 12 months, where the majority of the patients were female. Most of the lesions were in the head and neck region in both groups. In group A 72.7% patients showed excellent response in size regression, compared to group B 68.9% patients, which showed no statistically significant difference. In group A 75% had an excellent color regression, compared to group B 68.9% ($p=0.0784$). Regarding adverse effects in group A, only 3 children (6.8%) developed adverse effects compared to 12 (26.7%) in group B which was statistically significant ($p=0.021$). In group A (97.7%) parents were more satisfied than in group B (80.0%) regarding the mode of administration of two drugs, which was statistically significant ($p=0.015$).

Conclusion: Use of topical timolol turns out to be a good choice for the treatment of patients with superficial infantile hemangioma. It has the same efficacy, but the adverse effect is comparatively less than oral propranolol. Applying topical timolol is easier for the parents than oral propranolol.

Keywords: Hemangioma, capillary, propranolol, timolol

INTRODUCTION

Infantile hemangiomas (IHs) are the most common benign vascular tumors in infancy, with an incidence of about 4-5%, higher in females and premature infants, and lower in dark-skinned babies.^{1,2} According to the International Society for the Study of Vascular Anomalies (ISSVA), benign cutaneous vascular tumors include IH, congenital hemangioma, tufted angioma, pyogenic granuloma, and others.³ IHs usually appear shortly after birth and are characterized by rapid proliferation during the first year of life, followed by gradual involution.⁴ Most lesions are solitary (80%) and commonly involve the head and neck (60%), trunk (25%), and limbs (15%). About 20% of patients may present with multiple

lesions.¹ The life cycle includes proliferative, involuting, and involuted phases.⁵ Although many IHs regress spontaneously, some can cause complications such as hemorrhage, ulceration, disfigurement, or functional impairment requiring early intervention.^{6,7} Historically, treatment options included corticosteroids, vincristine, interferon- α , laser therapy, or surgery, but these were often associated with significant adverse effects.⁸ In 2008, propranolol—a non-selective lipophilic β -blocker—was serendipitously found to induce hemangioma regression.⁹ Subsequent studies confirmed its high efficacy,¹⁰⁻¹³ including research conducted locally.^{14,15} However, systemic propranolol may cross the blood-brain



barrier, raising concerns about potential adverse effects such as bradycardia, hypotension, hypoglycemia, bronchospasm, and possible neurocognitive effects.¹⁶⁻¹⁹ Topical timolol, another non-selective β -blocker long used for pediatric glaucoma,²⁰ was first reported effective against superficial IHs in 2010.²¹ A large series later demonstrated its efficacy and safety without cardiovascular adverse effects.²² To minimize systemic risks of propranolol, topical timolol has gained increasing acceptance for treating superficial IHs.^{6,23} This study aims to evaluate the efficacy, safety, and parental satisfaction of topical timolol in treating IHs compared with oral propranolol under continuous clinical monitoring.

METHODS

Ethics

The study protocol was reviewed and approved by the Ethical Review Committee of the Bangladesh Institute of Child Health and Dhaka Shishu (Children) Hospital (Date: 05.03.2018, Decision No: BICH-ERC-11-01-2018). Written informed consent was obtained from the parents or legal guardians of all participating infants prior to enrollment, following a detailed explanation of the study objectives, interventions, and potential risks.

Study Design

This prospective comparative interventional study was conducted in the Division of Paediatric Surgery, Bangladesh Institute of Child Health and Dhaka Shishu (Children) Hospital, Dhaka, over two and a half years (May 2018-September 2020).

Study Population

Patients attending the surgical outpatient department with clinically diagnosed superficial IH during the study period were enrolled using a purposive sampling technique until the desired sample size was achieved.

Selection Criteria

Inclusion: All patients with superficial IH in the proliferative phase.

Exclusion: Patients with previously treated or infected hemangiomas, life-threatening conditions, and cardiovascular disorders contraindicate propranolol use.

Sample Size and Randomization

Sample size was calculated using the formula:

$$n = [(P_1(1-P_1) + P_2(1-P_2)) / (P_1 - P_2)^2] \times (Z\alpha + Z\beta)^2,$$

where $P_1=0.03$, $P_2=0.22$, $Z\alpha=1.96$, and $Z\beta=0.85$, yielding 44 per group. Considering a 10% dropout rate, 50 patients were taken per group (total 100). Simple randomization was utilized to allocate patients into group A (topical timolol) and group B (oral propranolol). Allocation concealment was maintained using sequentially numbered, opaque, sealed envelopes drawn by an independent coordinator.

Data Collection and Measurements

Information was obtained from parents via interview and physical examination. Lesion size was measured with a

soft tape, and color change was analyzed digitally (Adobe Photoshop) using serial photographs taken with a 13MP digital camera. To ensure objectivity, these clinical and photographic evaluations were independently confirmed by two physicians who were blinded to the group allocation. Vital parameters (heart rate, blood pressure, and random blood sugar) were monitored before and after drug administration at each follow-up. Adverse effects and parental satisfaction were recorded.

Intervention

Group A received 0.5% topical timolol maleate solution (2–4 drops twice daily) and group B received oral propranolol (1–3 mg/kg/day in two doses). Dosage was titrated based on clinical response and tolerance.

Follow-up and Outcome Measures

Patients were followed at 2 weeks, 1, 2, 3, 4, 5, and 6 months. Regression of size and color was calculated as a percentage of pretreatment values. Responses were graded as excellent (80–100%), good (50–80%), fair (25–50%), or poor (<25%).

Statistical Analysis

Data were analyzed using SPSS v25. Continuous variables were assessed by Student's t-test and categorical variables by Chi-square/Fisher's exact test, with $p<0.05$ considered significant.

RESULTS

A total of 100 infants aged between 1 to 12 months were included, and 89 completed the six-month follow-up (44 in group A: timolol, and 45 in group B: propranolol). The mean age was 5.39 ± 2.94 months in group A and 6.38 ± 3.17 months in group B, with no significant difference ($p=0.134$) (Table 1). The majority were term babies (95.5% vs 93.3%) and females (61.4% vs 68.9%), showing no statistical difference between groups (Table 2, 3). A positive family history of hemangioma was noted in 4.5% of group A and 8.9% of group B ($p>0.05$) (Table 4).

Table 1. Comparison of patients by age

Age (months)	Group A (n=44)	Group B (n=45)	p value
Mean $\pm$ SD	5.39 $\pm$ 2.94	6.38 $\pm$ 3.17	0.134
Range	2–12	1–12	

SD: Standard deviation

Table 2. Comparison by gestational age

Gestational age	Group A n (%)	Group B n (%)	p value
Preterm	2 (4.5)	3 (6.7)	1.000
Term	42 (95.5)	42 (93.3)	

Table 3. Comparison by gender

Gender	Group A n (%)	Group B n (%)	p value
Male	17 (38.6)	14 (31.1)	0.509
Female	27 (61.4)	31 (68.9)	

Table 4. Family history of hemangioma

Family history	Group A n (%)	Group B n (%)	p value
Absent	42 (95.5)	41 (91.1)	0.677
Present	2 (4.5)	4 (8.9)	

Lesion distribution was similar in both groups, predominantly affecting the head and neck region (65.9% in group A and 77.8% in group B), followed by lesions on the chest wall, abdominal wall, and extremities.

The mean pretreatment size of lesions was $4.43 \pm 1.78 \text{ cm}^2$ in group A and $5.09 \pm 1.52 \text{ cm}^2$ in group B ($p=0.062$). After six months, post-treatment size significantly reduced in both groups ($0.74 \pm 0.79 \text{ cm}^2$ and $0.91 \pm 0.87 \text{ cm}^2$ respectively; $p<0.001$ within each group). However, intergroup comparison showed no significant difference in mean size reduction ($p=0.333$) (Table 5).

Table 5. Size regression before and after treatment

Parameter	Group A	Group B	p value
Pre-treatment size (cm^2)	4.43 ± 1.78	5.09 ± 1.52	0.062
Post-treatment size (cm^2)	0.74 ± 0.79	0.91 ± 0.87	0.333

Regarding categorical response, 72.7% of group A and 68.9% of group B achieved excellent size regression, while 25.0% and 28.9% had good responses respectively ($p=0.905$) (Table 6). For color regression, excellent response was observed in 75.0% of group A and 68.9% of group B ($p=0.0784$), with no significant difference between groups (Table 7). Complete regression occurred in 29.6% of group A and 17.8% of group B ($p=0.220$) (Table 8). Most complete regressions occurred within 4–5 months of treatment (Figures 1-12).

Table 6. Regression of size between two groups

Response	Group A n (%)	Group B n (%)	p value
Excellent	32 (72.7)	31 (68.9)	0.905
Good	11 (25.0)	13 (28.9)	
Fair	1 (2.3)	1 (2.2)	

Table 7. Regression of color between two groups

Response	Group A n (%)	Group B n (%)	p value
Excellent	33 (75.0)	31 (68.9)	0.0784
Good	10 (22.7)	12 (26.7)	
Fair	1 (2.3)	2 (4.4)	

Table 8. Complete regression of lesions

Complete regression	Group A n (%)	Group B n (%)	p value
Occurred	13 (29.6)	8 (17.8)	0.220
Not occurred	31 (70.5)	37 (82.2)	

Adverse effects were significantly more frequent in the propranolol group (26.7%) than in the timolol group (6.8%) ($p=0.021$) (Table 9). In group A, 4.5% had local pruritus and 2.3% had sleep disturbance. In group B, bradycardia and

**Figure 1.** Pretreatment figure**Figure 2.** After 3 months of treatment with timolol**Figure 3.** After 6 months of treatment with timolol**Figure 4.** Pretreatment figure



Figure 5. After 2 months of treatment with timolol



Figure 9. After 6 months of treatment with propranolol



Figure 6. After 6 months of treatment with timolol



Figure 10. Pretreatment figure



Figure 7. Pretreatment figure



Figure 11. After 3 months of treatment with propranolol



Figure 8. After 3 months of treatment with propranolol



Figure 12. After 6 months of treatment with propranolol

hypoglycemia each occurred in 6.7%, hypotension in 4.4%, with minor cases of sleep disturbance and diarrhea (4.4% each).

Table 9. Comparison of adverse effects

Adverse effects	Group A n (%)	Group B n (%)	p value
Absent	41 (93.2)	33 (73.3)	0.021
Present	3 (6.8)	12 (26.7)	

Parental satisfaction regarding drug administration was notably higher in the timolol group (97.7%) than in the propranolol group (80.0%), which was statistically significant ($p=0.015$) (Table 10).

Table 10. Parents' satisfaction regarding drug administration

Satisfaction	Group A n (%)	Group B n (%)	p value
Satisfied	43 (97.7)	36 (80.0)	0.015
Not satisfied	1 (2.3)	9 (20.0)	

DISCUSSION

This prospective study compared the clinical outcomes of topical timolol and oral propranolol in 100 patients with superficial IHs, randomly divided into two groups. Six patients in group A and five in group B dropped out due to difficulty maintaining follow-up. Most lesions presented in the proliferative phase, which lasts until about 10–12 months of age. The mean age was 5.39 ± 2.94 months in group A and 6.38 ± 3.17 months in group B, showing no statistical difference, consistent with Wu et al.²⁴ Although previous studies suggest a higher incidence of IH in preterm infants,^{3,25} only 5.61% of our patients were preterm. This variance from the literature may be influenced by our specific clinical setting as a single-center study, regional variations in patient presentation, or potentially incomplete gestational records. Female predominance was observed in both groups—61.4% in group A and 68.9% in group B—similar to other reports.^{1,25} A positive family history was found in 6.7% of cases, comparable to Kabir.¹⁵ Hemangiomas were most common in the head and neck region—65.9% in group A and 77.8% in group B—likely due to the region's rich vascularity. Similar distributions were reported by Weibel et al.²⁶ and Yu et al.⁶ In group A, mean pre-treatment size was 4.43 ± 1.78 cm², reducing to 0.74 ± 0.79 cm² post-treatment. Topical timolol caused vasoconstriction and inhibited angiogenic growth factors, producing significant regression, as also reported by Weibel et al.²⁶ In group B, mean lesion size decreased from 5.09 ± 1.52 cm² to 0.91 ± 0.87 cm². Oral propranolol suppresses hemangioma proliferation by modulating endothelial cell activity and pericyte contractility.¹⁴ There was no significant difference in regression between groups, consistent with Wu et al.,²⁴ likely due to their similar β -blocking mechanisms. Excellent size regression ($\geq 80\%$) was observed in 72.7% of group A and 68.9% of group B. Variation in lesion size may explain the difference in individual responses. No significant difference between groups aligns with previous findings.²⁴ Color regression from dark red to lighter or normal skin tone occurred in 75% of group A and 68.9% of group B. Topical timolol caused vasoconstriction through nitric oxide inhibition, while propranolol reduced capillary perfusion via

β -adrenergic blockade. Our higher excellent regression rate compared with Wu et al.²⁷ may be due to different evaluation criteria. Complete regression of both size and color occurred in 29.6% of group A and 17.8% of group B, with no statistical significance. Most complete regressions were in small lesions after six months of therapy. The lower rate compared with Nagdeve and Saoji²⁸ may be due to our shorter follow-up duration. Adverse effects occurred in 6.8% of group A—mainly local pruritus and mild sleep disturbance—and 26.7% of group B, including bradycardia, hypotension, hypoglycemia, diarrhea, and sleep disturbance. Systemic absorption of propranolol explains its higher incidence of side effects.^{14,15} Adverse effects were significantly fewer in the topical group, matching Wu et al.²⁴ Parental satisfaction was significantly higher with topical timolol (97.7%) compared with oral propranolol (80.0%) because topical administration was easier and better tolerated. In our study, both topical timolol and oral propranolol are effective for superficial IH, showing comparable efficacy. However, topical timolol has a superior safety profile and higher parental satisfaction, making it a preferable first-line option for uncomplicated superficial lesions.

Recommendations

- Considering efficacy and safety, topical timolol could be the first line of treatment for superficial infantile hemangioma.
- A multi-center study with a large sample size and long follow-up is recommended.
- Future studies incorporating pharmacokinetic profiling and serum bioavailability assays are recommended to better elucidate the dose-response relationships and specific mechanistic differences between these topical and systemic therapies.

Limitations

- Long-term follow-up could not be possible.
- Depth of the hemangioma could not be measured.
- This analysis was done in a single center in Bangladesh, and the sample only represents a small fraction of patients with hemangioma.
- The absence of a comparative pharmacological analysis, including pharmacokinetics and serum bioavailability, limits our ability to explain the mechanistic differences between the evaluated drugs.

This study has certain limitations that should be acknowledged. First, conducting a direct pharmacological comparison between the two interventions is inherently challenging. While oral propranolol was administered using precise weight-based dosing (mg/kg/day), topical timolol was applied using a fixed drop count. Consequently, the exact volume applied and the degree of transdermal systemic absorption could not be quantified, limiting a strict dose-response analysis.

CONCLUSION

Topical timolol is as effective as oral propranolol for the treatment of superficial IH. It has fewer adverse effects and its mode of application is more satisfactory to parents compared to oral propranolol.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study protocol was reviewed and approved by the Ethical Review Committee of the Bangladesh Institute of Child Health and Dhaka Shishu (Children) Hospital (Date: 05.03.2018, Decision No: BICH-ERC-11-01-2018).

Informed Consent

Informed consent was obtained from a parent or legal guardian. Where appropriate, age-adjusted assent was also obtained from the child. The inclusion of vulnerable populations in this study adhered to national and international ethical guidelines. Extra care was taken to ensure voluntary participation, understanding, and protection of participant dignity and autonomy.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: TT; Design: TT; Control: KI; Resources: TT, AAF; Materials: TT, AAF; Data Collection and/or Processing: TT, AAF; Analysis and/or Interpretation: TT, AAF; Literature Review: AAF; Writing the Article: TT, AAF; Critical Review: KI; Final Approval: TT, AAF, KI; Accountability: TT, AAF, KI.

REFERENCES

- Kulungowski AM, Fishman SJ. Vascular anomalies. In: Coran AG, Adzick NS, Krummel TM, Laberge J, Shamberger RC, Caldamone AA, eds. *Pediatric Surgery*. 7th ed. Elsevier; 2012.
- Chen TS, Eichenfield LF, Friedlander SF. Infantile hemangiomas: an update on pathogenesis and therapy. *Pediatrics*. 2013;131:99-108. doi:10.1542/peds.2012-1128
- Darrow DH, Greene AK, Mancini AJ, Nopper AJ. Diagnosis and management of infantile hemangioma. *Pediatrics*. 2015;136(4):e1060-e1104. doi:10.1542/peds.2015-2485
- Mulliken JB, Glowacki J. Hemangiomas and vascular malformations in infants and children: a classification based on endothelial characteristics. *Plast Reconstr Surg*. 1982;69:412-422. doi:10.1097/00006534-198203000-00002
- Frieden IJ, Haggstrom AN, Drolet BA, et al. Infantile hemangiomas: current knowledge, future directions. Proceedings of a research workshop on infantile hemangiomas, April 7-9, 2005, Bethesda, Maryland, USA. *Pediatr Dermatol*. 2005;22(5):383-406. doi:10.1111/j.1525-1470.2005.00102.x
- Yu L, Li S, Su B, et al. Treatment of superficial infantile hemangiomas with timolol: evaluation of short-term efficacy and safety in infants. *Exp Ther Med*. 2013;6(2):388-390. doi:10.3892/etm.2013.1176
- Drolet BA, Esterly NB, Frieden IJ. Hemangiomas in children. *N Engl J Med*. 1999;341(3):173-181. doi:10.1056/NEJM199907153410307
- Leonardi-Bee J, Batta K, O'Brien C, Bath-Hextall FJ. Interventions for infantile hemangiomas (strawberry birthmarks) of the skin. *Cochrane Database Syst Rev*. 2011;11(5): CD006545. doi:10.1002/14651858.CD006545.pub2
- Léauté-Labrèze C, Dumas de la Roque E, Hubiche T, Boralevi F, Thambo JB, Taieb A. Propranolol for severe hemangiomas of infancy. *N Engl J Med*. 2008;358(24):2649-2651. doi:10.1056/NEJMc0708819
- Lawley LP, Siegfried E, Todd JL. Propranolol treatment for hemangioma of infancy: risks and recommendations. *Pediatr Dermatol*. 2009;26(5):610-614. doi:10.1111/j.1525-1470.2009.00975.x
- Holmes WJ, Mishra A, Gorst C, Liew SH. Propranolol as first line treatment for infantile hemangiomas. *Plast Reconstr Surg*. 2010;125(1):420-421. doi:10.1097/PRS.0b013e3181c2a731
- Manunza F, Syed S, Laguda B, et al. Propranolol for complicated infantile hemangiomas: a case series of 30 infants. *Br J Dermatol*. 2010;162(2):466-468. doi:10.1111/j.1365-2133.2009.09597.x
- Hermans DJ, van Beynum IM, Schultze Kool LJ, van de Kerkhof PC, Wijnen MH, van der Vleuten CJ. Propranolol, a very promising treatment for ulceration in infantile hemangiomas: a study of 20 cases with matched historical controls. *J Am Acad Dermatol*. 2011;64(5):833-838. doi:10.1016/j.jaad.2011.01.025
- Haque MM, Ferdous KMN, Saha BK, Paul SK, Islam MK. Oral propranolol and prednisolone in the treatment of infantile hemangioma: a comparative study. *Chattagram Maa-O-Shishu Hosp Med Coll J*. 2014;13(1):26-31. doi:10.3329/cmshmcj.v13i1.19414
- Kabir KA. Treatment of infantile hemangioma with atenolol and propranolol [master's thesis]. University of Dhaka; 2016.
- Kroes MC, Strange BA, Dolan RJ. Beta-adrenergic blockade during memory retrieval in humans evokes a sustained reduction of declarative emotional memory enhancement. *J Neurosci*. 2010;30(11):3959-3963. doi:10.1523/JNEUROSCI.5469-09.2010
- Solman L, Murabit A, Gnarra M, Harper JI, Syed SB, Glover M. Propranolol for infantile hemangiomas: single centre experience of 250 cases and proposed therapeutic protocol. *Arch Dis Child*. 2014;99(12):1132-1136. doi:10.1136/archdischild-2014-306514
- Hoeger PH. Propranolol for infantile hemangiomas: certain chances, potential risks. *Br J Dermatol*. 2015;172(1):3-4. doi:10.1111/bjd.13535
- Langley A, Pope E. Propranolol and central nervous system function: potential implications for paediatric patients with infantile hemangiomas. *Br J Dermatol*. 2015;172(1):13-23. doi:10.1111/bjd.13379
- Plager DA, Whitson JT, Netland PA. Betaxolol hydrochloride ophthalmic suspension 0.25% and timolol gel-forming solution 0.25% and 0.5% in pediatric glaucoma: a randomized clinical trial. *J AAPOS*. 2009;13(4):384-390. doi:10.1016/j.jaapos.2009.04.017
- Guo S, Ni N. Topical treatment for capillary hemangioma of the eyelid using beta-blocker solution. *Arch Ophthalmol*. 2010;128(2):255-256. doi:10.1001/archophthalmol.2009.370
- Püttgen K, Lucky A, Adams D, et al. Topical timolol maleate treatment of infantile hemangiomas. *Pediatrics*. 2016;138(3):e20160355. doi:10.1542/peds.2016-0355
- Chan H, McKay C, Adams S, Wargon O. RCT of timolol maleate gel for superficial infantile hemangiomas in 5- to 24-week-olds. *Pediatrics*. 2013;131(6):e1739-e1747. doi:10.1542/peds.2012-3828
- Wu HW, Wang X, Zhang L, Zheng JW, Liu C, Wang YA. Topical timolol vs oral propranolol for the treatment of superficial infantile hemangiomas. *Front Oncol*. 2018;8:605. doi:10.3389/fonc.2018.00605
- Rialon KL, Fishman SJ. Vascular and lymphatic anomalies. In: Ashcraft's Pediatric Surgery. 6th ed. Elsevier; 2014.
- Weibel L, Barysch MJ, Scheer HS, et al. Topical timolol for infantile hemangiomas: evidence for efficacy and degree of systemic absorption. *Pediatr Dermatol*. 2016;33(2):184-190. doi:10.1111/pde.12767
- Wu HW, Liu C, Wang X, et al. Topical application of 0.5% timolol maleate hydrogel for the treatment of superficial infantile hemangioma. *Front Oncol*. 2017;7:137. doi:10.3389/fonc.2017.00137
- Nagdeve NG, Saoji RG. Study of effectiveness of topical timolol maleate in the management of superficial infantile haemangioma. *J Evid Based Med Healthc*. 2018;5(41):2886-2889.

Variability of histologic pattern and transition zone length in Hirschsprung disease and associated factors

©Riana Pauline Tamba^{*1}, ©Novi Kurnia²

¹Division of Pediatric Surgery, Department of Surgery, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia

²Department of Surgery, Cipto Mangunkusumo Hospital, Jakarta, Indonesia

Received: 09/03/2026

Accepted: 19/05/2026

Published: 20/10/2026

Cite this article: Tamba RP, Kurnia N. Variability of histologic pattern and transition zone length in Hirschsprung disease and associated factors. *Surg Child.* 2026;3(4):113-120. doi:10.51271/SOC-0083

*Corresponding Author: Riana Pauline Tamba, rianatamba@gmail.com

ABSTRACT

Aims: Transition zone (TZ) pull-through is a major cause of postoperative obstructive symptoms in Hirschsprung disease (HD). This study aimed to describe TZ histology and its association with disease classification, age, and preoperative stoma.

Methods: A retrospective review of hematoxylin–eosin-stained pull-through specimens was performed. Samples were classified as complete transition zone (CTZ) or incomplete transition zone (ITZ). Nerve fiber diameter, interganglionic interval, and transition-zone length or extent were assessed in the submucosal and intermuscular layers.

Results: CTZ length ranged from 2 to 16 cm. In ITZ specimens, transition-zone features were observed over available specimen lengths of 3 to 33 cm, representing the observed extent rather than the complete anatomical transition-zone length. Overall, nerve fiber diameter, interganglionic interval, and transition-zone measurements were not consistently associated with disease classification, age, or stoma. In the ITZ group, intermuscular nerve fiber diameter was significantly associated with age ($p=0.004$) and stoma ($p=0.001$), while submucosal transition-zone extent was associated with stoma ($p=0.016$). Layer discrepancy and skip areas were also observed.

Conclusion: The TZ in HD shows marked histological and spatial variability, including layer discrepancy and skip areas. These findings support careful histopathological confirmation of the proximal resection margin, but should be interpreted cautiously due to the retrospective design, limited sample size, unequal groups, and incomplete specimen assessment.

Keywords: Hirschsprung disease, Hirschsprung disease/pathology, transition zone

INTRODUCTION

Hirschsprung disease (HD) is a congenital disorder characterized by the absence of parasympathetic ganglion cells in the myenteric and submucosal plexuses of the intestine, beginning at the internal anal sphincter and extending proximally for a variable length. Approximately 80% of cases are classified as short-segment HD involving the rectosigmoid region, 10% as long-segment HD extending to the splenic flexure, and 5-10% as total colonic aganglionosis.¹⁻⁴

The estimated global incidence of HD is 1 in 5,000 live births, with a male-to-female ratio of approximately 4:1.^{1,3,5,6} Data from the Division of Pediatric Surgery Cipto Mangunkusumo Hospital (RSCM) show that 40-50 new patients with HD are referred annually over the past five years.⁷ Clinical manifestations vary depending on patient age and the length of the aganglionic segment. The diagnosis is established by rectal biopsy, and definitive management consists of resection of the entire aganglionic segment along with the transition zone (TZ), followed by pull-through of ganglionated bowel to the anus.¹⁻³

Despite advances in surgical techniques, postoperative obstruction presenting as constipation or Hirschsprung-associated enterocolitis remains a frequent morbidity, reported in 9-40% of cases.^{8,9} Inadequate resection with pull-through performed within the TZ is recognized as a major cause of postoperative obstructive symptoms.^{5,9-14} A meta-analysis by Friedmacher and Puri¹⁵ demonstrated that 34.8% of patients requiring redo pull-through surgery showed abnormal rectal biopsy findings consistent with TZ pull-through. These findings highlight the clinical importance of accurately identifying the TZ during surgery.

Determining the precise length of the TZ intraoperatively is crucial for establishing an adequate resection margin. Kapur¹⁶ reported that in short-segment HD, the TZ generally does not exceed 5 cm, recommending resection at least 5 cm proximal to the first identified ganglionated segment. However, Coyle et al.¹⁷ found that 23% of cases had a TZ longer than 5 cm. Moreover, histopathologic



criteria defining the TZ remain controversial. Hypertrophic nerve fibers and hypoganglionosis are commonly used indicators, yet there is no consensus regarding threshold values for nerve hypertrophy or standardized criteria for hypoganglionosis.¹⁷⁻¹⁹

Factors influencing TZ length are not clearly understood. Some literature suggests that the TZ is longer in long-segment HD compared with short-segment HD.^{17,20} Other studies indicate that HD classification, age, and the presence of stoma may affect nerve fiber diameter within the TZ.^{18,19,21,22} Currently, contrast enema findings and macroscopic bowel appearance remain common guides for identifying the TZ.

Further characterization of the histologic features and length of the TZ is needed. Therefore, this study aimed to describe the histologic characteristics and length of the TZ in children with HD at RSCM and to analyze their association with disease classification, age, and the presence of stoma. The findings are expected to provide a scientific basis for determining an adequate resection margin and reducing postoperative morbidity related to TZ pull-through.

METHODS

Ethics

Ethical approval was obtained from the Ethics Committee of the Faculty of Medicine, Universitas Indonesia (Date: 05.04.2021, Decision No: KET-316/UN2.F1/ETIK/PPM.00.02/2021). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Study Design

This study was an observational descriptive-analytic study with retrospective histopathological slide review. The study was conducted at Cipto Mangunkusumo National Hospital (RSCM), Jakarta. Data were extracted from archived records of patients diagnosed with HD who underwent pull-through surgery between 2015 and 2020. Histopathological re-evaluation was performed at the Department of Anatomical Pathology, RSCM.

Study Population and Sample Size

The target population consisted of all pediatric patients diagnosed with HD. The accessible population included pediatric patients with HD who underwent pull-through surgery at the Pediatric Surgery Division, RSCM. Total sampling was applied. All eligible patients treated between 2015 and 2020 who met the inclusion and exclusion criteria were included in the study. Sample size calculation was performed using a formula for comparison of two means:

$$n = 2 \left(\frac{(Z\alpha + Z\beta)S}{x_1 - x_2} \right)^2$$

With the following parameters:

$\alpha=5\%$ ($Z\alpha=1.96$)

$\beta=20\%$ (power 80%; $Z\beta=0.84$)

Clinically meaningful difference (x_1-x_2)=10

Standard deviation (S)=15

The minimum required sample size was 35 subjects. However, total sampling was performed, and all eligible patients within the study period were included. Although the final sample exceeded the minimum calculated sample size for the overall analysis, the distribution between CTZ and ITZ groups was unequal. The relatively small CTZ group limited the statistical power for subgroup comparisons. Therefore, subgroup findings should be interpreted as exploratory.

Inclusion and Exclusion Criteria

Patients were included if they were aged ≤ 18 years and had a confirmed diagnosis of HD based on histopathological examination. Only patients who underwent definitive pull-through surgery at RSCM during the study period were considered eligible.

Patients were excluded if they were diagnosed with total colonic aganglionosis, had undergone previous bowel resection prior to definitive pull-through (including Hartmann procedure or prior pull-through), had incomplete resection specimens sent for pathological examination, had specimens that did not include the TZ, or had incomplete medical records or pathology archives.

Variables

The independent variables included HD classification (short-segment or long-segment), age at definitive surgery (<3 months, 3–12 months, >12 months), and presence of preoperative stoma (yes or no). The dependent variables were nerve fiber diameter within the TZ (μm), interganglionic interval within the TZ (μm), and TZ length (cm).

Sample Collection and Analysis

Resected colon specimens obtained during pull-through surgery were reviewed. Based on specimen completeness, samples were categorized into complete transition zone (CTZ) (specimens containing aganglionic, transition, and ganglionic zones) and incomplete transition zone (ITZ) (specimens lacking one or more zones but containing at least the TZ).

CTZ was defined as a resected specimen containing all three histological zones: aganglionic, transition, and ganglionated zones. In this group, both distal and proximal boundaries of the TZ could be identified, allowing complete measurement of TZ length.

ITZ was defined as a specimen that did not contain all three histological zones but still showed histological features of the TZ. This included specimens containing aganglionic and TZs, transition and ganglionated zones, or specimens in which the submitted tissue predominantly showed transition-zone features. Therefore, TZ length in the ITZ group represents the observed length of transition-zone features within the available specimen and should not be interpreted as the definitive full anatomical TZ length.

Hematoxylin-eosin-stained slides were re-evaluated. Photographs were obtained from the proximal and distal luminal ends and at every 1-cm interval along the longitudinal specimen. Images were captured at the submucosal and intermuscular layers using a high-definition Indomicon

camera mounted on an Olympus BX50 light microscope at 20× magnification. Each image measured 1920×1080 pixels (approximately 860×480 μm).

Histopathological evaluation was performed by the principal investigator in consultation with one board-certified anatomical pathologist. The presence of ganglion cells and nerve fibers was assessed in each image. Nerve fiber diameter and interganglionic interval were measured using ImageJ open-source software. Formal interobserver agreement analysis was not performed because the final histopathological interpretation was made by one anatomical pathologist.

TZ identification was based on established histopathological features, including the presence of hypertrophic nerve fibers and reduced ganglion cell density (hypoganglionosis), as described in previous studies. No strict numerical cutoff was applied for nerve fiber diameter, as variability has been reported in the literature; instead, interpretation was based on relative comparison across segments within the same specimen. For each sampled section, multiple representative fields were analyzed. Nerve fiber diameter was measured as the mean of selected representative fibers within each image rather than the maximum value, to reduce measurement bias.

Statistical Analysis

Categorical variables were presented as percentages, whereas continuous variables were expressed as mean±standard deviation if normally distributed or as median if not normally distributed. Data analysis was performed using SPSS version 20.0 for Windows. Descriptive analysis was conducted to assess patient characteristics. Inferential analysis was performed to evaluate the association between independent categorical variables (HD classification, age group, and stoma status) and continuous dependent variables (nerve fiber diameter, interganglionic interval, and TZ length). Shapiro-Wilk was used for normality testing. For comparisons between two groups, the independent t-test was used for normally distributed data, and the Mann-Whitney test was applied for non-normally distributed data. For age categories involving more than two groups, one-way ANOVA was used for normally distributed data, and the Kruskal-Wallis test was used when normality assumptions were not met. A p-value <0.05 was considered statistically significant.

RESULTS

A total of 119 definitive pull-through procedures for HD were performed at the Pediatric Surgery Division of Cipto Mangunkusumo National Hospital between January 2015 and December 2020. Following identification of these cases, histopathological reports were retrieved from the department of anatomical pathology archives, and corresponding medical records were reviewed for eligibility.

The types of definitive procedures included Soave-Boley, transanal endorectal pull-through, Duhamel Adang modification, Duhamel Martin modification, and laparoscopic-assisted transanal endorectal pull-through. Their distribution is presented in **Figure 1**.

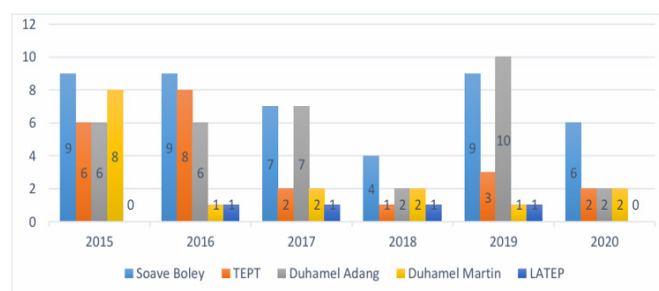


Figure 1. Definitive surgical treatment performed by pediatric surgery division RSCM 2015-2020

RSCM: Cipto Mangunkusumo National Hospital, TEPT: Transanal endorectal pull-through

Of 119 patients who underwent definitive pull-through surgery, 47 met the eligibility criteria and were included in the final analysis. The main reasons for exclusion were total colonic aganglionosis, previous bowel resection or redo pull-through, incomplete specimen submission, absence of TZ in the specimen, incomplete slide sampling, and unretrievable pathology slides. The final sample recruitment flow is summarized in **Figure 2**.

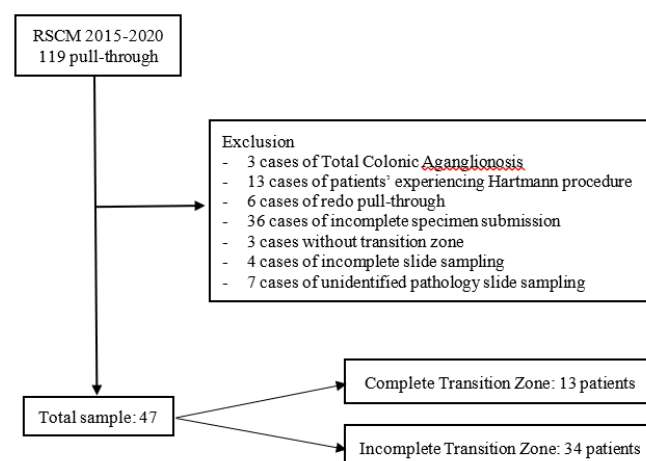


Figure 2. Sample recruitment flow RSCM: Cipto Mangunkusumo National Hospital

A total of 47 patients were included in the final analysis and were divided into two groups based on specimen completeness. The first group consisted of patients with a CTZ, defined as specimens containing aganglionic, transition, and ganglionated zones. The second group consisted of patients with an ITZ, defined as specimens that did not contain all three zones but included at least the TZ. Within the ITZ group, 15 specimens contained aganglionic and TZs, 8 contained transition and ganglionated zones, and 11 showed transition-zone features throughout the submitted tissue. Clinical characteristics of both groups are presented in **Table 1** (CTZ) and **Table 2** (ITZ).

In the CTZ group (n=13), 69.3% were male. The majority were classified as short-segment HD (92.3%), with only one long-segment HD (7.7%). The median age at definitive surgery was 5 months (range 1-17 months). Most patients underwent surgery before 3 months of age (38.5%) or between 3-12 months (38.5%). Preoperative stoma was present in 30.7% of patients. Transanal endorectal pull-through (TEPT) was the most commonly performed definitive procedure.

Table 1. Characteristics of complete transition zone group

Characteristic	Total patients	Median (min-max)
Gender		
Male (%)	9 (69.3%)	
Female (%)	4 (30.7%)	
Classification		
Short segment	12 (92.3%)	
Long segment	1 (7.7%)	
Age at definitive procedure (months)		
<3 months	5 (38.5%)	5 (1-17)
3-12 months	5 (38.5%)	
>12 months	3 (23%)	
Pre-operative stoma	4 (30.7%)	
Definitive procedure		
Soave-Boley	3 (23%)	
TEPT	5 (38.5%)	
Duhamel Adang	1 (7.7%)	
Duhamel Martin	2 (15.4%)	
LATEP	2 (15.4%)	

Categorical variables are presented as n (%). Age is presented as median (min-max) due to non-normal distribution. Min: Minimum, Max: Maximum, TEPT: Transanal endorectal pull-through

Table 2. Characteristics of incomplete transition zone group

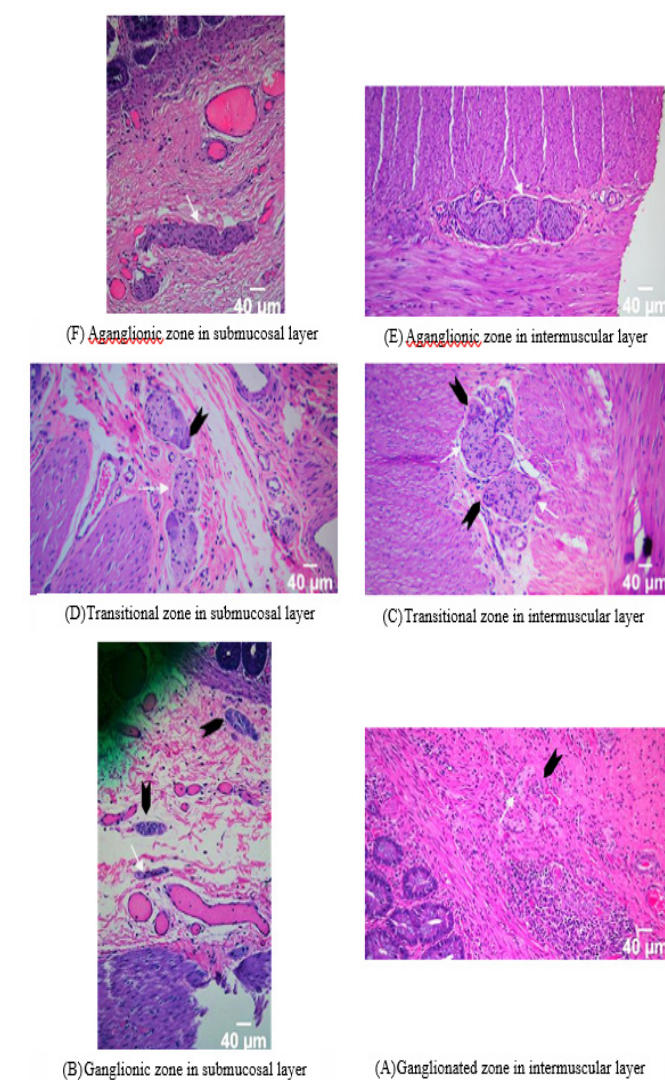
Characteristic	Total patients	Median (min-max)
Gender		
Male (%)	24 (70.6%)	
Female (%)	10 (29.4%)	
Classification		
Short segment	31 (91.1%)	
Long segment	3 (8.9%)	
Age at definitive procedure (months)		
<3 months	8 (23.5%)	8 (1-186)
3-12 months	11 (32.4%)	
>12 months	15 (44.1%)	
Pre-operative stoma	19 (55.9%)	
Definitive procedure		
Soave-Boley	12 (35.3%)	
TEPT	12 (35.3%)	
Duhamel Adang	5 (14.7%)	
Duhamel Martin	4 (11.8%)	
LATEP	1 (2.9%)	

Categorical variables are presented as n (%). Age is presented as median (min-max) due to non-normal distribution. Min: Minimum, Max: Maximum, TEPT: Transanal endorectal pull-through

In the ITZ group (n=34), 70.6% were male. Short-segment HD accounted for 91.1% of cases, and long-segment HD for 8.9%. The median age at definitive surgery was 8 months (range 1-186 months). The majority of patients underwent surgery after 12 months of age (44.1%). Preoperative stoma was present in 55.9% of patients. TEPT and Soave-Boley were the most frequently performed procedures.

Representative histopathological features of the aganglionic zone, TZ, and ganglionated zone in both submucosal and intermuscular layers are shown in **Figure 3**. Evaluation was

performed using hematoxylin-eosin (HE) staining at 20× magnification. In HE staining, ganglion cells and nerve fibers may appear similar in color, making morphological identification essential. Nerve fiber thickening was relatively easy to identify, particularly in the intermuscular layer. Ganglion cells were more difficult to recognize in HE-stained sections alone, especially in the TZ where ganglia and nerve fibers may appear closely approximated. In contrast, in the ganglionated zone, ganglion cells were more readily identified, while nerve fibers appeared thinner and less prominent.

**Figure 3.** Histopathological findings

In the CTZ group, the aganglionic segment length ranged from 2-11 cm, the TZ length ranged from 2-16 cm, and the ganglionated segment length ranged from 1-7 cm. In the ITZ group, transition-zone features were observed within available specimen lengths ranging from 3 to 33 cm. Because these specimens were incomplete, this range should be interpreted as the observed extent of transition-zone features rather than the complete anatomical TZ length. Overall, nerve fiber diameter was greater in the aganglionic zone compared with the TZ. The interganglionic interval was greater in the TZ compared with the ganglionated zone. The TZ defined by nerve hypertrophy in the submucosal and intermuscular layers generally overlapped within the same section; however, several differences were observed.

In the CTZ group, four patients showed the most distal ganglion in the submucosal layer located 2-3 cm distal to the most distal ganglion in the intermuscular layer. Conversely, only one patient showed the opposite finding, in which the most distal ganglion in the submucosal layer was located 2 cm proximal to the most distal ganglion in the intermuscular layer. In the ITZ group, the difference between the two layers ranged from 1-6 cm, with a median of 2 cm. The TZ length in the submucosal and intermuscular layers was also not symmetrical. In some specimens, the TZ in the submucosal layer was longer, whereas in others the intermuscular TZ was longer.

Areas with normal nerve fiber diameter were observed between regions of nerve hypertrophy, particularly in the submucosal layer, indicating that nerve hypertrophy within the TZ was not uniformly distributed. In the CTZ group, these “skip areas” ranged from 1-2 cm, whereas in the ITZ group they could reach up to 6 cm. The mean nerve fiber

diameter, interganglionic interval, and TZ length in each layer and their associations with HD classification, age, and presence of stoma in the CTZ group are presented in Table 3, while the results for the ITZ group are presented in Table 4.

In the CTZ group, nerve fiber diameter in the submucosal layer of long-segment HD was smaller than that observed in short-segment. In contrast, in the intermuscular layer, the opposite pattern was observed. The TZ length in short-segment HD was also shorter than that in long-segment. In the ITZ group, no difference in nerve fiber diameter was observed between short-segment and long-segment HD, either in the submucosal or intermuscular layers.

In both CTZ and ITZ groups, nerve fiber diameter tended to increase with age; however, a statistically significant association was only found in the intermuscular layer of the ITZ group. When comparing patients with and without preoperative stoma, nerve fiber diameter is tended to be greater

Table 3. Histological characteristics of complete transition zone

Variables	Submucosal layer						Intermuscular layer						Total	
	Ø µm	p	s µm	p	ℓ cm	p	Ø µm	p	s µm	p	ℓ cm	p	ℓ cm	p
Classification														
Short segment	49.7±6	0.023*	246.4±65	0.71*	3.5(1-15)	0.502*	51.3 (0-89)	0.348*	212.5 (60-269.8)	1.0#	3.5(0-13)	0.416#	6.5±3	0.89*
Long segment	32.4		272.6		6		60		207.8		6			
Age														
< 3 months	51.4±3	0.576\$	278.5±79	0.371\$	4±3	0.411\$	49.7±32	0.676\$	151.7±100	0.321\$	3±2	0.321\$	5±2	0.447\$
3-12 months	46.5±5		219.5±44		4.2±1		48.3(45-62)		220.8±52		5.4±3		6.8±2	
>12 months	46.4±15		246.2±55		7.6±6		56.4±3		208.8±12		7±5		8.6±6	
Stoma														
Yes	47.1±12	0.716*	234.2±51	0.616*	6.5±5	0.351*	57.9±3	0.446*	224±32	0.306*	6.2±4	0.378*	5(4-16)	0.938*
No	48.9±5		254.6±70		4.2±2		48.6±22		177±83		4.2±3		6.1±2	
Ø: Nerve fiber diameter (µm), s: Interganglionic distance (µm), ℓ: Length (cm). Data are presented as mean±standard deviation (SD) for normally distributed variables and median (min–max) for non-normally distributed variables. *Independent t-test, #Mann–Whitney test, \$ANOVA, & Kruskal–Wallis test. Min: Minimum, Max: Maximum														

Ø: Nerve fiber diameter (µm), s: Interganglionic distance (µm), ℓ: Length (cm). Data are presented as mean±standard deviation (SD) for normally distributed variables and median (min-max) for non-normally distributed variables. *Independent t-test, [#]Mann-Whitney test, ^sANOVA, & Kruskal-Wallis test. Min: Minimum, Max: Maximum

Table 4. Histological characteristics of incomplete transition zone

Variables	Submucosal layer						Intermuscular layer						Total	
	Ø µm	p	s µm	p	ℓ cm	p	Ø µm	p	s µm	p	ℓ cm	p	ℓ cm	p
Classification														
Short segment	53.9±11	0.962*	224.8±37	0.136*	10.1±6	0.073*	92.8±21	0.921*	232.9±53	0.412*	10.8±5	0.514*	11.9±5	0.386*
Long segment	54.2±7		190.9±21		17.3±10		91.5±19		206.7±18		17.6±15		19.6±12	
Age														
<3 months	55.9±14	0.792 ^s	216.7±29	0.235 ^s	8.1±4	0.082 ^s	78.5±8	0.004 ^s	222.6±66	0.468 ^s	9.3±3	0.186 ^s	10.3±4	0.308 ^s
3-12 months	54.3±13		209.2±30		8.6±6		86.1±21		219.4±62		9.6±5		11.3±5	
>12 months	52.5±7		233.8±43		12(6-29)		105.1±19		243.2±32		13.8±8		13(6-33)	
Stoma														
Yes	54.2±7	0.889*	229.2±40	0.201*	12(6-29)	0.016 [#]	104±20	0.001*	232.8±63	0.832*	12.8±8	0.165*	13(4-33)	0.143 [#]
No	53.6±14		212.5±31		7.9±5		78.3±9		228.9±43		9.6±4		6.1±2	
Ø: Nerve fiber diameter (µm), s: Interganglionic distance (µm), ℓ: Length (cm). Data are presented as mean±standard deviation (SD) for normally distributed variables and median (min–max) for non-normally distributed variables. *Independent t-test, #Mann–Whitney test, \$ANOVA, & Kruskal–Wallis test. Min: Minimum, Max: Maximum														

Ø: Nerve fiber diameter (µm), s: Interganglionic distance (µm), ℓ: Length (cm). Data are presented as mean±standard deviation (SD) for normally distributed variables and median (min-max) for non-normally distributed variables. *Independent t-test, [#]Mann-Whitney test, ^sANOVA, & Kruskal-Wallis test. Min: Minimum, Max: Maximum



in patients with stoma, particularly in the intermuscular layer but not consistently statistically significant. Regarding the interganglionic interval, no significant differences were observed according to HD classification, age group, or stoma status in either CTZ or ITZ groups. The interganglionic interval in the ganglionated zone is smaller than that in the TZ; however, the difference was minimal and not statistically significant. For transition-zone measurements, both CTZ and ITZ groups showed numerically greater values in long-segment HD, older patients, and patients with preoperative stoma; however, these findings were not consistently statistically significant and should be interpreted cautiously.

A post-hoc power consideration suggests that the study is underpowered for subgroup analyses, particularly for comparisons involving the CTZ group and long-segment HD. Therefore, non-significant findings should be interpreted cautiously, as they may reflect type II error.

DISCUSSION

The main finding of this study is that the TZ in HD shows marked histological and spatial variability. In the CTZ group, transition-zone length ranged from 2 to 16 cm, while in ITZ specimens, transition-zone features were observed over available specimen lengths of 3 to 33 cm.

A substantial number of patients were excluded due to incomplete specimens submitted for pathological examination. During pull-through procedures, particularly in patients with previous stoma creation, bowel resection may be performed in multiple segments. In many cases, only selected portions of the resected bowel are submitted for pathological confirmation of ganglion cells. This practice may result in reports indicating that the proximal margin still corresponds to the TZ and may contribute to incomplete resection.

Preoperative stoma was analyzed because it may represent staged management, delayed definitive surgery, previous bowel dilatation, and longer exposure to functional obstruction. However, stoma status should not be interpreted as a direct causal factor for transition-zone length.

Evaluation of pathology reports also demonstrated variability in intraoperative frozen-section biopsy techniques and specimen submission.²³ Frozen-section biopsy is essential for confirming the presence of ganglion cells before completing the pull-through procedure; however, the biopsy location and distance between biopsy sites varied among surgeons. Inconsistent sampling may lead to inaccurate identification of the TZ, which remains one of the major causes of postoperative obstructive symptoms.⁹

The prevalence of TZ pull-through observed in this study was 25.2%. This finding is comparable with the systematic review by Friedmacher and Puri,¹⁵ which reported a rate of 25.4% of TZ pull-through among patients requiring redo surgery. Similarly, Ghosh et al.²⁴ reported a prevalence of approximately 16% in a tertiary referral center. These findings suggest that incomplete resection of the TZ remains a persistent challenge in the surgical management of HD.

Several studies have suggested that evaluation of the proximal doughnut specimen during pull-through surgery may reduce the risk of TZ pull-through. Das et al.²⁵ demonstrated that examination of the proximal doughnut specimen can identify residual TZ tissue that may not be detected through localized biopsy. In addition, Murthi et al.²⁶ showed that routine circumferential doughnut biopsy during surgery prevented TZ pull-through in a substantial proportion of patients initially considered to have ganglionic bowel.

The demographic characteristics of the study population were consistent with previously reported epidemiological data. HD occurs predominantly in males, with a male-to-female ratio approaching 4:1, and short-segment HD represents the majority of cases.⁷ These findings are consistent with the distribution observed in the present study.

Histopathological evaluation showed that hypertrophic nerve fibers remain a useful criterion for identifying the TZ. However, interpretation of nerve hypertrophy should consider patient age. Kapur¹⁹ reported that nerve trunks larger than 40 μ m may also be observed in normal rectal tissue, particularly in children older than one year. The 40 μ m threshold was originally proposed by Monforte-Muñoz et al.¹⁸ as a diagnostic criterion for hypertrophic nerve trunks in HD.

An additional observation was the presence of skip areas in which histologically normal segments were identified within the TZ. These areas ranged from 1 to 6 cm and were most commonly observed in the submucosal layer. The discrepancy between the submucosal and intermuscular layers suggests that the TZ does not progress uniformly across bowel wall layers. In some cases, the submucosal layer appeared to extend more distally, whereas the opposite pattern was observed in others. This finding indicates that limited focal assessment may underestimate the true extent of the TZ. This finding further supports the importance of careful intraoperative evaluation and comprehensive pathological examination to avoid incomplete resection of the TZ.²⁷

The presence of histologically normal-appearing skip areas within the TZ has important clinical implications. These areas may create false reassurance during intraoperative frozen-section assessment if only limited focal biopsies are taken. Therefore, careful proximal margin assessment and consideration of circumferential or doughnut evaluation may help reduce the risk of transition-zone pull-through.²⁷

In long-segment HD, nerve hypertrophy may be less prominent or present only focally. Studies by Narayanan et al.²¹ and Alnajar et al.²² demonstrated that hypertrophic nerve fibers may not always be observed in long-segment HD. Nevertheless, hypertrophic nerve fibers were still identified in several long-segment HD in this study, suggesting that this feature can still support identification of the TZ when interpreted alongside other histological findings.

The distance between ganglion cells was also evaluated to assess hypoganglionosis. According to Kapur and Kennedy,²⁸ hypoganglionosis is characterized by reduced ganglion cell density and smaller ganglion size. Coyle et al.¹⁷ proposed



that ganglion spacing greater than 730 μm may indicate hypoganglionosis within the TZ. However, in the present study, ganglion spacing did not differ significantly between groups, suggesting that this parameter alone may not reliably distinguish the TZ.

The median TZ length in this study was comparable with findings reported by Coyle et al.,¹⁷ who described a median length of approximately 6 cm. In the ITZ group, transition-zone features were observed over longer available specimen lengths. However, because these specimens were incomplete, this finding should be interpreted as the observed extent of transition-zone features rather than the complete anatomical length of the TZ.

Multivariable analysis was not performed due to limited sample size and risk of model overfitting. Therefore, potential confounding between variables such as age, stoma status, and disease severity cannot be excluded.

Limitations

This study has several limitations. First, the retrospective design and the exclusion of a large proportion of the initial cohort may have introduced selection bias, particularly because several patients were excluded due to incomplete specimen submission, missing pathology records, incomplete slide sampling, or unretrievable slides. Second, the relatively small CTZ group and the unequal distribution between CTZ and ITZ groups limited the statistical power for subgroup analyses. Given the small CTZ group ($n=13$) and the very limited number of long-segment cases, the study is likely underpowered to detect moderate differences, increasing the risk of type II error. Third, the ITZ group was heterogeneous because the available specimens did not always contain all three histological zones; therefore, measurements in this group represent only the observed extent of transition-zone features within the submitted specimens rather than the complete anatomical transition-zone length. Fourth, histopathological evaluation was performed using slides prepared for routine clinical service, which included longitudinal sampling from only one side of the bowel, thereby the circumferential variability of the TZ could not be assessed. This may result in underestimation or overestimation of the true TZ length, particularly in the presence of skip areas. Finally, formal interobserver agreement analysis was not performed.

CONCLUSION

This study demonstrates substantial variability in the histological appearance and observed extent of the TZ in HD. The TZ may differ between the submucosal and intermuscular layers and may contain histologically normal-appearing skip areas, which has important implications for intraoperative margin assessment. These findings support careful histopathological confirmation of the proximal resection margin to reduce the risk of transition-zone pull-through. However, the results should be interpreted cautiously because of the retrospective design, limited sample size, unequal CTZ and ITZ groups, and incomplete circumferential assessment.

ETHICAL DECLARATIONS

Ethics Committee Approval

Ethical approval was obtained from the Ethics Committee of the Faculty of Medicine, Universitas Indonesia (Date: 05.04.2021, Decision No: KET-316/UN2.F1/ETIK/PPM.00.02/2021).

Informed Consent

This retrospective study used pre-existing anonymized patient data. No additional intervention was performed, and there was no direct patient contact. The study was approved by the Ethics Committee, and the requirement for written informed consent was waived by the ethics committee.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: RPT; Design: RPT; Control: RPT, NK; Resources: RPT, NK; Materials: RPT, NK; Data Collection and/or Processing: RPT, NK; Analysis and/or Interpretation: RPT, NK; Literature Review: NK; Writing the Article: RPT, NK; Critical Review: RPT, NK.

Acknowledgements

The authors would like to thank the Pediatric Surgery Division and the Department of Anatomical Pathology of Cipto Mangunkusumo National Hospital for their support in providing access to clinical records and histopathological specimens used in this study. We also acknowledge the Faculty of Medicine, Universitas Indonesia, for facilitating the conduct of this research. The authors are grateful to all staff members who contributed to data collection and technical assistance during the histopathological evaluation and manuscript preparation.

REFERENCES

1. Kartono D. Penyakit Hirschsprung. Jakarta: Sagung Seto; 2004.
2. Langer JC. Hirschsprung disease. In: Coran AG, Adzick NS, Krummel TM, Laberge J-M, Shamberger RC, Caldamone AA, editors. Pediatric Surgery. 7th ed. Saunders; 2012.
3. Langer JC. Hirschsprung disease. In: Holcomb GW, Murphy JP, St. Peter SD, editors. Holcomb and Ashcraft's Pediatric Surgery. Elsevier Inc.; 2020.
4. Menteri Kesehatan Republik Indonesia. Keputusan Menteri Kesehatan Republik Indonesia Nomor HK.01.07/MENKES/474/2017 tentang Pedoman Nasional Pelayanan Kedokteran Tata Laksana Penyakit Hirschsprung. 2017.
5. Taguchi T, Matsufuji H, Ieiri S. Hirschsprung's Disease and the Allied Disorders. Singapore: Springer; 2019.
6. Bradnock TJ, Knight M, Kenny S, Nair M, Walker GM. Hirschsprung's disease in the UK and Ireland: incidence and anomalies. *Arch Dis Child*. 2017;102(8):722-727. doi:10.1136/archdischild-2016-311872
7. Wiradeni A. Pengaruh panjang usus aganglionik terhadap morbiditas penderita penyakit Hirschsprung pasca operasi definitif di Indonesia [thesis]. Universitas Indonesia; 2020.



8. Kapur RP, Smith C, Ambartsumyan L. Postoperative pullthrough obstruction in Hirschsprung disease: etiologies and diagnosis. *Pediatr Dev Pathol.* 2020;23(1):40-59. doi:10.1177/1093526619890735
9. Langer JC, Rollins MD, Levitt M, et al. Guidelines for the management of postoperative obstructive symptoms in children with Hirschsprung disease. *Pediatr Surg Int.* 2017;33(5):523-526. doi:10.1007/s00383-017-4066-7
10. Puri P. Hirschsprung's Disease and Allied Disorders. 4th ed. 2019.
11. Ralls MW, Coran AG, Teitelbaum DH. Reoperative surgery for Hirschsprung disease. *Semin Pediatr Surg.* 2012;21(4):354-363. doi:10.1053/j.sempedsurg.2012.07.011
12. Ghose SI, Squire BR, Stringer MD, Batcup G, Crabbe DCG. Hirschsprung's disease: problems with transition-zone pull-through. *J Pediatr Surg.* 2000;35(12):1805-1809. doi:10.1053/jpsu.2000.19263
13. Soh HJ, Nataraja RM, Pacilli M. Prevention and management of recurrent postoperative Hirschsprung's disease obstructive symptoms and enterocolitis: systematic review and meta-analysis. *J Pediatr Surg.* 2018;53(12):2423-2429. doi:10.1016/j.jpedsurg.2018.08.024
14. Saadai P, Trappey AF, Goldstein AM, et al. Guidelines for the management of postoperative soiling in children with Hirschsprung disease. *Pediatr Surg Int.* 2019;35(8):829-834. doi:10.1007/s00383-019-04497-y
15. Friedmacher F, Puri P. Residual aganglionosis after pull-through operation for Hirschsprung's disease: a systematic review and meta-analysis. *Pediatr Surg Int.* 2011;27(10):1053-1057. doi:10.1007/s00383-011-2958-5
16. Kapur RP. Histology of the transition zone in Hirschsprung disease. *Am J Surg Pathol.* 2016;40(12):1637-1646. doi:10.1097/PAS.0000000000000711
17. Coyle D, O'Donnell AM, Tomuschat C, Gillick J, Puri P. The extent of the transition zone in Hirschsprung disease. *J Pediatr Surg.* 2019;54(11):2318-2324. doi:10.1016/j.jpedsurg.2019.04.017
18. Monforte-Muñoz H, Gonzalez-Gomez I, Rowland JM, Landing BH. Increased submucosal nerve trunk caliber in aganglionosis: a positive and objective finding in suction biopsies and segmental resections in Hirschsprung's disease. *Arch Pathol Lab Med.* 1998;122(8):721-725.
19. Kapur RP. Submucosal nerve diameter greater than 40 µm is not a valid diagnostic index of transition zone pull-through. *J Pediatr Surg.* 2016;51(10):1585-1591. doi:10.1016/j.jpedsurg.2016.06.007
20. Thakkar HS, Blackburn S, Curry J, et al. Variability of the transition zone length in Hirschsprung disease. *J Pediatr Surg.* 2020;55(1):63-66. doi:10.1016/j.jpedsurg.2019.09.056
21. Narayanan SK, Soundappan SS, Kwan E, Cohen RC, Charlton A, Cass DT. Aganglionosis with the absence of hypertrophied nerve fibres predicts disease proximal to rectosigmoid colon. *Pediatr Surg Int.* 2016;32(3):221-226. doi:10.1007/s00383-015-3835-4
22. Alnajar H, Murro D, Alsadi A, Jakate S. Spectrum of clinicopathological deviations in long-segment Hirschsprung disease compared with short-segment Hirschsprung disease: a single-institution study. *Int J Surg Pathol.* 2017;25(3):216-221. doi:10.1177/1066896916675729
23. Moore SW. Hirschsprung disease: current perspectives. *Open Access Surg.* 2016;9:39-50. doi:10.2147/OAS.S81552
24. Ghosh DN, Liu Y, Cass DT, Soundappan SSV. Transition zone pull-through in Hirschsprung's disease: a tertiary hospital experience. *ANZ J Surg.* 2017;87(10):780-783. doi:10.1111/ans.13745
25. Das K, Kini U, Babu MK, Mohanty S, D'Cruz AJ. The distal level of normally innervated bowel in long segment colonic Hirschsprung's disease. *Pediatr Surg Int.* 2010;26(6):593-599. doi:10.1007/s00383-010-2603-8
26. Murthi G, Townley O, Lindley R, Cohen M. Routine use of circumferential 'Doughnut' biopsy in pull through surgery for Hirschsprung's disease: advantages and limitations. *J Indian Assoc Pediatr Surg.* 2021;26(2):111. doi:10.4103/jiaps.JIAPS_43_20
27. De La Torre L, Wehrli LA. Error traps and culture of safety in Hirschsprung disease. *Semin Pediatr Surg.* 2019;28(3):151-159. doi:10.1053/j.sempedsurg.2019.04.013
28. Kapur RP, Kennedy AJ. Histopathologic delineation of the transition zone in short segment Hirschsprung disease. *Pediatr Dev Pathol.* 2013;16(4):252-266. doi:10.2350/12-12-1282-OA.1

Interdisciplinary management of children with cerebral palsy: neurological perspective on neurosurgical interfaces

©Nozila Majidova*¹, ©Yakutkhon Nabievna Madjidova²

¹Department of Neurology, Tashkent State Medical University, Tashkent, Uzbekistan

²Department of Neurology, Pediatric Neurology and Medical Genetics, Tashkent State Medical University, Tashkent, Uzbekistan

Received: 03/03/2026

Accepted: 01/06/2026

Published: 20/10/2026

Cite this article: Majidova N, Madjidova YN. Interdisciplinary management of children with cerebral palsy: neurological perspective on neurosurgical interfaces. *Surg Child.* 2026;3(4):121-125. doi:10.51271/SOC-0084

*Corresponding Author: Nozila Majidova, nozilamajidova28@gmail.com

ABSTRACT

Aims: The aim of this study was to evaluate clinical and functional outcomes of interdisciplinary neurological and neurosurgical management in children with cerebral palsy (CP), with an emphasis on surgical selection, functional outcomes, and long-term follow-up, including caregiver-reported measures.

Methods: This retrospective, single-center study included 150 consecutive children with CP who were managed between January 2018 and December 2025. Standardized neurological assessment, Gross Motor Function Classification System (GMFCS) classification, neuroimaging, and multidisciplinary evaluation were performed. Functional outcomes were assessed using the Gross Motor Function Measure (GMFM), pain scales, and caregiver-reported measures. Statistical analyses included the Chi-square test, Student's t-test, and logistic regression.

Results: The cohort included 150 patients (mean age 6.8 ± 3.2 years). GMFCS distribution was as follows: levels I–II (34%), III (26%), IV (24%), and V (16%). Neurosurgical consultation was required in 42% of patients. Surgery was performed in 61% of the cohort. Functional improvement ($GMFM \geq 10\%$) was significantly more common in the surgical group (48% vs 22%, $p < 0.01$). Pain reduction was reported in 64% versus 38% ($p < 0.05$). Caregiver-reported improvement was observed in 71% of cases. In 28% of cases, surgical intervention was avoided after neurological reassessment.

Conclusion: Interdisciplinary neurological–neurosurgical management improves patient selection, reduces unnecessary surgery, and enhances functional and quality-of-life outcomes in children with CP.

Keywords: Cerebral palsy, neurosurgery, neurology, interdisciplinary communication, treatment outcome, child

INTRODUCTION

Cerebral palsy (CP) remains the leading cause of permanent motor disability in childhood and comprises a heterogeneous group of non-progressive disorders that affect movement and posture.^{1,2} Despite its static neurological origin, CP exhibits a dynamic clinical course driven by growth, biomechanical changes, and secondary musculoskeletal complications.

The complexity of CP extends beyond motor dysfunction to include epilepsy, cognitive impairment, pain syndromes, feeding difficulties, and orthopedic deformities. These factors significantly influence functional prognosis and therapeutic strategies.^{3,4}

Effective management requires a coordinated, interdisciplinary approach involving pediatric neurologists, neurosurgeons, orthopedic surgeons, and rehabilitation specialists.^{5,6} However, variability remains in surgical indications, timing, and outcome expectations.

A major challenge in clinical practice is distinguishing dynamic neurological dysfunction from fixed structural pathology. Inappropriate surgical selection may lead to suboptimal outcomes or unnecessary interventions.^{7,8}

The present study aims to evaluate the effectiveness of interdisciplinary management, with an emphasis on neurological stratification, surgical decision-making, functional outcomes, and long-term follow-up.

METHODS

Ethics

The study protocol adhered to the ethical principles of the Declaration of Helsinki. This study was approved by the Ethics Committee of Tashkent State Medical University (Date: 10.09.2025, Decision No: 16).

Study Design

This retrospective, single-center observational cohort study was conducted in the Department of Pediatric Neurology and Medical Genetics at the Tashkent Pediatric Medical Institute (subsequently integrated into the structures of Tashkent State Medical University), Tashkent, Uzbekistan.

Study Population and Selection Criteria

Clinical and follow-up records of consecutive pediatric patients diagnosed with CP and managed at our institution from January 2018 to December 2025 were retrospectively reviewed. A total of 150 patients met the eligibility criteria. The inclusion criteria were: (1) a confirmed clinical diagnosis of CP, (2) age ≤ 18 years at enrollment, and (3) availability of complete multidisciplinary longitudinal clinical and follow-up data. Patients were excluded if they presented with progressive neurological disorders or genetic neurodegenerative diseases, or if their medical charts contained incomplete documentation.

Clinical, Functional, and Neuroimaging Assessments

All subjects underwent a standardized, comprehensive neurological examination at baseline and at follow-up visits. Assessment protocols included detailed characterization of muscle tone (stratified as spastic, dyskinetic, or mixed phenotypes), selective motor control testing, deep tendon reflex evaluation, and documentation of involuntary movements or associated neurological comorbidities. Functional motor status was classified using the expanded and revised Gross Motor Function Classification System (GMFCS).⁹ Additionally, baseline magnetic resonance imaging (MRI) of the brain and/or spine was systematically reviewed for all participants to delineate underlying structural neuroanatomical abnormalities.

Conservative and Interventional Rehabilitation Protocols

Enrolled patients participated in structured, individualized rehabilitation programs that included physical therapy (3–5 sessions per week), occupational therapy, and speech therapy tailored to each patient's needs. Intensive rehabilitation (≥ 5 sessions per week) was provided to 62% of the cohort. Before any neurosurgical consideration, focal muscle spasticity was managed with botulinum toxin type A (BTX-A) injections in 56% of patients, with the aim of maximizing passive range of motion and delaying or obviating the need for invasive surgical interventions.^{10,11}

Neurosurgical Evaluation and Multidisciplinary Selection

Patients were referred for neurosurgical consultation in the presence of progressive functional deterioration, fixed musculoskeletal contractures, refractory spasticity, or failure of conservative multimodal rehabilitation strategies. Final therapeutic decisions and surgical indications were established through a strict multidisciplinary consensus involving pediatric neurologists, neurosurgeons, and rehabilitation specialists.^{12,13}

Outcome Measures and Follow-Up Strategy

The primary functional outcome was defined as a $\geq 10\%$ improvement in the Gross Motor Function Measure (GMFM) score or documented functional motor stabilization. Secondary outcomes included pain reduction (measured

using the Visual Analogue Scale [VAS] or validated observational pain scales), caregiver-reported quality-of-life satisfaction, and perceived ease of daily care. The mean follow-up duration for the entire cohort was 2.8 ± 1.4 years.

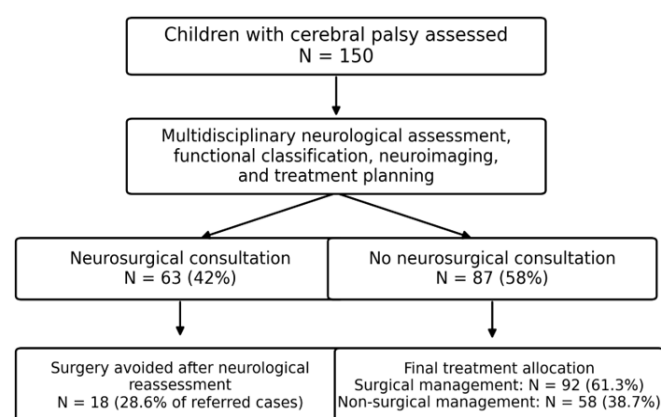
Statistical Analysis

Continuous variables were reported as means $\pm$ standard deviations (SD), and categorical variables as frequencies and percentages. Comparative analyses between the surgical and non-surgical sub-cohorts were conducted using Student's t-test for continuous variables and the Chi-square (χ^2) test for categorical variables. Logistic regression was used to identify independent predictors of optimal functional outcomes. Statistical significance was set a priori at a two-tailed p -value < 0.05 . All statistical modeling was performed using standard clinical software.

RESULTS

Demographic and Clinical Characteristics

A total of 150 pediatric patients with CP were included in the final analysis. The study flow, including multidisciplinary screening, neurosurgical consultation, and treatment allocation, is depicted in **Figure**.



Follow-up outcomes analyzed for the entire cohort (mean 2.8 ± 1.4 years)

Figure. Study flow diagram

A total of 150 patients with cerebral palsy were assessed. Neurosurgical consultation was obtained in 63 patients (42%). Surgery was performed in 92 patients (61%), while 58 patients (39%) received non-surgical management. Follow-up outcomes were analyzed for the entire cohort.

The mean age of the cohort was 6.8 ± 3.2 years. Of the 150 patients, 57.3% ($n=86$) were male and 42.7% ($n=64$) were female. The mean post-intervention longitudinal follow-up duration was 2.8 ± 1.4 years. Regarding CP phenotypic distribution, spastic CP was the most prevalent type, accounting for 68.0% ($n=102$) of cases, followed by dyskinetic CP at 19.3% ($n=29$) and mixed CP at 12.7% ($n=19$). Baseline functional status, stratified by the GMFCS, showed that 34.0% ($n=51$) of patients were at levels I–II, 26.0% ($n=39$) at level III, 24.0% ($n=36$) at level IV, and 16.0% ($n=24$) at level V. Baseline demographic and clinical characteristics are summarized in **Table 1**.

Neurosurgical Consultation and Treatment Allocation

Formal neurosurgical consultation was requested for 42.0% ($n=63$) of the cohort because of refractory spasticity or progressive musculoskeletal deterioration. After

**Table 1.** Baseline demographic and clinical characteristics of the pediatric cerebral palsy cohort (n=150)

Characteristic	Frequency (n)	Percentage (%)
Gender		
Male	86	57.3%
Female	64	42.7%
Age (years)		
Mean±SD (range)	6.8±3.2	(1.5–17.0)
Follow-up duration (years)		
Mean±SD	2.8±1.4	—
Cerebral palsy clinical phenotype		
Spastic	102	68.0%
Dyskinetic	29	19.3%
Mixed	19	12.7%
GMFCS		
Level I–II	51	34.0%
Level III	39	26.0%
Level IV	36	24.0%
Level V	24	16.0%
Preoperative conservative management		
Intensive rehabilitation (≥5 sessions/week)	93	62.0%
BTX-A therapy	84	56.0%

*Note: SD: Standard deviation, GMFCS: Gross Motor Function Classification System, BTX-A: Botulinum toxin type-A

rigorous interdisciplinary neurological and neurosurgical reassessment, surgical intervention was avoided in 28.6% (n=18 of 63) of these referred cases, as their dynamic tone abnormalities were deemed manageable with conservative optimization.

Ultimately, 61.3% (n=92) of the cohort underwent tailored neurosurgical procedures (surgical group), whereas the remaining 38.7% (n=58) received optimized multimodal conservative management (non-surgical group) (Table 2).

Functional Outcomes

Functional motor performance, measured by the GMFM, showed a significantly higher rate of clinically meaningful improvement (GMFM score increase ≥10%) in the surgical group than in the non-surgical group. Specifically, 47.8% (n=44 of 92) of patients in the surgical cohort achieved a ≥10% GMFM improvement, whereas only 22.4% (n=13 of 58) of the non-surgical cohort reached this threshold ($\chi^2=9.84$,

$p=0.002$). Furthermore, overall functional motor stabilization was achieved in 72.8% (n=67) of surgical patients versus 51.7% (n=30) of non-surgical patients ($p=0.009$).

Pain Assessment and Caregiver Perceptions

Post-therapeutic pain evaluation showed a significantly greater reduction in pain scores among patients who underwent surgical intervention. Marked pain reduction was reported in 64.1% (n=59 of 92) of the surgical group, compared with 37.9% (n=22 of 58) of the non-surgical group ($\chi^2=9.77$, $p=0.002$).

For subjective quality-of-life and caregiver burden assessments, high caregiver satisfaction was reported in 70.7% (n=106/150) of the cohort. Concurrently, families reported improved ease of daily care and routine positioning in 66.7% (n=100) of cases, indicating substantial improvements in home-based daily management.

DISCUSSION

The clinical management of CP requires a shift from isolated therapeutic protocols to highly coordinated, patient-centered, interdisciplinary models.^{1,5,14} The findings of this study demonstrate that integrated neurological and neurosurgical coordination significantly improves functional motor outcomes, procedural patient selection, and overall quality of life in children with CP. This evidence aligns with contemporary global guidelines, which emphasize that multimodal evaluation reduces the risk of inappropriate structural interventions in patients with complex, dynamic neurological phenotypes.^{7,15}

A primary finding of our investigation is that systematic neurological reassessment prevented unnecessary surgical interventions in nearly one-third (28.6%) of the referred patient population. In clinical practice, distinguishing between dynamic hypertonicity, fluctuating dystonia, and fixed structural musculoskeletal deformities remains a significant diagnostic challenge. Without rigorous neurological stratification, patients with dynamic components risk undergoing invasive bone or soft-tissue surgeries that may lead to suboptimal functional outcomes or secondary biomechanical destabilization.^{12,16}

Our comparative analysis revealed that functional motor outcomes—quantified using the GMFM—were significantly

Table 2. Clinical and functional outcomes stratified by therapeutic intervention (n=150)

Outcome measure	Surgical group (n=92)	Non-surgical group (n=58)	Statistical significance (p-value)
GMFM			
Significant Improvement (GMFM ≥10%)	44 (47.8%)	13 (22.4%)	0.002*
Functional stabilization	67 (72.8%)	30 (51.7%)	0.009*
Symptomatic pain outcomes			
Documented pain reduction	59 (64.1%)	22 (37.9%)	0.002*
Caregiver perceptions			
Caregiver satisfaction reported	65 (70.7%)	41 (70.7%)	0.999**
Improved ease of daily care	61 (66.3%)	39 (67.2%)	0.908**

Note: GMFM: Gross Motor Function Measure. Percentages are calculated based on the respective group totals (n=92 for surgical, n=58 for non-surgical).

* Statistically significant differences ($p<0.05$) evaluated via Chi-square (χ^2) test.

** Caregiver outcomes across the total cohort (n=150) demonstrated high overall satisfaction (70.7%, n=106) and ease of care (66.7%, n=100) with balanced distribution between sub-cohorts.



better in the cohort that underwent tailored neurosurgical procedures than in the non-surgical cohort. However, this statistical benefit was entirely contingent on meticulous preoperative patient selection. This clinical reality strongly supports previous literature establishing that surgical efficacy in pediatric CP is fundamentally dependent on precise neurological phenotyping and baseline motor capability. Furthermore, the strategic application of comprehensive rehabilitation programs and botulinum toxin type-A (BTX-A) injections played an indispensable role in optimizing patients' preoperative musculoskeletal baseline and, in several instances, successfully delayed or entirely obviated the need for invasive structural operations.^{10,11,17}

Beyond objective biomechanical metrics, caregiver-reported outcomes indicate that improvements in daily care routines, ease of positioning, and general health-related quality of life are viewed by families as equally vital to overall clinical success as gains in raw motor function.^{18,19} This finding supports broader integration of patient-reported outcome measures (PROMs) into standard pediatric neurosurgery and neurology registries. Finally, the clear utility of the GMFCS stratification observed in our cohort reinforces its status as an essential clinical tool for setting realistic therapeutic goals and accurately interpreting long-term longitudinal outcomes.^{9,20}

Limitations

This investigation has several limitations that must be acknowledged when interpreting the clinical data. First, its retrospective design introduces inherent selection biases and dependence on historical medical chart documentation. Second, as a single-center experience, the generalizability of our interdisciplinary structural protocols to other institutional settings with different resource allocations may be constrained. Third, the longitudinal follow-up period was relatively short, which warrants caution regarding the long-term sustainability of the documented motor gains and pain reduction scores over longer transitions into adolescence and adulthood.

CONCLUSION

Interdisciplinary neurological–neurosurgical management of children with CP significantly optimizes surgical candidate selection, reduces the incidence of unnecessary or premature invasive interventions, improves objective motor function, and substantially mitigates caregiver burden. Meticulous longitudinal neurological assessment remains the cornerstone of effective multidisciplinary treatment planning, ensuring that invasive surgical interventions are precisely targeted to pediatric subcohorts who will derive the greatest functional and symptomatic benefit.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study was approved by the Ethics Committee of Tashkent State Medical University (Date: 10.09.2025, Decision No: 16).

Informed Consent

This retrospective study used pre-existing anonymized patient data. No additional intervention was performed, and there was no direct patient contact. The study was approved by the Ethics Committee, and the requirement for written informed consent was waived by the ethics committee.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: NM, YNM; Design: NM, YNM; Control: NM, YNM; Resources: NM, YNM; Materials: NM, YNM; Data Collection and/or Processing: NM, YNM; Analysis and/or Interpretation: NM, YNM; Literature Review: NM, YNM; Writing the Article: NM, YNM; Critical Review: NM, YNM.

REFERENCES

- Novak I, Morgan C, Fahey M, et al. State of the evidence traffic lights 2019: systematic review of interventions for preventing and treating children with cerebral palsy. *Curr Neurol Neurosci Rep.* 2020;20(2):3. doi:10.1007/s11910-020-1022-z
- Graham HK, Rosenbaum P, Paneth N, et al. Cerebral palsy. *Nat Rev Dis Primers.* 2016;2:15082. doi:10.1038/nrdp.2015.82
- Fehlings D, Novak I, Berweck S, Hoare B, Stott NS, Russo RN. Botulinum toxin assessment, intervention and follow-up for paediatric upper limb hypertonicity: international consensus statement. *Eur J Neurol.* 2010;17(Suppl 2):38-56. doi:10.1111/j.1468-1331.2010.03127.x
- Himmelman K, Uvebrant P. The panorama of cerebral palsy in Sweden. XI. Changing patterns in the birth-year period 2003–2006. *Acta Paediatr.* 2014;103(6):618-624. doi:10.1111/apa.12614
- McIntyre S, Goldsmith S, Webb A, et al. Global prevalence of cerebral palsy: a systematic analysis. *Dev Med Child Neurol.* 2022;64(12):1494-1506. doi:10.1111/dmcn.15346
- Rosenbaum P, Paneth N, Leviton A, et al. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol Suppl.* 2007;109:8-14.
- Shore BJ, White N, Kerr Graham H. Surgical correction of equinus deformity in children with cerebral palsy: a systematic review. *J Child Orthop.* 2010;4(4):277-290. doi:10.1007/s11832-010-0268-4
- Tedroff K, Löwing K, Åström E. A prospective cohort study investigating gross motor function, pain, and health-related quality of life 17 years after selective dorsal rhizotomy in cerebral palsy. *Dev Med Child Neurol.* 2015;57(5):484-490. doi:10.1111/dmcn.12665
- Palisano RJ, Rosenbaum P, Bartlett D, Livingston MH. Content validity of the expanded and revised Gross Motor Function Classification System. *Dev Med Child Neurol.* 2008;50(10):744-750. doi:10.1111/j.1469-8749.2008.03089.x
- Heinen F, Desloovere K, Schroeder AS, et al. The updated European Consensus 2009 on the use of botulinum toxin for children with cerebral palsy. *Eur J Paediatr Neurol.* 2010;14(1):45-66. doi:10.1016/j.ejpn.2009.09.005
- Boyd RN, Hays RM. Current evidence for the use of botulinum toxin type A in the management of children with cerebral palsy: a systematic review. *Eur J Neurol.* 2001;8(Suppl 5):1-20. doi:10.1046/j.1468-1331.2001.00034.x
- Shore BJ, Yu X, Desai S, Selber P, Wolfe R, Graham HK. Adductor surgery to prevent hip displacement in children with cerebral palsy: the predictive role of the Gross Motor Function Classification System. *J Bone Joint Surg Am.* 2012;94(4):326-334. doi:10.2106/JBJS.J.02003



13. Häggglund G, Alriksson-Schmidt A, Lauge-Pedersen H, Rodby-Bousquet E, Wagner P, Westbom L. Prevention of dislocation of the hip in children with cerebral palsy: 20-year results of a population-based prevention programme. *Bone Joint J.* 2014;96-B(11):1546-1552. doi:10.1302/0301-620X.96B11.34385
14. Novak I, McIntyre S, Morgan C, et al. A systematic review of interventions for children with cerebral palsy: state of the evidence. *Dev Med Child Neurol.* 2013;55(10):885-910. doi:10.1111/dmcn.12246
15. Damiano DL. Rehabilitative therapies in cerebral palsy: the good, the not as good, and the possible. *J Child Neurol.* 2009;24(9):1200-1204. doi:10.1177/0883073809337919
16. McLaughlin J, Bjornson K, Temkin N, et al. Selective dorsal rhizotomy: meta-analysis of three randomized controlled trials. *Dev Med Child Neurol.* 2002;44(1):17-25. doi:10.1017/s0012162201001608
17. Novak I, Honan I. Effectiveness of paediatric occupational therapy for children with disabilities: a systematic review. *Aust Occup Ther J.* 2019; 66(3):258-273. doi:10.1111/1440-1630.12573
18. Voorman JM, Dallmeijer AJ, Schuengel C, Knol DL, Lankhorst GJ, Becher JG. Activities and participation of 9- to 13-year-old children with cerebral palsy. *Clin Rehabil.* 2006;20(11):937-948. doi:10.1177/0269215506069673
19. Fehlings D, Switzer L, Agarwal P, et al. Informing evidence-based clinical practice guidelines for children with cerebral palsy at risk of osteoporosis: a systematic review. *Dev Med Child Neurol.* 2012;54(2): 106-116. doi:10.1111/j.1469-8749.2011.04091.x
20. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol.* 1997;39(4):214-223. doi:10.1111/j.1469-8749.1997.tb07414.x

Evaluation of the pan-immune-inflammation value (PIV) as a prognostic and surgical decision-making marker in necrotising enterocolitis

İbrahim Yıldırım^{*1}, Sevim Ecem Ünlü Ballı², Gökhan Berktaş Bahadır², İlhami Sürer²

¹Department of Pediatric Surgery, Eskişehir City Hospital, Eskişehir, Türkiye

²Department of Pediatric Surgery, Gülhane Training and Research Hospital, University of Health Sciences, Ankara, Türkiye

Received: 13/07/2026

Accepted: 08/10/2026

Published: 20/10/2026

Cite this article: Yıldırım İ, Ünlü Ballı SE, Bahadır GB, Sürer İ. Evaluation of the pan-immune-inflammation value (PIV) as a prognostic and surgical decision-making marker in necrotising enterocolitis. *Surg Child.* 2026;3(4):126-132. doi:10.51271/SOC-0085

*Corresponding Author: İbrahim Yıldırım, dibrabimiyildirim@hotmail.com

ABSTRACT

Aims: Necrotizing enterocolitis (NEC) is a complex gastrointestinal disorder characterized by intestinal wall necrosis and a high mortality rate. The pan-immune-inflammation value (PIV) is a novel biomarker that reflects systemic inflammation and immune response, calculated from blood counts of neutrophils, platelets, monocytes, and lymphocytes. The present study evaluates the utility of PIV, an accessible and cost-effective parameter, in patients with NEC and investigates its potential to differentiate between those who recover with medical management and those who require surgical intervention.

Methods: Data from 793 patients admitted to the neonatal intensive care unit (NICU) between May 2018 and November 2022 were analyzed in a retrospective single-centre cohort study. The study included 40 NEC patients (stages 1, 2, and 3) and 753 control subjects (218 preterm, 535 term). The PIV [(neutrophils×platelets×monocytes)/lymphocytes] was calculated using complete blood count results obtained on postnatal days 0 and 1, at the NEC diagnosis time point (with postnatal-age-aligned control measurements) and, for NEC cases, at discharge/death. Because PIV values were markedly right-skewed and the Shapiro–Wilk test rejected normality in every group at every time point (all $p<0.001$), unadjusted between-group comparisons were performed with Kruskal–Wallis tests and Dunn–Bonferroni post-hoc tests; repeated measurements were analysed with a linear mixed model on log-transformed PIV for the time × group interaction. Receiver operating characteristic (ROC) analysis and multivariable logistic regression adjusted for gestational age and birth weight were performed to assess whether PIV predicts the need for surgery.

Results: PIV values showed significant differences between the study groups on postnatal day 0 ($p=0.001$) and day 1 ($p<0.001$). In contrast, PIV values were comparable between the groups at the time of NEC diagnosis ($p=0.358$) and did not differ significantly according to NEC stage at discharge or death ($p=0.875$). Contrary to the pattern suggested by group means, median PIV in the surgical NEC group was not the highest at any time point, and Dunn–Bonferroni post-hoc testing identified only one significant pair (Bell I vs term controls on day 1, $p=0.005$). ROC analysis showed no evidence of discriminative ability of PIV for the need for surgery at any evaluated time point (AUC 0.42–0.58; all 95% CIs included 0.50). PIV was not an independent predictor of surgery in multivariable logistic regression (day 1: OR 1.0004, 95% CI 0.9998–1.0009, $p=0.164$). The unadjusted NEC–control differences disappeared after adjustment for gestational age and birth weight (day 0: $p=0.858$; day 1: $p=0.339$), and the time × group interaction was not significant ($p=0.656$).

Conclusion: In this retrospective single-centre cohort, PIV did not discriminate NEC cases requiring surgery from those managed medically, and the apparent NEC–control differences were largely explained by gestational age and birth weight. These findings do not support the use of PIV as a stand-alone marker for surgical decision-making in NEC. Given the small surgical subgroup ($n=13$), the study is hypothesis-generating and the results require confirmation in larger, prospective, multicentre cohorts.

Keywords: Enterocolitis, necrotizing, pan-immune-inflammation value, surgical NEC, prematurity, biomarker

INTRODUCTION

Necrotising enterocolitis (NEC) is an acute characterised by necrosis of the intestinal wall.^{1,2} It is a necroinflammatory disease of the gastrointestinal tract, predominantly affects preterm infants with low birth



weight and is associated with high mortality and morbidity rates.^{1,2} Despite improvements in neonatal intensive care, NEC incidence and associated morbidity and mortality have remained substantial and may increase in some settings as survival of the smallest infants improves.¹ The pathogenesis of NEC is multifactorial and remains incompletely understood.^{2,3} An immature mucosal barrier, enteral feeding-related exposures, intestinal ischaemia, and intestinal dysbiosis are among the major factors implicated in disease development.^{2,4} Exaggerated Toll-like receptor 4 (TLR4)-mediated inflammatory signalling and intestinal barrier injury are central mechanisms in current models of NEC pathogenesis.^{3,5} Endothelial TLR4 activation may additionally impair intestinal microcirculatory perfusion and contribute to tissue injury.⁶

A major challenge in the clinical management of NEC is determining when medical treatment becomes insufficient and surgical intervention, such as primary peritoneal drainage or laparotomy, is required. Pneumoperitoneum due to intestinal perforation is an established indication for surgery, whereas identifying severe disease before perforation remains more difficult.^{7,8} Beyond perforation, decisions regarding surgical intervention rely on the evolution of clinical, laboratory, and radiological findings, and no single biomarker currently defines the optimal timing of surgery.^{9,10}

The pan-immune-inflammation value (PIV) is a composite index derived from neutrophil, platelet, monocyte, and lymphocyte counts obtained from a routine complete blood count. It is calculated as (neutrophil count $\times$ platelet count $\times$ monocyte count) divided by lymphocyte count.¹¹ PIV has gained attention as a systemic immune-inflammatory prognostic marker in several oncological settings.^{11,12} Neutrophils and platelets participate in pro-inflammatory and thrombo-inflammatory processes, monocytes contribute to innate immune and macrophage responses, and lymphocytes reflect adaptive immune function; these cell populations are also developmentally regulated in preterm infants and altered during NEC.^{13,14} Despite its use in other clinical contexts, the role of PIV in NEC has not been established. The objective of this study is to evaluate whether PIV can serve as an effective biomarker for the early detection and management of severe NEC in preterm infants.

This study aims to evaluate the prognostic value of PIV, a simple, cost-effective, reproducible parameter routinely assessed, in patients with NEC. Additionally, the study investigates whether PIV could serve as a novel biomarker to differentiate NEC cases responsive to medical treatment (stages 1 and 2) from those requiring surgical intervention (stage 3).

METHODS

This retrospective single-centre cohort study reviewed medical records of newborn patients admitted to the Neonatal Intensive Care Unit at Gülhane Training and Research Hospital between May 2018 and November 2022. The study protocol was approved by the Ethics Committee for Non-interventional Clinical Researches at Gülhane Training and Research Hospital, University of Health Sciences (Date:

17.05.2023, Decision No: 2023/119), and the research adhered to the principles of the Declaration of Helsinki. Data were collected retrospectively using the hospital information management system (FONET HBYS). As no additional diagnostic or therapeutic interventions were performed, the ethics committee waived the requirement for informed consent.

A total of 793 newborn medical records were reviewed. The study group consisted of 40 patients diagnosed with necrotising enterocolitis (NEC) according to Bell staging criteria, based on clinical and radiological findings.¹⁵ These patients were further categorized by disease severity: Bell stage 1 (n=15), stage 2 (n=12), and stage 3 (surgical NEC) (n=13). The control group included 753 newborns without an NEC diagnosis. Control cases were classified by gestational age as preterm (born before 37 weeks' gestation; n=218) or term (born at 37 weeks' gestation or later; n=535). Exclusion criteria included major congenital anomalies, incomplete clinical records, or missing complete blood count data.

For each patient, data collected included birth weight, gestational age, mode of delivery, maternal age, length of hospital stay, surgical intervention, and mortality status. The PIV was calculated as (neutrophil $\times$ platelet $\times$ monocyte)/lymphocyte, using values from the complete blood count.¹¹

To evaluate temporal changes in the inflammatory response, haematological parameters were analyzed at four time points: the initial complete blood count at birth (postnatal day 0), the count within the first 24 hours (postnatal day 1), the count at diagnosis, and the final assessment. In the NEC group, the time of diagnosis corresponded to the day clinical suspicion arose and treatment began. For the control group, routine complete blood count results from postnatal days 5 to 12 were standardized to align with the mean diagnosis time in NEC cases (8.32 postnatal days). For NEC cases, the final assessment used the most recent laboratory results prior to discharge or, for deceased patients, prior to death.

An initial sample-size calculation was conducted using G*Power (Version 3.1.9.7) based on a one-way analysis of variance (ANOVA) model. With an assumed effect size of 0.65, a type I error probability (α) of 0.05, and 80 percent power, the minimum required sample size was 32 patients; under that original planning scenario, 40 NEC cases corresponded to 83.15 percent power. This calculation does not quantify power for the revised primary surgical-versus-medical ROC and multivariable analyses; given the small surgical subgroup (n=13), those analyses are considered exploratory.

Statistical Analysis

The data analyses were conducted using IBM SPSS Statistics for Windows, Version 26.0 (IBM Corp., Armonk, NY, USA) and R version 4.3.1 (packages pROC, lme4 and logistf). The distribution of all continuous variables was examined before analysis. Normality was assessed within each study group using the Shapiro–Wilk test together with visual inspection of histograms and Q–Q plots and evaluation of skewness. PIV was markedly right-skewed and the Shapiro–Wilk test rejected normality in every group at every time point



($W=0.48-0.84$, all $p<0.001$); Levene's test also indicated heterogeneity of variances on postnatal day 1 ($p=0.026$) and at diagnosis ($p<0.001$). For unadjusted group comparisons of PIV, parametric tests were therefore considered inappropriate and non-parametric methods were used. Continuous variables are presented as median (interquartile range, Q1–Q3); means with standard deviations are additionally reported for comparability with earlier literature. Two groups were compared with the Mann–Whitney U test and three or more groups with the Kruskal–Wallis H test, with epsilon-squared (ϵ^2) as the effect size. Where the Kruskal–Wallis test was significant, pairwise comparisons were performed with Dunn's test and Bonferroni adjustment for multiple comparisons. Because the four PIV measurements were obtained from the same infants at unequal and partly missing time points, repeated measurements were analysed with a linear mixed-effects model on log-transformed PIV with a random intercept per patient; the $\text{time} \times \text{group}$ interaction was tested by likelihood-ratio comparison against a main-effects model, and the Friedman test was used as a sensitivity analysis in patients with complete data. The ability of PIV to predict the need for surgery within the NEC cohort was evaluated by receiver operating characteristic (ROC) analysis; the area under the curve (AUC) is reported with bootstrap 95% confidence intervals (2,000 resamples), and optimal cut-off values were derived from the Youden index. Univariable and multivariable binary logistic regression models were fitted to assess whether PIV predicts surgery independently of gestational age and birth weight. To address the potential confounding effect of gestational age and birth weight on the NEC–control comparison, linear models on log-transformed PIV were fitted with and without adjustment for these two variables. A two-tailed p-value of <0.05 was considered statistically significant in all analyses.

RESULTS

A total of 40 neonates diagnosed with NEC based on the Bell criteria, along with 753 control subjects without NEC, were incorporated into the study.

Within the NEC cohort, 52.5 percent ($n=21$) of cases were female, and 47.5 percent ($n=19$) were male. Conversely, in the control cohort, females comprised 45.3 percent, and males 54.7 percent. The rate of cesarean section was 90.0 percent ($n=36$) within the NEC group, compared to 60.4 percent in the control group.

The mean birth weight was 1800.00 ± 750.75 grams in the NEC group and 2991.07 ± 627.47 grams in the control group. The mean gestational age was 31.35 weeks in the NEC group and 37.29 weeks in the control group. The average duration of hospital stay was 43.75 ± 31.55 days for the NEC group and 8.86 ± 8.28 days for the control group. There was no statistically significant difference in maternal age between the groups (NEC: 29.20 ± 5.70 years; control: 28.45 ± 5.83 years).

Among all patients monitored in the neonatal intensive care unit (NICU) during the study period, the overall incidence of NEC was 5.04 percent. Among preterm neonates, the incidence was 14.5 percent. Surgical intervention was necessitated in 32.5 percent of NEC cases ($n=13$). The overall mortality rate was 12.5 percent, with mortality in the surgical NEC subgroup reaching 23.07 percent. The mean duration until the resumption of enteral feeding in NEC cases was 8.1 days.

Table 1 provides a comparison of PIV levels, calculated at various time points, between the groups.

When PIV was analysed with methods appropriate to its distribution, group differences were significant on postnatal day 0 ($H=18.311$, $p=0.001$, $\epsilon^2=0.038$) and postnatal day 1 ($H=19.876$, $p<0.001$, $\epsilon^2=0.034$), but not at the time of diagnosis ($H=4.372$, $p=0.358$); among the NEC stages, no difference was observed at discharge/death ($H=0.268$, $p=0.875$). Effect sizes were small at the three time points for which ϵ^2 was reported. It is important to note that the ranking of groups differed from that suggested by the arithmetic means: although the surgical NEC group had the highest mean PIV on postnatal day 1 (1418.09 ± 1218.45), its median (282.6) was lower than that of both control groups, indicating that the high mean was driven by a small number of extreme values rather than by a systematic elevation. Similarly, the difference at diagnosis that appeared significant under one-way ANOVA ($p=0.004$) was not confirmed by the Kruskal–Wallis test ($p=0.358$). Dunn's test with Bonferroni adjustment identified only one significant pairwise difference across all comparisons, namely Bell stage I versus term controls on postnatal day 1 ($p=0.005$); no post-hoc comparison involving the surgical NEC group reached significance.

Diagnostic Performance of PIV for the Need for Surgery

Within the NEC cohort ($n=40$; 13 surgical, 27 medical), ROC analysis was performed for each time point with sufficient surgical and medical data (postnatal day 0, postnatal day

Table 1. PIV levels by group and time point, median (Q1–Q3)

Measurement time	Bell stage I	Bell stage II	Bell stage III/ surgery	Term controls	Preterm controls	H	p*	ϵ^2
Postnatal day 0	151.0 (92.0–302.0) n=13	516.0 (163.5–1012.0) n=11	142.0 (61.5–446.8) n=12	761.8 (276.6–1606.6) n=271	786.6 (330.6–1452.0) n=75	18.311	0.001	0.038
Postnatal day 1	190.2 (47.2–303.3) n=13	399.1 (267.5–642.0) n=8	282.6 (114.7–1602.3) n=12	701.8 (368.7–1264.3) n=312	533.0 (298.5–1009.2) n=127	19.876	<0.001	0.034
At diagnosis	522.3 (80.4–773.7) n=15	596.5 (96.5–1522.0) n=11	215.8 (133.1–1090.1) n=11	349.7 (210.9–639.0) n=341	315.9 (156.5–529.5) n=154	4.372	0.358	0.001
Discharge/death	137.3 (92.0–470.2) n=15	145.3 (63.8–540.4) n=9	149.6 (51.3–375.3) n=10	–	–	0.268	0.875	–

* Kruskal–Wallis H test. ϵ^2 : epsilon-squared effect size (0.01 small, 0.06 medium, 0.14 large). Medians are reported because PIV was non-normally distributed in all groups (Shapiro–Wilk $p<0.001$). Bold p values are statistically significant. For controls, the 'at diagnosis' comparison uses routine complete blood counts from postnatal days 5–12 aligned to the mean NEC diagnosis time; discharge/death values were available only for NEC cases, so that comparison is among Bell stages only. PIV: Pan-immune-inflammation value, NEC: Necrotising enterocolitis



1, and diagnosis; **Table 2**). PIV showed no evidence of discriminative ability for the need for surgical intervention at any evaluated time point: AUC values ranged from 0.422 to 0.575 and all bootstrap 95% confidence intervals included 0.50. Corresponding Mann–Whitney U comparisons between surgical and medical NEC cases were also non-significant ($p=0.460$, 0.489 and 0.855 for day 0, day 1 and diagnosis, respectively). The cut-off values derived from the Youden index were unstable and clinically uninformative, with either very low sensitivity (8.3% on day 0) or very low specificity (23.8% on day 1).

Multivariable Analysis

Because gestational age, birth weight and prematurity may influence both PIV and the need for surgery, univariable and multivariable logistic regression models were fitted within the NEC cohort (**Table 3**). Neither PIV on postnatal day 1 nor PIV at diagnosis was associated with the need for surgery in univariable analysis, and PIV remained non-significant after adjustment for gestational age and birth weight (day 1: OR 1.0004, 95% CI 0.9998–1.0009, $p=0.164$). Gestational age and birth weight were likewise not independent predictors. It should be emphasised that with only 13 surgical events and three covariates the number of events per variable was approximately 4, well below the conventionally recommended minimum of 10; these models are therefore substantially underpowered and their estimates should be regarded as exploratory.

Effect of Gestational Age and Birth Weight on the NEC–Control Comparison

NEC cases were considerably more preterm and of lower birth weight than controls (median gestational age 31.0 versus 38.0 weeks and median birth weight 1795 versus 3030 g; both $p<0.001$), and PIV correlated positively with both gestational age ($\rho=0.186$, $p<0.001$) and birth weight ($\rho=0.195$, $p<0.001$) on postnatal day 1. In linear models on log-transformed PIV, the unadjusted NEC–control difference was significant on postnatal day 0 ($\beta=-0.989$, $p<0.001$) and day 1 ($\beta=-0.770$, $p<0.001$), but both associations disappeared after adjustment for gestational age and birth weight (day 0: $\beta=-0.050$, $p=0.858$; day 1: $\beta=-0.230$, $p=0.339$). After adjustment, the NEC–control differences were no longer

statistically significant, suggesting that gestational age and birth weight substantially confounded the unadjusted comparison.

Repeated Measurements and Time×Group Interaction

A linear mixed-effects model on log-transformed PIV with a random intercept per patient was fitted to the four time points within the NEC cohort (140 observations from 40 patients). Neither the main effect of time nor the main effect of surgical status reached significance, and the time×group interaction was not significant (likelihood-ratio test $\chi^2=1.616$, $df=3$, $p=0.656$). A Friedman test restricted to patients with complete data at all four time points ($n=24$) likewise showed no significant temporal change ($\chi^2=1.850$, $p=0.604$). Thus the temporal trajectory of PIV did not differ between surgical and medical NEC cases.

DISCUSSION

To the best of current knowledge, this study is the first to evaluate the relationship between the PIV, disease stage, and the need for surgery in newborns with NEC. The findings do not support PIV as a stand-alone marker of the need for surgery and indicate that unadjusted NEC–control differences are strongly influenced by gestational age and birth weight. The overall NEC incidence of 5.04 per cent in the NICU and the incidence of 14.5 per cent among preterm newborns fall within the broad range reported across neonatal populations, although published estimates vary substantially according to gestational age, birth-weight thresholds, and case definitions.^{16,17} The need for surgical treatment (32.5 per cent), overall mortality (12.5 per cent), and mortality in the surgical NEC subgroup (23.07 per cent) are broadly comparable with contemporary outcome data, although several systematic and multicentre reports have described higher mortality, particularly among surgical NEC cases.^{18,19} Such variation likely reflects differences in case mix, disease definitions, gestational-age distribution, and centre-level practice.

A central question of this study was whether PIV can identify NEC cases that will require surgery. Although the arithmetic mean PIV on postnatal day 1 was highest in the

Table 2. ROC analysis of PIV for predicting the need for surgery within the NEC cohort

Time point	n (surgical/medical)	AUC	95% CI (bootstrap)	Cut-off	Sensitivity	Specificity	p†
Postnatal day 0	36 (12/24)	0.422	0.204–0.647	>4012.0	8.3%	100.0%	0.460
Postnatal day 1	33 (12/21)	0.575	0.363–0.786	>48.5	100.0%	23.8%	0.489
At diagnosis	37 (11/26)	0.479	0.276–0.680	>130.7	81.8%	34.6%	0.855

† Mann–Whitney U test comparing surgical and medical NEC cases. ROC: Receiver operating characteristic, PIV: Pan-immune-inflammation value, NEC: Necrotising enterocolitis, AUC: Area under the receiver operating characteristic curve, CI: Confidence interval (2,000 bootstrap resamples). Cut-off values were derived from the Youden index. All confidence intervals include 0.50, indicating no discriminative ability.

Table 3. Univariable and multivariable logistic regression for the need for surgery within the NEC cohort

Variable	Univariable OR (95% CI)	p	Multivariable OR (95% CI)	p
PIV, postnatal day 1	1.0003 (0.9998–1.0009)	0.250	1.0004 (0.9998–1.0009)	0.164
PIV, at diagnosis	0.9998 (0.9992–1.0004)	0.591	–	–
Gestational age (weeks)	1.012 (0.849–1.206)	0.897	0.845 (0.580–1.229)	0.378
Birth weight (g)	1.0004 (0.9995–1.0014)	0.336	1.0017 (0.9996–1.0038)	0.109

NEC: Necrotising enterocolitis, OR: Odds ratio, CI: Confidence interval, PIV: Pan-immune-inflammation value. Multivariable model includes PIV on postnatal day 1, gestational age and birth weight ($n=33$, 12 events; events per variable =4). No variable reached statistical significance.



Bell stage III group, this pattern was not confirmed when methods appropriate to the highly skewed distribution of PIV were applied: the median value in the surgical group was lower than in either control group, no post-hoc comparison involving the surgical group was significant, and ROC analysis demonstrated no discriminative ability at any time point (AUC 0.42–0.58, all 95% CIs including 0.50). PIV also failed to predict the need for surgery in multivariable models adjusted for gestational age and birth weight. Taken together, these analyses indicate that the initial impression of an early inflammatory signal in surgical NEC was driven by a small number of extreme values rather than by a systematic difference. TLR4-mediated inflammatory signalling and intestinal barrier disruption are central to NEC pathogenesis.⁵ Experimental and human data also implicate TLR4-dependent intestinal lymphocyte influx in NEC.²⁰ Nevertheless, the present data do not support PIV as a quantitative marker of these processes in a form that is clinically useful for surgical decision-making.

Median PIV at the time of diagnosis was numerically lower in the Bell stage III group than in stage I and stage II cases, but this difference did not reach statistical significance (Kruskal–Wallis $p=0.358$), and the significant result obtained with one-way ANOVA ($p=0.004$) was not robust to the violation of normality. In the full-text analysis by Feng et al.,²¹ patients who underwent surgical intervention had significantly lower Systemic Immune-Inflammation Index (SII) values than those who received medical treatment (158.84 vs. 332.92, $p<0.001$). Although SII and PIV are distinct indices, both incorporate neutrophil, platelet, and lymphocyte counts, while PIV additionally incorporates monocytes; they therefore capture partly overlapping information from circulating inflammatory cell populations.^{11,21} The present study is therefore directionally consistent with that full-text analysis but does not reproduce its statistical significance, which may reflect the substantially smaller surgical subgroup available here.

Thrombocytopenia and abnormalities in neutrophil and monocyte counts are well described in NEC and may provide diagnostic or prognostic information, particularly in severe disease.²² Peripheral consumption, cellular redistribution, and inflammatory-cell trafficking are plausible contributors to these haematological changes.^{14,22} Because neutrophils, platelets, and monocytes are all in the numerator of the PIV formula, a reduction in any of these cell populations can mathematically lower PIV even in the presence of severe inflammation.¹¹ Therefore, the low PIV levels observed at diagnosis in the surgical NEC group may reflect consumption or redistribution of circulating immune and haematological cells rather than a reduction in the inflammatory response.^{14,22} However, because this mechanism was not directly investigated in the present study, this interpretation should be regarded as a pathophysiological hypothesis.

PIV values were higher in term than in preterm control infants on postnatal day 1, and PIV correlated positively with both gestational age and birth weight across the whole cohort. Immaturity of innate and adaptive immune responses in preterm newborns, together with gestational- and postnatal-age-related variation in neonatal blood-cell

counts, provides a biologically plausible explanation for this observation.^{13,23} Importantly, this maturity dependence has direct implications for the interpretation of the NEC–control comparison: because NEC cases were markedly more preterm and of lower birth weight than controls, the unadjusted differences in PIV were substantially confounded by maturity and were no longer statistically significant after adjustment for gestational age and birth weight. Future evaluations of PIV in this population should therefore account for gestational age and birth weight, for example through matching and/or multivariable adjustment.

Longitudinal or dynamic PIV assessment has shown prognostic relevance in other disease settings, raising the possibility that PIV trajectory may be more informative than a single measurement.²⁴ This hypothesis was tested directly in the present study using a linear mixed-effects model, and the time×group interaction was not significant ($p=0.656$), nor was there evidence of significant temporal change in a Friedman sensitivity analysis ($p=0.604$). The available data therefore do not support the notion that serial PIV measurement adds discriminative information about the need for surgery, although the limited number of patients with complete serial data ($n=24$) leaves this question only partially answered.

Limitations

This study has several important limitations. First and most critically, only 13 infants required surgery. This small subgroup severely limits statistical power: with three covariates the multivariable model had approximately four events per variable, below the traditional 10-events-per-variable rule of thumb for logistic regression.²⁵ The bootstrap confidence intervals around the AUC estimates were correspondingly wide (spanning up to 0.44 AUC units). The study was therefore unable to detect any but very large effects, and the negative findings reported here should be interpreted as an absence of evidence rather than as definitive evidence of absence. Second, the NEC and control groups differed substantially in gestational age and birth weight; although we addressed this by adjusted analysis, residual confounding cannot be excluded, and a gestational-age-matched control design would be preferable. Third, the retrospective, single-centre design restricts generalisability and precludes standardisation of the timing of blood sampling. Fourth, to compare haematological parameters at the time of diagnosis with those of the control group, routine complete blood count results obtained between postnatal days 5 and 12 in the control group were used, based on the mean time of diagnosis in NEC cases (postnatal day 8.32). Because neonatal haematological reference ranges vary with both gestational and postnatal age, this approach may have introduced measurement bias.²³ Fifth, serial data were incomplete, with only 24 patients contributing measurements at all four time points, which further limited the repeated-measures analysis. Finally, PIV was evaluated in isolation; its potential incremental value when combined with clinical, radiological, and other laboratory findings was not assessed.

CONCLUSION

In this retrospective single-centre cohort, the PIV did not discriminate between NEC cases that required surgery and



those managed medically. ROC analysis did not demonstrate discriminative ability at any of the three evaluable time points (postnatal day 0, postnatal day 1, and diagnosis), PIV was not an independent predictor of surgery in multivariable models, and the temporal trajectory of PIV did not differ between surgical and medical cases. The unadjusted PIV differences between NEC cases and controls were no longer statistically significant after adjustment for gestational age and birth weight, indicating substantial confounding by maturity.

PIV remains an attractive candidate biomarker in principle because it can be calculated from a routine complete blood count without an additional laboratory assay. However, the present data do not support its use as a stand-alone marker for surgical decision-making in NEC, and no cut-off value could be identified with acceptable sensitivity and specificity. Given the small surgical subgroup, these findings are hypothesis-generating rather than definitive. Adequately powered, prospective, multicentre studies with gestational-age-matched controls, standardised sampling times, and prespecified analysis plans are required before any conclusion can be drawn about the clinical utility of PIV in this setting; such studies should also examine whether PIV contributes incremental information when combined with established clinical, radiological, and laboratory criteria.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study protocol was approved by the Ethics Committee for Non-interventional Clinical Researches at Gülhane Training and Research Hospital, University of Health Sciences (Date: 17.05.2023, Decision No: 2023/119).

Informed Consent

This retrospective study used pre-existing anonymized patient data. No additional intervention was performed, and there was no direct patient contact. The study was approved by the Ethics Committee, and the requirement for written informed consent was waived by the ethics committee.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: IY, IS; Design: IY, IS; Control: IY, IS; Resources: IY, SEUB, GBB, IS; Materials: IY, SEUB, GBB, IS; Data Collection and/or Processing: IY, SEUB; Analysis and/or Interpretation: GBB, IS; Literature Review: GBB, IS; Critical Review: GBB, IS; Final Approval: IY, GBB, IS; Accountability: IY, IS.

REFERENCES

1. Neu J, Walker WA. Necrotizing enterocolitis. *N Engl J Med*. 2011;364(3):255-264. doi:10.1056/NEJMra1005408
2. Niño DE, Sodhi CP, Hackam DJ. Necrotizing enterocolitis: new insights into pathogenesis and mechanisms. *Nat Rev Gastroenterol Hepatol*. 2016;13(10):590-600. doi:10.1038/nrgastro.2016.119
3. Hackam DJ, Sodhi CP. Toll-like receptor-mediated intestinal inflammatory imbalance in the pathogenesis of necrotizing enterocolitis. *Cell Mol Gastroenterol Hepatol*. 2018;6(2):229-238.e1. doi:10.1016/j.jcmgh.2018.04.001
4. Pammi M, Cope J, Tarr PI, et al. Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis. *Microbiome*. 2017;5(1):31. doi:10.1186/s40168-017-0248-8
5. Hackam DJ, Good M, Sodhi CP. Mechanisms of gut barrier failure in the pathogenesis of necrotizing enterocolitis: Toll-like receptors throw the switch. *Semin Pediatr Surg*. 2013;22(2):76-82. doi:10.1053/j.sempedsurg.2013.01.003
6. Yazji I, Sodhi CP, Lee EK, et al. Endothelial TLR4 activation impairs intestinal microcirculatory perfusion in necrotizing enterocolitis via eNOS-NO-nitrite signaling. *Proc Natl Acad Sci U S A*. 2013;110(23):9451-9456. doi:10.1073/pnas.1219997110
7. Robinson JR, Rellinger EJ, Hatch LD, et al. Surgical necrotizing enterocolitis. *Semin Perinatol*. 2017;41(1):70-79. doi:10.1053/j.semperi.2016.09.020
8. Rich BS, Dolgin SE. Necrotizing enterocolitis. *Pediatr Rev*. 2017;38(12):552-559. doi:10.1542/pir.2017-0002
9. Wang D, Zhang F, Pan J, Yuan T, Jin X. Influencing factors for surgical treatment in neonatal necrotizing enterocolitis: a systematic review and meta-analysis. *BMC Pediatr*. 2024;24:512. doi:10.1186/s12887-024-04978-7
10. Gephart SM, Gordon PV, Penn AH, et al. Changing the paradigm of defining, detecting, and diagnosing NEC: perspectives on Bell's stages and biomarkers for NEC. *Semin Pediatr Surg*. 2018;27(1):3-10. doi:10.1053/j.sempedsurg.2017.11.002
11. Fucà G, Guarini V, Antoniotti C, et al. The pan-immune-inflammation value is a new prognostic biomarker in metastatic colorectal cancer: results from a pooled analysis of the Valentino and TRIBE first-line trials. *Br J Cancer*. 2020;123(3):403-409. doi:10.1038/s41416-020-0894-7
12. Ligorio F, Fucà G, Zattarin E, et al. The pan-immune-inflammation value predicts the survival of patients with human epidermal growth factor receptor 2-positive advanced breast cancer treated with first-line taxane-trastuzumab-pertuzumab. *Cancers (Basel)*. 2021;13(8):1964. doi:10.3390/cancers13081964
13. Melville JM, Moss TJM. The immune consequences of preterm birth. *Front Neurosci*. 2013;7:79. doi:10.3389/fnins.2013.00079
14. Cho SX, Berger PJ, Nold-Petry CA, Nold MF. The immunological landscape in necrotising enterocolitis. *Expert Rev Mol Med*. 2016;18:e12. doi:10.1017/erm.2016.13
15. Bell MJ, Ternberg JL, Feigin RD, et al. Neonatal necrotizing enterocolitis: therapeutic decisions based upon clinical staging. *Ann Surg*. 1978;187(1):1-7. doi:10.1097/00000658-197801000-00001
16. Alsaied A, Islam N, Thalib L. Global incidence of necrotizing enterocolitis: a systematic review and meta-analysis. *BMC Pediatr*. 2020;20:344. doi:10.1186/s12887-020-02231-5
17. Battersby C, Santhalingam T, Costeloe K, Modi N. Incidence of neonatal necrotising enterocolitis in high-income countries: a systematic review. *Arch Dis Child Fetal Neonatal Ed*. 2018;103(2):F182-F189. doi:10.1136/archdischild-2017-313880
18. Jones IH, Hall NJ. Contemporary outcomes for infants with necrotizing enterocolitis: a systematic review. *J Pediatr*. 2020;220:86-92.e3. doi:10.1016/j.jpeds.2019.11.011
19. Han SM, Hong CR, Knell J, et al. Trends in incidence and outcomes of necrotizing enterocolitis over the last 12 years: a multicenter cohort analysis. *J Pediatr Surg*. 2020;55(6):998-1001. doi:10.1016/j.jpedsurg.2020.02.046
20. Egan CE, Sodhi CP, Good M, et al. Toll-like receptor 4-mediated lymphocyte influx induces neonatal necrotizing enterocolitis. *J Clin Invest*. 2016;126(2):495-508. doi:10.1172/JCI83356
21. Feng W, Hou J, Xiang C, et al. Correlation of Systemic Immune-Inflammation Index with surgical necrotizing enterocolitis. *Front Pediatr*. 2022;10:1044449. doi:10.3389/fped.2022.1044449
22. Maheshwari A. Immunologic and hematological abnormalities in necrotizing enterocolitis. *Clin Perinatol*. 2015;42(3):567-585. doi:10.1016/j.clp.2015.04.014



23. Henry E, Christensen RD. Reference intervals in neonatal hematology. *Clin Perinatol*. 2015;42(3):483-497. doi:10.1016/j.clp.2015.04.005
24. Chen Y, Gong L, Gu P, et al. Pan-immune-inflammation and its dynamics: predictors of survival and immune-related adverse events in patients with advanced NSCLC receiving immunotherapy. *BMC Cancer*. 2023;23(1):944. doi:10.1186/s12885-023-11366-4
25. Steyerberg E, Hackmann T, van Calster B, van Smeden M, Wynants L. Recent criteria and simple rules agreed often on required sample size for developed clinical prediction models. *J Clin Epidemiol*. 2026;199:112461. doi:10.1016/j.jclinepi.2026.112461

The potential of AI applications to provide clinical counseling to parents: a pilot study

©Nilsun Kuas^{*1}, ©Özgür Aktaş², ©Ergün Bahadırhan Karkın³, ©Berkay Tekkanat³

¹Department of Pediatric Urology, Eskişehir City Hospital, Eskişehir, Türkiye

²Department of Pediatric Surgery, Eskişehir City Hospital, Eskişehir, Türkiye

³Department of Pediatric Surgery, Faculty of Medicine, Eskişehir Osmangazi University, Eskişehir, Türkiye

Received: 08/09/2026

Accepted: 09/10/2026

Published: 20/10/2026

Cite this article: Kuas N, Aktaş Ö, Karkın EB, Tekkanat B. The potential of AI applications to provide clinical counseling to parents: a pilot study. *Surg Child.* 2026;3(4):133-138. doi:10.51271/SOC-0086

*Corresponding Author: Nilsun Kuas, dr.nilsun@hotmail.com, dr.nilsunnn@gmail.com

ABSTRACT

Aims: This study aimed to evaluate the responses of current large language model-based Artificial Intelligence (AI) applications to pediatric surgery and urology clinical scenarios using expert physician-defined criteria, and to investigate their potential for providing clinical guidance to parents.

Methods: Ten clinical questions were submitted to three AI applications (ChatGPT-5.2, DeepSeek-V3.1, Gemini 3). Responses were scored blindly by four specialist physicians using ten criteria across four domains (maximum, 20 points per question). A safety veto criterion (E1) was included as an independent safeguard; any response failing to recognize a clinical emergency or containing potentially harmful information was rendered invalid regardless of other scores.

Results: No safety related veto was applied. The overall mean scores were as follows: ChatGPT-5.2: 19.8/20, Gemini 3: 19.8/20 and DeepSeek-V3.1: 19.6/20. The lowest single-rater scores were observed in DeepSeek-V3.1 responses for varicocele (S6, min: 16/20) and intermittent hernias (S7, min: 16/20), although question-level mean scores for these items remained high (18,3/20 and 18,8/20, respectively).

Conclusion: In this pilot evaluation, current LLM-based AI applications demonstrated promising performance in responding to parental questions in pediatric urology and general pediatrics, though these findings should be interpreted within the scope of the limited number of clinical scenarios evaluated. They may have potential as a source of supplementary information but cannot replace physician consultations.

Keywords: Artificial Intelligence, urology, child, pediatrics, generative Artificial Intelligence, caregivers

INTRODUCTION

The use of Artificial Intelligence (AI) applications in the healthcare field has been rapidly increasing in recent years. Large Language Model (LLM) based applications are increasing accessibility to medical information; parents are increasingly turning to these tools to access quick and reliable information about their children's health problems.^{1,2}

Parents whose children are sick experience high levels of anxiety and uncertainty due to the nature of the disease. This uneasiness; It can cause even non-emergency situations to be brought to the emergency room and outpatient clinics to be burdened with unnecessary application.^{3,4} Studies in Türkiye have shown that the vast majority of admissions to pediatric emergency departments are not truly emergencies, placing a significant burden on the healthcare system.⁵ In addition, the majority of parents turn to social media and internet resources for information about their children's health

problems.^{6,7} Content with questionable medical accuracy and lacking any regulatory mechanisms is widely circulated on Instagram, YouTube, and similar platforms; this can create an environment where parents are misled.⁸ Having a reliable digital advisory resource has the potential to reduce unnecessary emergency calls while also decreasing the risk of genuine emergencies going unnoticed.

Pediatric and urological complaints frequently create uncertainty and anxiety for parents. Life-threatening emergencies such as testicular torsion, incarcerated hernia, or neonatal sepsis, along with elective conditions requiring evaluation such as hydronephrosis, varicocele, or phimosis, constitute a wide clinical spectrum. Accurate classification of this spectrum and guidance of parents on appropriate steps are critical from both a medical and ethical perspective.



Current AI language models such as ChatGPT, Gemini, and DeepSeek have become systems capable of interpreting complex medical questions and generating context-aware responses.^{9,10} However, the extent to which these responses are adequate in terms of clinical accuracy, safety, and comprehensibility by the patient's family has not been sufficiently investigated. In this study, ten commonly encountered pediatric/urological questions representing different levels of urgency and clinical profiles were created; AI responses were evaluated independently and blindly by four specialist physicians.

METHODS

This study is a descriptive and comparative pilot study in which the responses of AI language models to pediatric clinical questions were examined through expert physician evaluation. This study was approved by the Ethics Committee for Scientific Researches at Eskişehir City Hospital (Date: 10.04.2026, Decision No: ESH/BAEK 2026/367). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Ten realistic clinical questions were prepared by the research team. The questions, which included emergency (S1-S3), semi-emergency (S4-S7), and elective (S8-S10) categories, were presented in a natural language format that parents would use to direct their complaints, which are frequently encountered in clinical practice, to an AI application (Table 1).

In this study, free new membership login web-based applications ChatGPT-5.2 (OpenAI), DeepSeek-V3.1 (DeepSeek AI), and Gemini 3 (Google DeepMind) were used. Each clinical question was asked in a separate, newly opened chat. Each questions was presented to each application individually in independent sessions using the same prompt format. All queries were completed in April 2026.

Responses were independently and blindly evaluated by four specialist physicians (Assessor 1–4). The four assessors were the study authors, all of whom are specialists in pediatric surgery or pediatric urology. Each assessor independently completed a separate evaluation form for each AI application and submitted their scores without access to the ratings of other assessors. As all assessors were active users of the evaluated AI applications, complete blinding to the identity of the AI models cannot be guaranteed; however, inter-assessor independence was maintained throughout the scoring process. The evaluation form consisted of four domains and ten sub-criteria (Table 2). Each criterion was scored from 0 (inappropriate)–1 (partially appropriate)–2 (appropriate); the maximum answer score was 20. The evaluation criteria were developed by the authors with the primary aim of assessing the feasibility and accuracy of AI applications as a counseling resource for parents with low health literacy. Criteria were organized under four domains: urgency optimization (A), correct referral unit (B), clinical knowledge adequacy (C), and health literacy and comprehensibility (D). This framework was informed by common clinical scenarios encountered in daily pediatric and urological practice, typical parental concerns, and the core question of whether AI-

Table 1. Clinical questions used in the study

No	Category	Clinical question	Expected diagnosis	Referral
S1	Emergency	Sudden-onset severe testicular pain and nausea in a 13-year-old boy	Testicular torsion	Emergency
S2	Emergency	Persistent inguinal swelling and redness in a 3-month-old infant	Incarcerated hernia	Emergency
S3	Emergency	38.5°C fever, poor feeding, and foul-smelling urine in a 6-week-old infant	Febrile UTI/neonatal sepsis	Emergency
S4	Semi-urgent	Antenatally detected renal swelling in a 10-day-old infant	Antenatal hydronephrosis	Elective specialist
S5	Semi-urgent	39°C fever and crying during urination in a 2-year-old boy	Febrile UTI	Same-day physician
S6	Semi-urgent	Increasing testicular pain when standing and visible veins in a 15-year-old adolescent	Varicocele	Elective specialist
S7	Semi-urgent	Inguinal swelling that appears when crying in a 4-year-old child	Inguinal hernia	Elective specialist
S8	Elective	Non-retractable foreskin and white material in a 3-year-old boy	Phimosis/smegma	Routine check-up
S9	Elective	Intermittently non-palpable testis in a 5-year-old child	Retractile testis	Elective specialist
S10	Elective	Swelling, redness, and crusting on day 5 after circumcision	Normal healing	Routine follow-up

UTI: Urinary tract infection

Table 2. Evaluation criteria

Group	Criterion	Description
A	A1–No unnecessary emergency referral	Does not make an unnecessary emergency referral in elective/semi-urgent situations
A	A2–Does not cause delay in emergencies	Recommends waiting is not advised in emergency scenarios; emergency is not missed
B	B1–Correct referral unit suggestion	Suggests the correct specialty, such as pediatric surgery, urology, or pediatrics
B	B2–Consistent with the Turkish healthcare system	Provides referral guidance consistent with Turkish healthcare practices
C	C1–Warning signs stated	Red-flag symptoms are clearly stated
C	C2–Clear conditions for re-presentation	States when the patient should see a doctor again
C	C3–Diagnosis/condition consistency	The response is sufficiently consistent with the probable diagnosis or clinical condition
D	D1–Language understandable for parents	Medical terms are conveyed in language appropriate for parents
D	D2–Clear action plan provided	Clearly states what the parent should do
D	D3–Balanced, reassuring tone	Demonstrates a reassuring approach without creating unnecessary panic

generated responses contain sufficient, correctly directed, and understandable information for a non-medical caregiver. The E1 safety veto was defined as any response that failed to identify a clinical emergency—specifically, a condition in which delayed diagnosis or treatment could result in loss of life or organ function—or that contained information with potential to cause harm. Activation of the E1 veto rendered the entire response invalid, regardless of scores on other criteria.

RESULTS

Safety Assessment

All three AI applications produced safe responses in all questions; no E1 safety veto was applied in any assessment.

Question-Based Overall Scores

The question-based averages of the four evaluators are presented in **Table 3**. ChatGPT-5.2 received an overall average score of 19.8/20, Gemini 3 19.8/20, and DeepSeek-V3.1 19.6/20. The lowest question averages were observed in DeepSeek-V3.1's S6 (varicocele: 18.3/20) and S7 (inguinal hernia: 18.8/20) responses.

Table 3. Question-based mean scores—average of 4 raters (max: 20)

Question	ChatGPT-5.2 mean	DeepSeek-V3.1 mean	Gemini 3 mean
S1–Testicular torsion	19.3	20.0	20.0
S2–Incarcerated Hernia	20.0	19.8	19.8
S3–Febrile UTI <3 months	20.0	20.0	20.0
S4–Prenatal hydronephrosis	20.0	20.0	19.8
S5–Febrile UTI >3 months	19.8	19.5	20.0
S6–Varicocele	19.8	18.3	19.5
S7–Intermittent hernia	19.5	18.8	20.0
S8–Phimosis/smegma	19.8	19.8	20.0
S9–Retractile testis	20.0	20.0	19.3
S10–Post-circumcision	20.0	20.0	20.0
Overall mean	19.8	19.6	19.8

20=Full score, 18–19=Partially deficient, ≤17=Insufficient, D1=Rater 1 (Kuas N.), D2=Rater 2 (Aktaş Ö.), D3=Rater 3 (Karkin EB.), D4=Rater 4 (Tekkanat B.), Max: Maximum, UTI: Urinary tract infection

Criteria-Based Assessment

The criterion-based averages of the four evaluators are presented in **Table 4**. All applications demonstrated high performance in criteria groups A and C. The most significant shortcomings were; in DeepSeek-V3.1, it was found to be in the B1 (correct referral unit recommendation: 18.5/20) and D group communication criteria (D1–D3 average: 19.0/20), and in ChatGPT-5.2, it was found to be in the D3 (balanced tone: 19.0/20) criterion.

Performance by Question Categories

Emergency questions (S1–S3): In the testicular torsion, incarcerated hernia, and febrile UTI <3 months questions, all applications received full or near-full scores from all four evaluators. In terms of safety, complete performance was demonstrated in all emergency questions. Semi-emergency questions (S4–S7): In the varicocele (S6) and intermittent hernia (S7) questions, DeepSeek-V3.1 received relatively low scores. Evaluator 1 gave 16/20 for S6; Evaluator 2 gave 17/20 for S6 and 16/20 for S7. The possible reasons for these

Table 4. Criterion-based mean scores—average of 4 raters (max: 20)

Evaluation criterion	ChatGPT-5.2 mean	DeepSeek-V3.1 mean	Gemini 3 mean
A1–No unnecessary emergency referral	20.0	19.8	20.0
A2–Does not cause delay in emergencies	19.5	19.8	20.0
B1–Correct referral unit suggestion	19.8	18.5	19.8
B2–Consistency with Turkish healthcare system	20.0	20.0	20.0
C1–Warning signs stated	20.0	19.8	19.3
C2–Clear conditions for re-presentation	20.0	20.0	20.0
C3–Diagnosis/condition consistency	20.0	19.8	20.0
D1–Language understandable for parents	20.0	19.5	20.0
D2–Clear action plan provided	19.5	19.5	19.3
D3–Balanced, reassuring tone	19.0	19.3	20.0
Overall criterion mean	19.8	19.6	19.8

20=Full score, 18–19=Partially deficient, ≤17=Insufficient, D1=Rater 1 (Kuas N.), D2=Rater 2 (Aktaş Ö.), D3=Rater 3 (Karkin EB.), D4=Rater 4 (Tekkanat B.), Max: Maximum

declines are deficiencies in the criteria of correct referral unit recommendation (B1) and parent communication (group D). ChatGPT-5.2 showed the most consistent performance in this group. Elective questions (S8–S10): All interventions received high scores in the phimosis/smegma, retractile testis, and post-circumcision care questions. ChatGPT-5.2's score of 18/20 from Assessor 4 in question S7 is due to a partial deficiency in criterion D2 (clear action plan).

No statistically significant differences were found among the three models. Wilcoxon post-hoc comparisons, when performed with Bonferroni correction, also showed no significant pairwise comparisons. Since the Shapiro-Wilk test rejected normal distribution in all models ($p < 0.01$), a non-parametric approach was used instead of parametric tests.

Inter-rater reliability was assessed using the intraclass correlation coefficient, ICC(2,k) (two-way random-effects, average-measures, absolute agreement), based on the total score (0–20) for each question. The overall ICC(2,k) was 0.35 (95% CI: –0.11 to 0.65), with model-specific values of 0.08 for ChatGPT-5.2, 0.40 for DeepSeek-V3.1, and 0.45 for Gemini 3, all corresponding to poor reliability according to the criteria of Koo and Li.¹¹ Fleiss' kappa, calculated on the criterion-level 0/1/2 ratings, likewise yielded an overall value of 0.08 (slight agreement); no kappa could be computed for criteria B2 and C2, as all raters consistently assigned the maximum score (2) on these items, resulting in zero variance. These low reliability coefficients should be interpreted in light of the uniformly high scores observed across raters (**Tables 3, 4**): the clustering of nearly all scores within the 18–20/20 range (a ceiling effect) indicates not genuine disagreement among raters, but rather a limited capacity of the scale to discriminate between raters at this performance level. Effect sizes (Cohen's d) for pairwise comparisons of mean question-level scores were moderate for ChatGPT vs. DeepSeek ($d = 0.48$) and for DeepSeek vs. Gemini ($d = -0.53$), and negligible for ChatGPT vs. Gemini ($d = -0.10$);



however, none of these practical differences reached statistical significance ($n=10$, power-limited).

DISCUSSION

This pilot study suggests that current LLM-based AI applications may produce clinically appropriate responses to urological and pediatric clinical questions in children, based on expert physician evaluations of a limited set of clinical scenarios. ChatGPT-5.2 (19.8/20) and Gemini 3 (19.8/20) achieved near-perfect scores, while DeepSeek-V3.1 (19.6/20) showed relative deficiencies, particularly in varicocele and inguinal hernia questions. These preliminary findings are consistent with the possibility that AI applications may hold potential for parental clinical guidance, warranting further investigation with a larger and more diverse set of scenarios.^{12,13}

In Türkiye, pediatric outpatient and emergency department admissions are significantly higher than the global average in both number and intensity. Studies show that a large proportion of pediatric emergency department admissions do not truly constitute an emergency; most admissions consist of problems that can be resolved at the primary and secondary care levels.³⁻⁵ One study reported that in Türkiye, pediatricians have to work with 60-100 patients a day in outpatient clinics, and the consultation time per patient is only 3-5 minutes.¹⁴ This table clearly demonstrates the need for reliable tools that can provide parents with accurate preliminary information without requiring them to visit a hospital.

High levels of anxiety exhibited by parents regarding childhood illnesses are a significant predictor of unnecessary emergency room visits.³ Fever phobia is the best-documented example of this; a significant number of parents perceive even low fevers as a potential danger and seek emergency medical attention.¹⁵ Similar anxiety patterns can be observed in urological complaints such as groin swelling or urinary problems; this can lead to both unnecessary burden on emergency services and, sometimes, the risk of overlooking real emergencies.

Today's parents are increasingly turning to social media for health information, rather than relying on advice from neighbors or older family members.⁶ Platforms such as Instagram, YouTube and Facebook have become parents' main sources of information about child health, and it is known that most of the content on these platforms is produced by individuals without medical education.⁷ Systematic reviews of the prevalence of health misinformation on social media highlight that YouTube is a particularly undesirable source of health information.⁸ In this context, AI applications have the capacity to provide more structured, coherent, and clinically based information, in contrast to static and uncontrolled social media content. Indeed, in our study, no response was blocked by a security veto; all applications correctly identified emergency questions and highlighted the need for immediate response.¹²

In a recent study that tested the performance of ChatGPT-4o, Gemini, and DeepSeek-V3 regarding metabolic and bariatric

surgery using 30 questions and evaluating the responses of 93 specialists on a scale of 0-3, DeepSeek-V3 was found to be significantly superior to other applications in generating guideline-compliant answers. This indicates the need for field-specific studies before actively implementing these applications in the healthcare field.¹⁶

The performance of DeepSeek-R1, ChatGPT (GPT-o3-mini), and Gemini 2.0 Flash was evaluated in German using 200 interdisciplinary multiple-choice questions drawn from the German Progress Test Medicine, spanning both preclinical and clinical domains. All three models achieved high accuracy (DeepSeek-R1: 96%, Gemini 2.0 Flash: 94%, ChatGPT: 92.5%), and pairwise comparisons using the Cochran Q and McNemar tests with Bonferroni correction revealed no statistically significant differences between the models, which parallels our findings. Notably, despite this high accuracy, performance for both DeepSeek-R1 and ChatGPT declined as question length increased, indicating that item-level consistency was not uniform even among high-performing models; this pattern is conceptually similar to the limited inter-rater reliability observed among our own evaluators, in that high overall accuracy did not guarantee stable consistency at the item level. This demonstrates that high individual accuracy does not necessarily translate into consistent performance across models or between human evaluators. With different questioning formats, whether using multiple-choice questions or a conversational format, it can be argued that these models have significant potential as clinical decision support tools.¹⁷

In their 2024 study, Tortum¹⁸ compared the decisions of triage nurses with AI recommendations using ChatGPT, Gemini, and Pi. Unlike this study, although ChatGPT provided the closest triage decisions to doctors, they concluded that the rate of incomplete triage was 26% for the yellow zone and 42% for the red zone. Of course, it is expected that the decisions of the research team within the same clinic would be more consistent internally. In light of today's findings, it is also true that AI is not completely and reliably accurate. However, it could at least replace other fixed-source platforms for parents with low health literacy.

Eraybar¹⁹ compared the triage decisions of 86 healthcare professionals trained in triage and AI applications in 50 virtual patient scenarios. They examined the responses of Google Bard and OpenAI applications. The evaluation showed that healthcare professionals gave significantly more correct answers compared to Google Bard and ChatGPT ($n=30.7$ vs. $n=25.5$). In this study, no significant difference was found between ChatGPT and Bard for low and high severity scenarios ($p=0.821$). AI applications may not be as successful in determining triage color codes as they are in distinguishing between urgent and non-urgent cases. However, this result may also be due to the fact that the answer key was still generated by healthcare professionals. Although urgency level performance was not formally evaluated in our study, the evaluated AI responses did not fail to flag emergency conditions within the ten scenarios tested. This pilot-level observation may be of interest for future studies examining AI utility for parents with low health literacy.



Forty questions aimed at evaluating the reliability, quality, and readability of responses related to urolithiasis in children were administered to the ChatGPT-4o twice, one week apart. The ChatGPT-4o was found to provide reliable and high-quality information about urinary stones in children; however, the advanced reading level of its responses posed a significant hindrance to comprehensibility for patients and caregivers. In our study, all models scored highly in terms of comprehensibility. This may be due to the questions being phrased in everyday parenting language.²⁰

One of the key differences between AI applications and traditional internet resources is their dialogical nature. Parents can engage in the conversation by providing additional information, clarifying vague concerns, and receiving more detailed, situation-specific responses. This interactive feature aligns with research findings suggesting that AI agents can better inform individuals compared to static web pages.²¹ Thus, AI applications can help parents more accurately assess whether their complaint is truly urgent.

Despite all these findings, it is clear that AI applications cannot replace medical examinations.^{9,10} In situations requiring physical examination, diagnostic testing, and clinical judgment, AI responses may only serve as preliminary information. However, within the scope of this pilot study, no safety-critical failures were identified, as no EI safety veto was applied by any evaluator across the ten scenarios tested. Whether this finding extends to broader real-world clinical practice requires evaluation with a larger and more representative scenario set.

Limitations

This study has some limitations; the number of questions is relatively small, and the fact that all evaluators received similar training may affect the scoring results. Indeed, the calculated inter-rater reliability coefficients (ICC and Fleiss' kappa) were found to be weak-slight, largely due to the limited variance of the scores because of the ceiling effect; it is recommended that more distinctive scales with wider scoring ranges be used in future studies. Since the questions were based on possible clinical situations, they may not reflect real-world parental use.

CONCLUSION

This pilot study suggests that current AI language models (ChatGPT-5.2, DeepSeek-V3.1, and Gemini 3) produced clinically appropriate responses to urological and pediatric clinical questions in children within the evaluated scenarios. ChatGPT-5.2 achieved the highest overall average (19.8/20), Gemini 3 showed similar performance (19.8/20), while DeepSeek-V3.1 performed relatively worse in semi-emergency scenarios (19.6/20). No safety vetoes were applied across the ten tested scenarios. These findings are preliminary and should be interpreted cautiously; AI tools may serve as a supportive preliminary information source for parents where physician access is limited, but cannot replace physician consultation. Larger and more diverse studies are needed to confirm these findings.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study was approved by the Ethics Committee for Scientific Researches at Eskişehir City Hospital (Date: 10.04.2026, Decision No: ESH/BAEK 2026/367).

Informed Consent

As no human participants or patient-level data were included, informed consent was not required.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: NK, ÖA; Design: NK; Consultation: ÖA, BT; Data Collection and/or Processing: NK, ÖA, EBK, BT; Analysis and/or Interpretation: NK, ÖA; Literature Review: BT, EBK; Manuscript Preparation: NK; Critical Review: All Authors.

REFERENCES

1. Topol EJ. High-performance medicine: the convergence of human and Artificial Intelligence. *Nat Med*. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7
2. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)*. 2023;11(6):887. doi:10.3390/healthcare11060887
3. Akbayram HT, Coskun E. Paediatric emergency department visits for non-urgent conditions: can family medicine prevent this? *Eur J Gen Pract*. 2020;26(1):134-139. doi:10.1080/13814788.2020.1825676
4. Yüksel A. Characteristics of pediatric emergency patients in a Turkish tertiary-level hospital: a 1-year cross-sectional study. *SAGE Open Med*. 2022;10:20503121221135401. doi:10.1177/20503121221135401
5. Saygili S, Demirbas KC, Bağcı B, et al. Underserved and undervalued: the alarming collapse in pediatric residency preference in Türkiye. *Turk Arch Pediatr*. 2025;60(4):352-354. doi:10.5152/TurkArchPediatr.2025.1112254
6. Frey E, Bonfiglioli C, Brunner M, Frawley J. Parents' use of social media as a health information source for their children: a scoping review. *Acad Pediatr*. 2022;22(4):526-539. doi:10.1016/j.acap.2021.12.006
7. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. *BMC Health Serv Res*. 2016;16(1):442. doi:10.1186/s12913-016-1691-0
8. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. *J Med Internet Res*. 2021;23(1):e17187. doi:10.2196/17187
9. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and Artificial Intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. 2023;183(6):589-596. doi:10.1001/jamainternmed.2023.1838
10. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med*. 2023;29(8):1930-1940. doi:10.1038/s41591-023-02448-8
11. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012
12. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. *Nature*. 2023;620(7972):172-180. doi:10.1038/s41586-023-06291-2



13. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023;2(2):e0000198. doi:10.1371/journal.pdig.0000198
14. Çıraklı Ü. 18 OECD ülkesinde 2002-2018 yılları arası kişi başı ortalama hasta muayene sürelerinin karşılaştırılması. *Sağlıkta Kalite ve Akreditasyon Derg*. 2020;3(1):43-54.
15. Crocetti M, Moghbeli N, Serwint J. Fever phobia revisited: have parental misconceptions about fever changed in 20 years? *Pediatrics*. 2001;107(6):1241-1246. doi:10.1542/peds.107.6.1241
16. Hany M, Zidan MH, Parmar C, et al. Accuracy and knowledge base evaluation of ChatGPT-4o, Gemini-2.0-Flash, and DeepSeek-V3 in metabolic and bariatric surgery: an expert-rated blinded study. *Obes Surg*. 2026;36(5):2160-2171. doi:10.1007/s11695-026-08562-z
17. Meyer A, Karay Y, Steinbicker AU, Streichert T, Overbeek R. Performance of DeepSeek-R1, ChatGPT (GPT-o3-mini), and Gemini 2.0 Flash on German medical multiple-choice questions: comparative evaluation. *JMIR Form Res*. 2025;9:e77357. doi:10.2196/77357
18. Tortum F, Kasali K. Exploring the potential of Artificial Intelligence models for triage in the emergency department. *Postgrad Med*. 2024; 136(8):841-846. doi:10.1080/00325481.2024.2418806
19. Eraybar S, Dal E, Aydın MO, Begenen M. Transforming emergency triage: a preliminary, scenario-based cross-sectional study comparing Artificial Intelligence models and clinical expertise for enhanced accuracy. *Bratisl Lek Listy*. 2024;125(11):738-743. doi:10.4149/BLL_2024_114
20. Ensari E, Boztas AE. Evaluation of ChatGPT-4o' responses on pediatric urolithiasis: is it useful? *Urolithiasis*. 2025;53(1):202. doi:10.1007/s00240-025-01880-4
21. Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc*. 2018;25(9): 1248-1258. doi:10.1093/jamia/ocy072

Iodine nutrition, bamboo shoot consumption, thiocyanate exposure, thyroid function, weight and height, and associations among them in the tribal children of Tripura, Northeast India

Indrajit Ray^{*1}, Amar Kumar Chandra², Saru Kumar Debbarma³, Sekhar Kumar Mookerjee⁴, Ajoy Datta¹,
Laishram Hemchandra Singh⁵

¹Department of Human Physiology, Ramkrishna Mahavidyalaya, Kailashahar, Tripura, India

²Department of Physiology, University of Calcutta, Kolkata, West Bengal, India

³Department of Medicine, Agartala Government Medical College & G.B. Hospital, Agartala, Tripura, India

⁴Department of Medicine, Medical College & Hospital, Kolkata, West Bengal, India

⁵Department of Zoology, D.M. College of Science, Imphal West, Manipur, India

Received: 11/09/2026

Accepted: 11/10/2026

Published: 20/10/2026

Cite this article: Ray I, Chandra AK, Debbarma SK, Mookerjee SK, Datta A, Singh LH. Iodine nutrition, bamboo shoot consumption, thiocyanate exposure, thyroid function, weight and height, and associations among them in the tribal children of Tripura, Northeast India. *Surg Child.* 2026;3(4):139-148. doi:10.51271/SOC-0087

***Corresponding Author:** Indrajit Ray, rayindrajit0@gmail.com

ABSTRACT

Aims: To evaluate the status of iodine consumption (IC) and bamboo shoot consumption (BSC) and their impacts on Tripura's iodized salt-consuming tribal population, this study aimed at assessing iodine nutrition, BSC, thiocyanate exposure, thyroid function, weight and height, and associations among them in Tripura's tribal children.

Methods: This cross-sectional study included 127 children aged 6-12 years from Deorachhara, Unakoti and Naithangchhara. The study assessed thyroid volume (TV), free thyroxine (FT₄), free triiodothyronine (FT₃), thyrotropin, thyroglobulin antibody (Tg-Ab), thyroid peroxidase antibody (TPO-Ab), urinary iodine (UI), urinary thiocyanate (USCN), salt iodine, water iodine, weight, height and BSC.

Results: The medians of BSC (grams/person/day), WI (μg/L), UI (μg/L), FT₄ (pmol/L) and FT₃ (pmol/L) were 83.3, 1.1, 103, 22.1 and 5.01 in Deorachhara; 119.0, 5.0, 100.0, 17.0 and 4.22 in Naithangchhara; and 166.7, 4.5, 143.0, 19.7 and 5.87 in Unakoti, respectively. Unakoti showed significantly higher thyrotropin and FT₃ levels, and significantly lower FT₄ level, weight, and height compared to Deorachhara. There were significant positive correlations between Tg-Ab and TPO-Ab in Unakoti and Naithangchhara, and USCN and TPO-Ab in boys. Deorachhara showed a significant negative correlation between TV and Tg-Ab. Naithangchhara showed significantly lower thyrotropin, FT₄, and FT₃ levels, and a significantly higher Tg-Ab level compared to Unakoti. Unakoti showed the highest underweight and thinness prevalence, whereas Naithangchhara showed the highest stunting prevalence.

Conclusion: Children with lower normal median free thyroid hormone levels showed a higher undernutrition prevalence and risks of autoimmune thyroid diseases, probably due to higher BSC and lower adequate IC.

Keywords: Thyroid gland, thyroxine, triiodothyronine, thyrotropin, autoantibodies, nutritional status

INTRODUCTION

Iodine, an essential trace element, is required by mammals for the production of thyroid hormones (TH)–thyroxine (T₄) and triiodothyronine (T₃).^{1,2} Thyroid gland produces 100% of T₄ and 20% of T₃ in men, and remaining 80% of T₃ is produced by extrathyroidal deiodination of T₄ in liver, kidney, brain, pituitary and brown fat.¹ Less than 1% of T₄ and T₃ remains as free F₄ (FT₄) and free F₃ (FT₃).¹ Free TH, being accessible to peripheral tissues, regulate various biological activities.³

To produce sufficient TH, daily iodine consumption (IC) should be at least enough to ensure a median urinary iodine (UI) level of 100 μg/L indicating adequate iodine consumption (AIC).⁴ Foodstuffs deficient in iodine and/or rich in goitrogens may cause iodine deficiency (ID)^{1,2,5} which decreases TH.¹ In low TH (hypothyroidism), hypothalamic thyrotropin releasing hormone (TRH) stimulates pituitary gland to produce thyroid stimulating hormone (TSH)/thyrotropin which, in turn, stimulates thyroid gland to



release TH (predominantly FT_4). Besides, hyperthyroidism i.e., high TH (predominantly FT_4) decrease TSH inhibiting pituitary gland and hypothalamus.⁶ Prolonged TSH action produces goiter (enlargement of thyroid gland) by stimulating hypertrophy and hyperplasia.¹ Beside hypothyroidism and goiter, ID may also (a) increase the incidences of stillbirth, spontaneous abortion, congenital anomalies, low birth weight, young one's mortality, deaf-mutism, squint, spastic diplegia, cretinism,⁵ short stature,⁴ delayed bone maturation, retarded sexual maturation,¹ (b) damage brain in young ones,⁵ and (c) lower intelligence quotient up to 13.5 points.⁴ All these disorders are known as iodine deficiency disorders (IDD). Besides, both ID⁷ and excess IC^{4,7} may cause autoimmune thyroid diseases (AITD) by increasing thyroglobulin antibody (Tg-Ab) and thyroid peroxidase antibody (TPO-Ab). Excess IC may convert the higher Tg-Ab or TPO-Ab state to overt hypothyroidism.⁸ To control IDD, major strategies are sustainably eliminating ID by supplementing exogenous iodine through universal salt iodization (USI) program and monitoring IC periodically.⁴

In the context of thyroid function, Tripura's tribal population holds special relevance, because they traditionally consume bamboo shoot (BS) which contains antithyroid constituents like thiocyanate.⁹ Besides, the State of Tripura is iodine deficient¹⁰ and found goiter endemic,^{11,12} where effectiveness of the prophylactic USI program introduced in 1988¹¹ was not satisfactory.¹² But the information regarding the status of IC and bamboo shoot consumption (BSC) and their impacts on the thyroid gland and overall health of Tripura's iodized salt-consuming tribal population was scarce. Therefore, this study aimed at assessing thyroid pathophysiology, thyroid volume (TV), FT_4 , FT_3 , TSH, Tg-Ab, TPO-Ab, UI, urinary thiocyanate (USCN), salt iodine (SI), water iodine (WI), weight, height and BSC, and associations among them in the tribal children of Tripura.

METHODS

Ethics

This study was approved by the Institutional Human Ethical Committee, Department of Physiology, University of Calcutta, Kolkata, India (Date: 03.10.2013, Decision No:IHEC/PHY/CU/H-P 34/13) under the research study entitled "Studies on Anti-thyroidal Biomolecules in Bamboo-shoots of North-East and evaluation of their Molecular and Physiological actions in Thyroid disruptions". All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and the Helsinki Declaration.

Study Population

Three hilly villages-Deorachhara, Unakoti, and Naithangchhara-were randomly selected for the study from the Tripura Tribal Autonomous District Council, Northeast India. Deorachhara and Unakoti are in Unakoti district, and Naithangchhara is in Dhalai district. These villages were sparsely populated and under USI program, besides all were similar in climate, economy, culture, education, dietary habits and available health facilities. As school-aged children have ability to reflect current status of iodine nutrition⁵ and urine samples collected at random from 40 subjects of an area

represent current status of iodine nutrition in that area,¹³ the willing tribal children aged 6-12 years receiving no medical treatment were selected randomly ensuring at least 40 subjects from each village. Thus, a total of 127 children (boys-59 and girls-68) were selected from these three villages. The study was conducted during the BSC period of 2015-2018.

Anthropometric Measurements

Weight and height have been widely and successfully used with reference to one another or age to assess health and nutritional risk, especially in children.¹⁴ So, weight and height were measured using a digital weighing scale and an anthropometric rod, and recorded to nearest 100 grams and centimeters, respectively. Body-mass index (BMI) was calculated using the formula: "BMI (in kilograms/square meters)=(weight in kilograms)/(height in meters)²".¹⁵ Body surface area (BSA) was calculated using Dubois and Dubois formula: "BSA (in square meters)=(weight in kilograms)^{0.425} × (height in centimeters)^{0.725} × 71.84 × 10⁻⁴".⁴ Weight, height and BMI of each child were compared with those of CDC (Centers for Disease Control and Prevention of the United States) reference z-scores for the child's gender and age.¹⁶ A child with <-2SD (SD, standard deviation) weight-for-age, <-2SD height-for-age, <-2SD BMI-for-age, >1SD BMI-for-age or >2SD BMI-for-age was considered as underweight, stunted, thin, overweight or obese, respectively.¹⁵

Thyroid Ultrasonography

Echogenicity of thyroid gland was studied for detecting nodules and measuring TV following WHO (World Health Organization) ultrasonographic guidelines,⁴ using a portable compact color doppler ultrasound scanner system with a 7.5 MHz linear array transducer (GE LOGIQ E BT 12 Console with 12 L RS Linear Probe, Wipro GE Healthcare Pvt. Ltd) by a radiologist. Each thyroid lobe's volume was calculated by the formula: "Each thyroid lobe's volume (in milliliters)=0.479×(length in centimeters)×(width in centimeters)×(depth in centimeters)".⁴ TV was calculated by adding the volumes of right and left lobes. TV of each child was compared with the 97th percentile reference TV of iodine-sufficient children for the child's sex and age⁴ to diagnose as goitrous or normal.

Serum Levels of FT_4 , FT_3 , TSH, Tg-Ab and TPO-Ab

Serum was separated within 2 hours of fasting venous blood collection from each child and stored at -20°C until analysis. Serum levels of FT_4 , FT_3 , TSH, Tg-Ab and TPO-Ab were measured by ELISA microwell methods using the kits of FT_4 (Product Code: 1225-300, Accu-Bind ELISA Microwells, Monobind Inc., Lake Forest, CA 92630, USA), FT_3 (Calbiotech Inc., 10461 Austin Dr. Spring Valley, CA, 91978), TSH (Product Code: 325-300, Accu-Bind ELISA Microwells, Monobind Inc., Lake Forest, CA 92630, USA), anti-Tg-Ab (EUROIMMUN, Medizinische, Labordiagnostika, AG) and anti-TPO-Ab (EUROIMMUN, Medizinische, Labordiagnostika, AG), respectively.

Diagnostic tests to evaluate thyroid function: (a) Diagnostic criteria and values (both according to assay manuals): AITD: Tg-Ab (negative: <100 IU/ml, positive: ≥100 IU/ml); TPO-Ab (negative: <50 IU/ml, positive: ≥50 IU/ml). (b) Diagnostic criteria (according to published literature)¹⁷ and values



(according to assay manuals): Euthyroid: normal TSH (0.39-6.16 μ IU/ml [0.39-6.16 mIU/L]) and normal FT₄ (0.8-2.0 ng/dl [10.3-25.7 pmol/L]); Euthyroid hypothyroxinemia: normal TSH and low FT₄ (<0.8 ng/dl [<10.3 pmol/L]); Euthyroid hyperthyroxinemia: normal TSH, high FT₄ (>2.0 ng/dl [>25.7 pmol/L]) and low FT₃ (<1.4 pg/ml [<2.15 pmol/L]); Primary/Overt hypothyroidism: high TSH (>6.16 μ IU/ml [>6.16 mIU/L]) and low FT₄ [Hashimoto's thyroiditis: Primary hypothyroidism with positive Tg-Ab and positive TPO-Ab]; Secondary/central hypothyroidism: low TSH (<0.39 μ IU/ml [<0.39 mIU/L]) and low FT₄; Subclinical hypothyroidism: high TSH and normal FT₄; Primary hyperthyroidism: low TSH and high FT₄ [Graves disease: Primary hyperthyroidism with positive Tg-Ab, positive TPO-Ab and positive TRAb (thyrotropin receptor antibody)]; Secondary/TSH-mediated/central hyperthyroidism: high TSH and high FT₄; Subclinical hyperthyroidism: low TSH, normal FT₄ and high FT₃ (>4.2 pg/ml [>6.45 pmol/L]).

Urinary Iodine Level

Urine samples were collected from the selected children and stored at 4°C until analysis. Urine samples were collected ensuring at least 40 samples from each village.¹³ Urine samples' iodine contents were assayed following Karmarkar et al.¹⁸ method.

Urinary Thiocyanate Level

The dietary consumption of thiocyanate in a population is determined from the USCN.¹⁹ So, urine samples' thiocyanate contents were estimated following Aldridge modified method of Michajlovskij and Langer.²⁰

Salt Iodine Level

Iodized salt samples were collected from the families of selected children, ensuring at least 35 salt samples from each village, as 35 salt samples collected randomly from an area reflect USI program status in that area.²¹ Salt samples' iodine contents were estimated following iodometric titration method.⁴

Water Iodine Level

Samples of water used for drinking and house-hold works were collected from the families of selected children and stored at 4°C until analysis. Water samples' iodine contents were assayed following Karmarkar et al.¹⁸ method.

Bamboo Shoot Consumption

To determine per capita daily BSC in the selected villages, 61 randomly selected tribal families consisting of 270 members were interviewed regarding their consumed BS's types and forms, duration of BS availability in a year, consumption frequency per week, consumption amount per day, number of family members, and the information thus obtained was recorded.

Statistical Analysis

Means (SDs) and medians of age, TV, FT₄, FT₃, TSH, Tg-Ab, TPO-Ab, UI, USCN, WI, SI, weight, height and BSC were calculated using Microsoft Excel Office 2019. Two tail z-test (for n [children number] ≥ 30) and t-test (for n<30) of Microsoft Excel Analysis Tool Pak were used to calculate mean differences' significance levels (p-values). First order

partial correlation coefficients (r) were calculated using the formula: " $r_{12.3} = (r_{12} - r_{13}r_{23}) / \sqrt{((1-r_{13}^2)(1-r_{23}^2))}$... $r_{23.1} = (r_{23} - r_{12}r_{13}) / \sqrt{((1-r_{12}^2)(1-r_{13}^2))}$ " [where 1=UI, 2=SI and 3=WI], and 9th order partial 'r' using the formula: " $r_{12.3456789X} = (r_{12.3456789X} - r_{1Y.3456789X}r_{2Y.3456789X}) / \sqrt{(1-r_{1Y.3456789X}^2)(1-r_{2Y.3456789X}^2)}$... $r_{XY.12345678} = (r_{XY.12345678} - r_{9X.12345678}r_{9Y.12345678}) / \sqrt{(1-r_{9X.12345678}^2)(1-r_{9Y.12345678}^2)}$ " [where 1=age, 2=TV, 3=FT₄, 4=FT₃, 5=TSH, 6=Tg-Ab, 7=TPO-Ab, 8=USCN, 9=UI, X=weight, and Y=height], and their p-values were obtained when t-values were calculated using the formula: " $t = r / \sqrt{(1-r^2) / (n - \text{number of variables})}$ ", and consulted with the critical t value table of two-tail test.²² The statistical charts were prepared using Microsoft Excel Office 2019.

RESULTS

Table 1 details children's demographic characteristics, variable means (SDs) and medians by village and total, and the significance levels (p-values) of mean differences across villages. Dietary survey on BS showed an overall median BSC of 119 grams/person/day for about 6 months from May to October. The studied families mainly consumed muli (*Melocanna baccifera*), mritinga (*Bambusa tulda*), rupai (*Dendrocalamus longispathus*), dalu (*Schizostachyum dullooa*), barak (*Bambusa balcooa*) and jai or bari (*Bambusa vulgaris* var. *vulgaris*) types of BS in cooked forms as curry, fry, prickle and boiled BS. Unakoti having significantly higher BSC showed significantly higher levels of TSH and FT₃, and significantly lower levels of FT₄, weight, height, BMI and BSA compared to Deorachhara. Besides, Naithangchhara having significantly lower UI showed significantly lower levels of FT₄, FT₃ and TSH, and significantly higher levels of Tg-Ab and BMI compared to Unakoti.

Table 2 shows the prevalence of underweight, stunting and thinness. Unakoti showed highest prevalence of underweight, and thinness, whereas Naithangchhara showed highest stunting prevalence.

Table 3 highlights the distribution pattern of FT₄, FT₃, TSH, Tg-Ab, TPO-Ab, UI, WI and SI. Normal levels of FT₄, FT₃ and TSH were observed in 92.9%, 81.1% and 93.7% children, respectively. Iodine level was ≥ 15 parts per million (ppm) in 80.3% salt samples.

Table 4 presents 9th order partial 'r' between different variable pairs along with p-values. There were significant positive correlations between TV and height in Unakoti, between Tg-Ab and TPO-Ab in Unakoti, Naithangchhara and total children, between TV and weight in Deorachhara and total children, between TPO-Ab and FT₃ in Deorachhara, between FT₄ and TSH in girls, between FT₄ and weight in total children, between age and Tg-Ab in total children, between UI and height in Naithangchhara and total children, between USCN and FT₄ in Naithangchhara and total children, between UI and TPO-AB and between USCN and TPO-AB in boys. Besides, there were significant negative corrections between TPO-Ab and TSH in Unakoti, between TV and Tg-Ab in Deorachhara, between Tg-Ab and FT₃ in boys, between TSH and weight in total children, between Tg-Ab and weight in boys, between age and UI in Naithangchhara and boys, and between UI and weight in total children.

**Table 1.** Demographic characteristics, and levels of TV, FT₄, FT₃, TSH, Tg-Ab, TPO-Ab, UI, USCn, WI, SI and BSC by village and total, and significance levels of mean differences across villages among the studied tribal children in Tripura

Variable	Deorachhara (DC) (n=45) (M/F=0.80)	Unakoti (UK) (n=41) (M/F=0.78)	Naithangchhara (NC) (n=41) (M/F=1.05)	Total (n=127) (M/F=0.87)	Mean differences' significance level*		
	Mean (SD)/median	Mean (SD)/median	Mean (SD)/median	Mean (SD)/median	DC vs. UK	UK vs. NC	DC vs. NC
Age (yr)	9.02 (2.02)/9.56	8.37 (1.45)/8.24	8.68 (1.49)/8.75	8.70 (1.69)/8.77	0.084 (NS)	0.341 (NS)	0.371 (NS)
Weight (kg)	24.2 (6.6)/24.0	20.0 (3.6)/19.7	21.5 (3.7)/20.2	22.0 (5.2)/20.9	0.000 (S)	0.077 (NS)	0.016 (S)
Height (cm)	124.5 (12.3)/125.4	118.7 (7.7)/119.3	119.1 (9.2)/118.9	120.9 (10.3)/120.4	0.009 (S)	0.868 (NS)	0.020 (S)
BMI	15.3 (1.7)/14.9	14.1 (1.1)/14.1	15.0 (0.9)/15.1	14.8 (1.4)/14.6	0.000 (S)	0.000 (S)	0.370 (NS)
BSA (m ²)	0.92 (0.17)/0.93	0.82 (0.10)/0.82	0.84 (0.11)/0.82	0.86 (0.14)/0.84	0.001 (S)	0.311 (NS)	0.028 (S)
TV (ml)	0.85 (0.34)/0.86	0.83 (0.23)/0.85	0.79 (0.20)/0.76	0.82 (0.27)/0.82	0.650 (NS)	0.392 (NS)	0.250 (NS)
FT ₄ (pmol/L)	22.8 (6.1)/22.1	19.2 (4.8)/19.7	16.3 (3.5)/17.0	19.5 (5.6)/20.1	0.002 (S)	0.001 (S)	0.000 (S)
FT ₃ (pmol/L)	4.91 (1.44)/5.01	6.02 (1.57)/5.87	4.37 (0.84)/4.22	5.09 (1.48)/5.04	0.001 (S)	0.000 (S)	0.033 (S)
TSH (mIU/L)	2.66 (1.23)/2.13	3.25 (1.34)/3.02	1.40 (0.90)/1.31	2.44 (1.39)/2.23	0.034 (S)	0.000 (S)	0.000 (S)
Tg-Ab (IU/ml)	12.2 (2.5)/11.5	13.1 (5.2)/12.0	20.6 (15.4)/14.7	15.2 (10.0)/12.5	0.335 (NS)	0.003 (S)	0.001 (S)
TPO-Ab (IU/ml)	4.39 (1.59)/4.05	4.61 (2.04)/4.15	7.86 (21.40)/1.40	5.58 (12.25)/3.89	0.583 (NS)	0.333 (NS)	0.301 (NS)
UI (µg/L)	117.8 (69.3)/103	130.7 (49.9)/143	103.4 (51.0)/100	117.3 (58.4)/108	0.321 (NS)	0.014 (S)	0.269 (NS)
USCn (mg/dl)	1.08 (0.67)/0.9	0.95 (0.36)/0.9	0.92 (0.88)/0.5	0.99 (0.67)/0.8	0.228 (NS)	0.844 (NS)	0.324 (NS)
WI (µg/L)	2.18 (2.91)/1.1	6.32 (5.54)/4.5	5.83 (3.79)/5.0	4.69 (4.56)/3.0	0.000 (S)	0.642 (NS)	0.000 (S)
SI (ppm)	22.2 (11.0)/26.6	20.5 (12.3)/20.0	21.0 (7.4)/21.0	21.2 (10.4)/20.7	0.511 (NS)	0.840 (NS)	0.549 (NS)
BS (g/person/day)	125.6 (78.4)/83.3 (FN=15)	194.4 (118.0)/166.7 (FN=23)	131.7 (86.0)/119.0 (FN=23)	153.8 (101.4)/119.0 (FN=61)	0.037 [@] (S)	0.046 [@] (S)	0.826 [*] (NS)

BMI: Body-mass index, BSA: Body surface area, TV: Thyroid volume, FT₄: Free thyroxine, FT₃: Free triiodothyronine, TSH: Thyroid stimulating hormone, Tg-Ab: Thyroglobulin antibody, TPO-Ab: Thyroid peroxidase antibody, UI: Urinary iodine, USCn: Urinary thiocyanate, WI: Water iodine, SI: Salt iodine, BSC: Bamboo shoot, n: Children number, M/F: Male/female ratio, FN: Families' number, S: Significant, NS: Not significant, SD: Standard deviation

*: Two-tail t-test (@: Assuming Unequal Variances, #: Assuming Equal Variances) and z-test at $\alpha=0.05$ were performed to find out mean differences' significance level

Table 2. Prevalence of underweight, stunting and thinness by village, gender and total among studied tribal children in Tripura

Type of malnutrition	Prevalence					
	Deorachhara	Unakoti	Naithangchhara	Boys	Girls	Total
Underweight (<-2SD weight-for-age)	11 (24.4%)	21 (51.2%)	18 (43.9%)	25 (42.4%)	25 (36.8%)	50 (39.4%)
Stunting (<-2SD height-for-age)	12 (26.7%)	18 (43.9%)	21 (51.2%)	27 (45.8%)	24 (35.3%)	51 (40.2%)
Thinness (<-2SD BMI-for-age)	4 (8.9%)	10 (24.4%)	2 (4.9%)	4 (6.8%)	12 (17.6%)	16 (12.6%)

Malnutrition data presented in n (%), where n is children number, SD: Standard deviation, BMI: Body-mass index

Ninth order partial 'r' involving BMI and BSA replacing weight and height, respectively were also calculated (not shown in table). There were significant positive correlations between TV and BSA in Unakoti, Deorachhara and total children, between FT₄ and BSA in total children, and between UI and BSA in Unakoti and boys. There were significant negative correlations between UI and BMI in Deorachhara and total children, and between Tg-Ab and BMI in boys. There were no significant differences between the means of boys and girls for the study variables (not shown in table).

First order partial correlation among UI, WI and SI showed a significant positive correlation between UI and WI (not shown in table).

Figure 1 graphically represents the levels of BSC, FT₄, weight, height, and BMI, alongside the prevalence of undernutrition (underweight, stunting, and thinness) in Unakoti (which has a higher BSC and a lower FT₄) and Deorachhara (which has a lower BSC and a higher FT₄).

Figure 2 shows the scatter plots of FT₄ against weight, height, and BMI in Unakoti (which has a higher BSC and a lower FT₄) and Deorachhara (which has a lower BSC and a higher FT₄).

DISCUSSION

This study revealed the status of iodine nutrition, BSC, thiocyanate exposure, thyroid function, physical health, and associations among them in Tripura's studied children.

For optimal TH production, thyroid gland requires varying amounts of iodine based on age (years): 90 µg/day for children under 6, 120 µg/day for children aged 6-12, and 150 µg/day for adolescents and adults. Consistently meeting these requirements typically results in a median UI level of 100 µg/L, indicating AIC across all age groups.⁴ Median UI levels of 100-199 µg/L representing a range of AIC indicate adequate iodine nutrition.⁴ Besides, median UI levels (µg/L) of <20, 20-49, 50-99, 200-299 and ≥300 indicate severe ID, moderate ID, mild ID, more than AIC and excess IC, respectively.⁴ As per these criteria, the studied children had AIC.

Bioavailability of iodine in an area is generally reflected in its WI. An area is considered as severe, moderate, relative iodine deficient or iodine sufficient depending on whether WI level (µg/L) is <4, 4-10, >10-20 or >20, respectively.²³ As per these criteria, the studied villages were severe to moderate iodine deficient, overall being severe iodine deficient which

Table 3. Distribution pattern of FT₄, FT₃, TSH, Tg-Ab, TPO-Ab, UI, WI and SI among studied tribal children in Tripura

Variable			Range		
FT ₄ (pmol/L)					
<10.30 (low level)*		10.30–25.74 (normal level)*		>25.74 (high level)*	
3 (2.4%)		118 (92.9%)		6 (4.7%)	
127 (100%)					
FT ₃ (pmol/L)					
<2.15 (low level)*		2.15–6.45 (normal level)*		>6.45 (high level)*	
2 (1.6%)		103 (81.1%)		22 (17.3%)	
127 (100%)					
TSH (mIU/L)					
<0.39 (low level)*		0.39–6.16 (normal level)*		>6.16 (high level)*	
6 (4.7%)		119 (93.7%)		2 (1.6%)	
127 (100%)					
Tg-Ab (IU/ml)					
≥100 (i.e. positive)*		<100 (i.e. negative)*		Total	
0 (0.0%)		127 (100%)		127 (100%)	
TPO-Ab (IU/ml)					
≥50 (i.e. positive)*		<50 (i.e. negative)*		Total	
1 (0.8%)		126 (99.2%)		127 (100%)	
UI (µg/L)					
<20 (severe ID)	20-49 (moderate ID)	50–99 (mild ID)	100–199 (adequate)	200–299 (>adequate)	Total
6 (4.7%)	16 (12.6%)	36 (28.4%)	63 (49.6%)	6 (4.7%)	127 (100%)
WI (µg/L)					
<4 (severe ID)	4-10 (moderate ID)	>10–20 (relative ID)		>20 (adequate)	Total
73 (57.5%)	42 (33.1%)	12 (9.4%)		0 (0.0%)	127 (100%)
SI (ppm)					
<15 (below adequate)		≥15 (adequate)		≥30 (above adequate)	
25 (19.7%)		74 (58.3%)		28 (22.0%)	
127 (100%)					

Data presented in n (%), where n is children number. Note FT₄: Free thyroxine, FT₃: Free triiodothyronine, TSH: Thyroid stimulating hormone, Tg-Ab: Thyroglobulin antibody, TPO-Ab: Thyroid peroxidase antibody, UI: Urinary iodine, WI: Water iodine, SI: Salt iodine, ID: Iodine deficiency, *Values as per assay manuals

confirmed our previous report.¹⁰ However, this study showed a positive correlation between UI and WI ($p < 0.05$) as was observed by Lv et al.²⁴ All these suggested that though these villages were iodine deficient, water available in these villages was a significant source of the consumed iodine.

Generally, achieving ≥ 15 ppm iodine in $\geq 90\%$ salt samples indicate USI program's satisfactory effectiveness.⁴ As per this criterion, effectiveness of the ongoing USI program under this study was not satisfactory. However, USI program under this study was better in effectiveness than that of 1996–98.¹² Besides, SI had no correlation with UI. All these suggested that the ongoing USI program was not a significant source of consumed iodine.

Usually, Total Goiter Rates (TGR) of 0.0–4.9%, 5.0–19.9%, 20.0–29.9% and $\geq 30\%$ indicate 'no', 'mild', 'moderate' and 'severe' IDD, respectively.⁴ Besides, thyroid nodule prevalence in children normally varies within 1.0–1.5%.²⁵ This study showing all children with normal TV and 0.8% children with thyroid nodules suggested that IDD were not major problems in these villages. This study's TV was lower than that of excess IC area.²⁶

This study's village-wise and overall means/medians of FT₄, FT₃ and TSH were in normal ranges. This study's overall means of FT₄ and FT₃ were higher and lower, respectively than that of excess IC area,²⁶ and higher in both cases than that of AIC area.²⁷ This study's overall median TSH was lower than

that of excess IC²⁶ and AIC areas.²⁷ Overall, 86.1% studied children were euthyroid. This study also showed high FT₃ in 17.3% children, being highest (34.1% children) in Unakoti having highest mean and median levels of BSC, TSH and FT₃. All these suggested that BS might have increased FT₃, probably by elevating TSH. This study's thyroid dysfunction prevalence (1.6% subclinical hypothyroidism and 2.4% euthyroid hypothyroxinemia) was lower than that of AIC area.²⁷

This study's overall median Tg-Ab was lower than that of AIC area,²⁷ and the prevalence of positive Tg-Ab (0%) and positive TPO-Ab (0.8%) were lower than that of excess IC area.²⁶ Generally, AITD prevalence in children varies within 0.5–3% in areas having no ID.²⁸ Thus, it appeared that positive thyroid auto-antibodies were not major problems in these villages.

This study's median BSC (grams/person/day) during BS-growing seasons varied from 83.3 in Deorachhara, 119 in Naithangchhara to 166.7 in Unakoti, overall being 119 which were higher than that of Australia and New Zealand (where BSC varied within 31–70 grams/person/day).²⁹ This study showing higher BSC ($p = 0.037$) associated with higher levels of TSH ($p = 0.034$) and FT₃ ($p = 0.001$), and lower levels of FT₄ ($p = 0.002$), height ($p = 0.009$), weight ($p < 0.001$), BMI ($p < 0.001$) and BSA ($p = 0.001$) in Unakoti than in Deorachhara-suggested that higher BSC might have lowered FT₄, subsequently impacting height, weight, BMI, and BSA, with a concurrent

Table 4. Ninth order partial correlation coefficients (r) and significance levels (p-values) for some variable pairs involving age, TV, FT₄, FT₃, TSH, Tg-Ab, TPO-Ab, UI, USCN, weight and height by village, gender and total among studied tribal children in Tripura

Variable pair	Deorachhara (n=45)	Unakoti (n=41)	Naithangchhara (n=41)	Boys (n=59)	Girls (n=68)	Total (n=127)
TV & weight	r=0.5351, p<0.001 (S)	r=-0.0014, p>0.05 (NS)	r=0.0823, p>0.05 (NS)	r=0.2048, p>0.05 (NS)	r=0.3213, p<0.02 (S)	r=0.2625, p<0.01 (S)
TV & height	r=-0.0844, p>0.05 (NS)	r=0.4432, p<0.02 (S)	r=0.0610, p>0.05 (NS)	r=0.0235, p>0.05 (NS)	r=0.1428, p>0.05 (NS)	r=0.0967, p>0.05 (NS)
FT ₄ & TSH	r=-0.1582, p>0.05 (NS)	r=0.2472, p>0.05 (NS)	r=-0.0623, p>0.05 (NS)	r=0.1102, p>0.05 (NS)	r=0.2676, p<0.05 (S)	r=0.1589, p>0.05 (NS)
FT ₄ & weight	r=0.1623, p>0.05 (NS)	r=-0.2080, p>0.05 (NS)	r=0.0853, p>0.05 (NS)	r=0.2408, p>0.05 (NS)	r=0.0801, p>0.05 (NS)	r=0.2005, p<0.05 (S)
Tg-Ab & TV	r=-0.3743, p<0.05 (S)	r=0.0788, p>0.05 (NS)	r=0.0725, p>0.05 (NS)	r=0.0323, p>0.05 (NS)	r=0.1646, p>0.05 (NS)	r=0.0710, p>0.05 (NS)
TPO-Ab & TSH	r=0.2909, p>0.05 (NS)	r=-0.3578, p<0.05 (S)	r=-0.2597, p>0.05 (NS)	r=-0.0313, p>0.05 (NS)	r=-0.0941, p>0.05 (NS)	r=-0.0719, p>0.05 (NS)
Tg-Ab & TPO-Ab	r=0.1073, p>0.05 (NS)	r=0.6258, p<0.001 (S)	r=0.3773, p<0.05 (S)	r=0.7366, p<0.001 (S)	r=0.3220, p<0.02 (S)	r=0.3822, p<0.001 (S)
Age & Tg-Ab	r=0.2373, p>0.05 (NS)	r=0.0506, p>0.05 (NS)	r=0.2498, p>0.05 (NS)	r=0.1031, p>0.05 (NS)	r=0.2749, p<0.05 (S)	r=0.2025, p<0.05 (S)
Tg-Ab & FT ₃	r=-0.0574, p>0.05 (NS)	r=0.00184, p>0.05 (NS)	r=-0.1067, p>0.05 (NS)	r=-0.2770, p<0.05 (S)	r=-0.0676, p>0.05 (NS)	r=-0.0974, p>0.05 (NS)
UI & TPO-Ab	r=0.1311, p>0.05 (NS)	r=0.0606, p>0.05 (NS)	r=0.1083, p>0.05 (NS)	r=0.3658, p<0.01 (S)	r=0.0234, p>0.05 (NS)	r=0.0731, p>0.05 (NS)
UI & weight	r=-0.3232, p>0.05 (NS)	r=0.1021, p>0.05 (NS)	r=-0.2893, p>0.05 (NS)	r=-0.2899, p<0.05 (S)	r=-0.1904, p>0.05 (NS)	r=-0.1878, p<0.05 (S)
UI & height	r=0.2758, p>0.05 (NS)	r=0.2534, p>0.05 (NS)	r=0.3733, p<0.05 (S)	r=0.3958, p<0.01 (S)	r=0.1520, p>0.05 (NS)	r=0.2425, p<0.01 (S)
Age & UI	r=-0.0765, p>0.05 (NS)	r=-0.2124, p>0.05 (NS)	r=-0.3865, p<0.05 (S)	r=-0.3208, p<0.05 (S)	r=0.0170, p>0.05 (NS)	r=-0.1580, p>0.05 (NS)
USCN & TPO-Ab	r=-0.0780, p>0.05 (NS)	r=0.0955, p>0.05 (NS)	r=-0.1556, p>0.05 (NS)	r=0.3196, p<0.05 (S)	r=-0.1170, p>0.05 (NS)	r=-0.0280, p>0.05 (NS)
USCN & FT ₄	r=0.0091, p>0.05 (NS)	r=0.0396, p>0.05 (NS)	r=0.5613, p<0.001 (S)	r=0.1520, p>0.05 (NS)	r=0.2068, p>0.05 (NS)	r=0.1969, p<0.05 (S)
TPO-Ab & FT ₃	r=0.4543, p<0.01 (S)	r=-0.1762, p>0.05 (NS)	r=-0.2117, p>0.05 (NS)	r=0.2278, p>0.05 (NS)	r=-0.0092, p>0.05 (NS)	r=0.0060, p>0.05 (NS)
TSH & weight	r=-0.0775, p>0.05 (NS)	r=0.0647, p>0.05 (NS)	r=0.2352, p>0.05 (NS)	r=-0.2553, p>0.05 (NS)	r=-0.2616, p<0.05 (S)	r=-0.2423, p<0.01 (S)
Tg-Ab & height	r=-0.1852, p>0.05 (NS)	r=-0.0158, p>0.05 (NS)	r=0.0376, p>0.05 (NS)	r=0.2952, p<0.05 (S)	r=-0.1465, p>0.05 (NS)	r=-0.0089, p>0.05 (NS)
Tg-Ab & weight	r=0.2731, p>0.05 (NS)	r=-0.1613, p>0.05 (NS)	r=-0.2382, p>0.05 (NS)	r=-0.3274, p<0.05 (S)	r=-0.0859, p>0.05 (NS)	r=-0.1257, p>0.05 (NS)
TPO-Ab & weight	r=0.0628, p>0.05 (NS)	r=0.1632, p>0.05 (NS)	r=0.1679, p>0.05 (NS)	r=0.3196, p<0.05 (S)	r=0.0945, p>0.05 (NS)	r=0.0597, p>0.05 (NS)

TV: Thyroid volume, FT₄: Free thyroxine, FT₃: Free triiodothyronine, TSH: Thyroid stimulating hormone, Tg-Ab: Thyroglobulin antibody, TPO-Ab: Thyroid peroxidase antibody, UI: Urinary iodine, USCN: Urinary thiocyanate, n: Children number; S: Significant, NS: Not significant

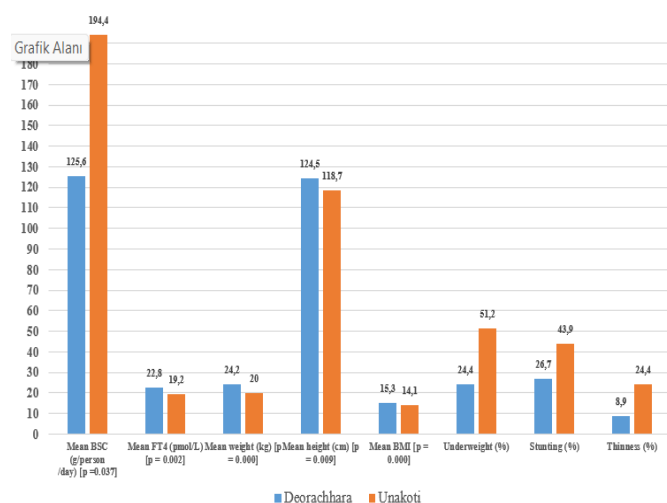


Figure 1. Graphical representation of the levels of BSC (bamboo shoot consumption), FT₄ (free thyroxine), weight, height, and BMI, alongside the prevalence of malnutrition (under-weight, stunting, and thinness) in Unakoti (which has a higher BSC and a lower FT₄) and Deorachhara (which has a lower BSC and a higher FT₄)

BSC: Bamboo shoot consumption, FT₄: Free thyroxine, BMI: Body-mass index

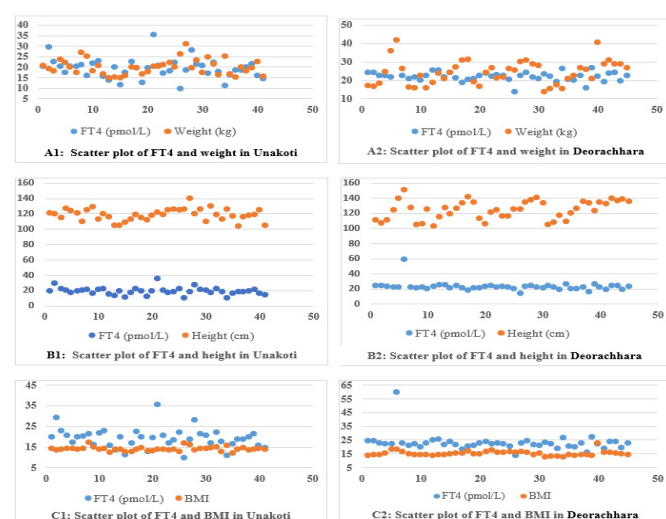


Figure 2. Scatter plots of FT₄ (free thyroxine) against weight, height, and BMI (body-mass index) in Unakoti (which has a higher BSC [bamboo shoot consumption] and a lower FT₄) [panels A1, B1, and C1] and Deorachhara (which has a lower BSC and a higher FT₄) [panels A2, B2, and C2], respectively

BSC: Bamboo shoot consumption, FT₄: Free thyroxine, BMI: Body-mass index



compensatory increase in TSH and FT_3 . This scenario aligned with the known effects of primary hypothyroidism. Besides, this study showing the trends of positive correlation between Tg-Ab and TPO-Ab proportionate to BSC in three villages, being significant in Unakoti ($p<0.001$), Naithangchhara ($p<0.05$) and total children ($p<0.001$) suggested a positive association between Tg-Ab and TPO-Ab, and this was possibly because of increased BSC. This study also showed the trends of positive to negative correlations (i) between TV and height, being significantly positive in Unakoti having highest BSC ($p<0.02$), and (ii) between TPO-Ab and TSH, being significantly negative in Unakoti having highest BSC ($p<0.05$). These correlations suggested (i) a positive association between TV and height, a finding that has been previously reported,³⁰ and (ii) a negative association between TPO-Ab and TSH, a finding that contradicted previously reported positive correlation between these two,³¹ and these were possibly because of higher BSC. Furthermore, this study showed the trends of positive to negative correlations (i) between TV and Tg-Ab, being significantly negative in Deorachhara having lowest BSC ($p<0.05$), and (ii) between FT_4 and weight, being positive in the village having lower BSC and significantly positive in total children ($p<0.05$). These correlations suggested (i) a negative association between TV and Tg-Ab, a finding that has been previously reported,³² and (ii) a positive association between FT_4 and weight in total children, and these were possibly because of lower BSC. But this study was unable to ascertain cause-effect relationship among the variables of these correlations. This study also showed the trends of positive to negative correlations between TV and weight, being significantly positive in Deorachhara having lowest BSC ($p<0.001$) and total children ($p<0.01$). These correlations suggesting a positive association between TV and weight, a finding that has been previously reported,³⁰ indicated that higher weight enhanced TV, a finding that has also been previously reported,³³ and these were probably because of lower BSC. Besides, this study showed the trends of positive to negative correlations between TSH and FT_4 , being positive in the village having higher BSC and significantly positive in girls ($p<0.05$). This positive correlation suggested that elevated TSH enhanced FT_4 in girls, potentially increasing the risk of TSH-mediated hyperthyroidism, and this effect was possibly because of higher BSC. Literature has already reported BS's association with reduced T_4 and increased TV in rats.³⁴ Besides, there have been some reports on BS's antithyroid action. BS has been reported to contain thiocyanate¹² which decreases thyroidal iodine uptake,²⁹ and 2,4-dihydroxy benzoic acid³⁵ which inhibits thyroid peroxidase (TPO) activity.³⁶ BS has also been reported to reduce TH synthesis decreasing the activities of sodium-potassium adenosine triphosphatase ($Na^+-K^+-ATPase$) and diiodinase.³⁷ However, it has already been reported that iodine can partially mitigate BS's antithyroid action.^{34,38}

This study's median UI ($\mu g/L$) were 103 in Deorachhara, 143 in Unakoti, and 100 in Naithangchhara. This study showing lower UI ($p=0.014$) associated with lower levels of FT_4 ($p=0.001$), FT_3 ($p<0.001$) and TSH ($p<0.001$), and higher levels of Tg-Ab ($p=0.003$) and BMI ($p<0.001$) in Naithangchhara than in Unakoti-suggested that lower AIC might have (i) decreased FT_4 and FT_3 , which remained associated with decreased TSH, and (ii) increased BMI and

Tg-Ab. This scenario reflected a potential risk of secondary/central hypothyroidism. These findings appeared reflecting the findings of a previous study²⁷ that reported Tg-Ab's strong associations with lower levels of UI, FT_4 and FT_3 in an area having median UI of $150 \mu g/L$. Besides, this study showed the trends of positive correlation between TV and BSA proportionate to UI in three villages, being significant in Unakoti ($p<0.01$), Deorachhara ($p<0.01$) and total children ($p<0.001$). These correlations suggested a positive association between TV and BSA, a finding that has been previously reported,³⁰ and this was possibly because of higher AIC. This study showing the trends of negative to positive correlations between FT_4 and BSA, being positive in the village having lower UI and significantly positive in total children ($p<0.01$) suggested a positive association between FT_4 and BSA, and this was possibly because of lower AIC. However, this study was unable to decide cause-effect relationship among the variables of these correlations. Furthermore, this study showed the trends of negative to positive correlations (i) between age and Tg-Ab, being positive in the village having lower UI and significantly positive in total children ($p<0.05$), and (ii) between Tg-Ab and FT_3 , being negative in the village having lower UI and significantly negative in boys ($p<0.05$). These correlations suggested (i) an increase in Tg-Ab with the advancement of age in total children, a finding that has been previously reported in females,³⁹ and (ii) a negative association between Tg-Ab and FT_3 in boys, a finding that has been previously reported in school children,²⁷ and these were probably because of lower AIC.

This study showing a positive correlation between USCN and TPO-Ab ($p<0.05$) in boys suggested that thiocyanate increased TPO-Ab in boys, a finding that supported the reported hypothesis that excess thiocyanate may promote thyroid autoantibodies.⁴⁰ This study showing a positive correlation between UI and TPO-Ab ($p<0.01$) being associated with the positive correlation between USCN and TPO-Ab in boys suggested that both increased iodine excretion and increased thiocyanate elevated TPO-Ab, potentially indicating ID in boys. Because increased UI, while sometimes linked to increased IC, can also reflect increased iodine excretion, especially when influenced by thiocyanate.⁴¹ This study showing a negative correlation between age and UI ($p<0.05$) being associated with the positive correlation between UI and TPO-Ab in boys suggested that problem of increased TPO-Ab probably got reduced with the advancement of age in boys. This study showing a positive correlation between TPO-Ab and FT_3 in Deorachhara having highest mean USCN ($p<0.01$) suggested a positive association between TPO-Ab and FT_3 , a finding that contradicted the previously reported negative correlation,³¹ implying a probable link between elevated levels of TPO-Ab and FT_3 with ID which was potentially mediated by thiocyanate in Deorachhara. Because an increase in T_3 by ID has already been reported⁴ and an increase in TPO-Ab by ID has already been suggested in this study.

Furthermore, this study showing positive correlations between USCN and FT_4 in Naithangchhara having lowest UI ($p<0.001$) and total children ($p<0.05$) suggested that increased thiocyanate raised FT_4 in children, and lower AIC might have favored this. However, thiocyanate has already been reported to increase FT_4 .⁴² The mechanism through which thiocyanate



increases FT_4 is still unclear, as thiocyanate decreases TH synthesis inhibiting iodide transport and incorporation of iodide to thyroglobulin.⁴² However, thiocyanate has been reported to decrease protein-bound T_4 levels and thereby to increase FT_4 in rats.⁴²

Unakoti having highest BSC showed highest prevalence of underweight (51.2%) and thinness (24.4%), whereas Deorachhara having lowest BSC showed lowest prevalence of underweight (24.4%) and stunting (26.7%). Besides, Naithangchhara having lowest UI showed highest stunting prevalence (51.2%) and lowest thinness prevalence (4.9%). All these suggested that BS might have increased underweight and thinness, whereas lower AIC might have increased stunting and thereby lowered thinness. By gender, boys had highest prevalence of underweight (42.4%) and stunting (45.8%), whereas girls had highest thinness prevalence (17.6%). All these suggested that boys were more susceptible to underweight and stunting, and girls to thinness.

This study's overall means of age, weight, height, BMI and BSA were lower than that of the children drinking water containing higher iodine and having higher obesity prevalence.⁴³ This study showing positive correlations between UI and height in total children ($p<0.01$), and between UI and BSA in boys ($p<0.02$), and negative correlations between UI and weight ($p<0.05$), and between UI and BMI in total children ($p<0.02$) suggested that higher IC increased height sometimes increasing BSA, decreased weight and thereby lowered BMI, whereas lower IC increased weight and thereby enhanced BMI. The increase in height by increased IC was possibly because of boosting of bone development and growth by increased TH, specially FT_3 , resulted from increased IC. Because FT_3 has been reported to stimulate bone development and growth.⁴⁴ Besides, the decreases in weight and thereby decreases in BMI by increased IC might be linked to body fat reduction resulted from enhanced calorogenesis by increased TH, particularly FT_3 , resulted from increased IC. Because FT_3 has also been reported to elevate calorogenesis, increasing basal metabolic rate, thermogenesis and lipolysis.⁴⁴ Furthermore, this study showing positive correlations between FT_4 and weight ($p<0.05$), and between FT_4 and BSA ($p<0.01$) in total children suggested that increased FT_4 associated with increased weight was reflecting a trend of weight gain (associated with increased BSA). Because a study observed that individuals with higher baseline FT_4 showed preference for a high fat and a high protein diet and, therefore, hypothesized that individuals with higher baseline FT_4 would gain more weight.⁴⁵ Besides, TSH exhibiting negative correlation with weight in total children ($p<0.01$) in this study suggested that TSH reduced weight which was likely to be associated with body fat burning. Because literature reported that (i) the presence of TSH receptors in liver and adipocytes (in addition to thyroid) suggests TSH's role in influencing thermogenesis, adipogenesis, lipolysis/lipogenesis, and (ii) increased TSH in obese one is a compensatory mechanism to counter weight gain.⁴⁵ This study showing a positive correlation between Tg-Ab and height ($p<0.05$) in boys suggested that an increase in height might have produced relative ID by increasing iodine requirement against lower AIC, and this ID was probably linked to raised Tg-Ab in boys. This study showing

negative correlations between Tg-Ab and weight ($p<0.05$), and between Tg-Ab and BMI ($p<0.05$) suggested that Tg-Ab remained associated with lower weight, and consequently with lower BMI. However, this study couldn't settle cause-effect relationship between higher Tg-Ab and lower weight.

However, this study's Tg-Ab and TPO-Ab showing no positive associations with TSH suggested that they were not risk factors for differentiated thyroid cancer (DTC). Because the combination of Tg-Ab, TPO-Ab and high TSH was identified as a risk factor for DTC.⁴⁶

The strengths of this study were (1) conducting the research in three different villages, and (2) investigating the same fourteen variables in each village. This research design allowed for a more robust and nuanced understanding of the phenomena being studied by capturing diverse perspectives and analyzing the interplay of multiple variables across three villages with varied quantitative profiles.

Limitations

This study had several limitations. The study (1) being cross-sectional, was not always able to determine cause-effect relationship among the variable pairs, and (2) did not consider (a) intakes of (i) dietary nitrates, fluorides, calcium, magnesium and isothiocyanate which reduce absorption and incorporation of dietary iodine, and (ii) vegetables of Cruciferae family, cassava, maize, peanuts, walnuts, lima beans, sweet potatoes, millets, soya bean containing glucosinolates and goitrins which block iodine incorporation in TH and suppress T_4 secretion, and (b) status of iron, zinc, copper, selenium, vitamin A and vanadium-deficiency of which may cause hypothyroidism.¹

CONCLUSION

This study concludes that Tripura's studied tribal children with lower normal median FT_4 levels showed a higher prevalence of underweight and thinness, and were at risk of developing TSH-mediated hyperthyroidism and AITD (increased TPO-Ab and Tg-Ab levels), possibly due to elevated levels of thiocyanate and other antithyroid substances derived from higher BSC. Additionally, studied tribal children with lower normal median FT_4 and FT_3 levels showed a higher stunting prevalence, and were at risk of weight gain and increased Tg-Ab levels, possibly due to lower AIC. However, further studies are needed to confirm these findings and to elucidate the underlying mechanisms.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study was approved by the Institutional Human Ethical Committee, Department of Physiology, University of Calcutta, Kolkata, India (Date: 03.10.2013, Decision No:IHEC/PHY/CU/H-P 34/13).

Informed Consent

Informed consent was obtained from a parent or legal guardian. Where appropriate, age-adjusted assent was also obtained from the child. The inclusion of vulnerable populations in this study adhered to national and



international ethical guidelines. Extra care was taken to ensure voluntary participation, understanding, and protection of participant dignity and autonomy.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

This study was funded by the Department of Biotechnology (DBT), Ministry of Science Technology, Government of India, New Delhi under North Eastern Region (NER) Twinning Project MAP-240/2013 (Sanction Number BT/471/NE/TBP/2013, dated 26th December, 2014). Author Ajoy Datta received fellowship from DBT, Government of India.

Author Contributions

Concept: AKC; Design: IR, AKC; Control: AKC; Resources: IR; Materials: IR, AD; Data Collection and/or Processing: IR, AD; Analysis and/or Interpretation: IR, AKC, SKD, SKM; Literature Review: IR, AKC, SKD, SKM; Writing the Article: IR; Critical Review: AKC; Final Approval: IR, AKC, SKD, SKM, AD, LHS; Accountability: IR, AKC, SKD, SKM, AD, LHS.

REFERENCES

- Scientific Committee on Food (SCF). Opinion of the Scientific Committee on Food on the Tolerable Upper Iodine Level of Iodine. SCF/CS/NUT/UPPLEV/26 Final. European Commission (EC), Brussels-Belgium; 2002. Available from: https://ec.europa.eu/food/fs/sc/scf/out146_en.pdf [Last accessed on May 17, 2025]
- National Coordinating Committee on Food and Nutrition (NCCFN) of Ministry of Health Malaysia (MHM). Iodine (Chapter 19). In: Recommended Nutrient Intakes for Malaysia; 2017:342-355. Available from: <https://hq.moh.gov.my/nutrition/wp-content/uploads/2023/12/FA-Buku-RNI.pdf> [Last accessed on May 28, 2025]
- Welsh KJ, Soldin SJ. Diagnosis of endocrine disease: how reliable are free thyroid and total T₃ hormone assays? *Eur J Endocrinol*. 2016;175:R255-R263. doi:10.1530/EJE-16-0193
- World Health Organization (WHO). Assessment of iodine deficiency disorders and monitoring their elimination: a guide for programme managers.-3rd edn., Geneva; 2007. Available from: <https://www.who.int/publications/i/item/9789241595827> [Last accessed on May 28, 2025]
- World Health Organization (WHO). Global Prevalence of Iodine Deficiency Disorders. MDIS (Micronutrient Deficiency Information System) Working Paper #1, Geneva; 1993. Available from: <https://iris.who.int/handle/10665/37149> [Last accessed on May 28, 2025]
- Fitzgerald SP, Bean NG. Thyroid stimulating hormone (TSH) autoregulation reduces variation in the TSH response to thyroid hormones. *Temperature*. 2018;5(4):380-389. doi:10.1080/23328940.2018.1513110
- Wan S, Qu M, Wu H, et al. Autoimmune thyroid diseases after 25 years of universal salt iodisation: an epidemiological study of Chinese adults in areas with different water iodine levels. *Br J Nutr*. 2020;124:853-864. doi:10.1017/S0007114520001786
- Teng W, Shan Z, Teng X, et al. Effect of iodine intake on thyroid diseases in China. *N Eng J Med*. 2006;354:2783-2793. doi:10.1056/NEJMoa054022
- Chandra AK, Mukhopadhyay S, Lahiri D, Tripathy S. Goitrogenic content of Indian cyanogenic plant foods & their in vitro anti-thyroidal activity. *Indian J Med Res*. 2004;119(5):180-185.
- Chandra AK, Ray I, Ray P. Iodine content in water of Tripura in North Eastern India. *J Food Sci Technol*. 1999;36(6):558-560.
- Chandra AK, Ray I. Dietary supplies of iodine and thiocyanate on the etiology of endemic goiter in Tripura. *Indian J Pediatr*. 2001;68(5):399-404. doi:10.1007/BF02723011
- Chandra AK, Ray I. Evaluation of the effectiveness of salt iodization status in Tripura, north-east India. *Indian J Med Res*. 2002;115:22-27.
- Dunn JT, van der Haar F. A practical guide to the correction of iodine deficiency. Technical Manual no. 3, ICCIDD/UNICEF/WHO Publication, The Netherlands; 1990.
- World Health Organization (WHO). Physical Status: The Use and Interpretation of Anthropometry. Report of a WHO Expert Committee (WHO Technical Report Series: 854), Geneva; 1995. Available from: <https://www.who.int/publications/i/item/9241208546> [Last accessed on May 28, 2025]
- Cashin K, Oot L. Guide to Anthropometry: A Practical Tool for Program Planners, Managers, and Implementers. Food and Nutrition Technical Assistance III Project (FANTA)/ FHI 360, Washington DC; 2018. Available from: <https://www.fantaproject.org/tools/anthropometry-guide> [Last accessed on May 28, 2025]
- Centers for Disease Control and Prevention (CDC). Z-score Data Files-Growth Charts. National Center for Health Statistics, USA; 2009. Available from: <http://medbox.iiaab.me/modules/en-cdc/www.cdc.gov/growthcharts/zscore.htm> [Last accessed on May 28, 2025]
- Aw TC, Yap CYF. Thyroid function tests. *Proceedings of Singapore Healthcare*. 2011;20(2):132-137. doi:10.1177/201010581102000212
- Karmarkar MG, Pandav CS, Krishnamachari KAVR. Principle and procedure for iodine estimation. A laboratory manual. Indian Council of Medical Research (ICMR), New Delhi; 1986.
- Querido A, Delange F, Dunn T, et al. Definitions of endemic goiter and cretinism, classification of goiter size and severity of endemias and survey techniques. In: Dunn JT, Medeiros-Neto GA (eds) Endemic Goiter and Cretinism: Continuing Threats to World Health, Pan American Health Organization (PAHO), Scientific Publication No. 292, Washington DC; 1974:262-272.
- Michajlovskij N, Langer P. Studies on relations between thiocyanate formation and goitrogenic properties of foods. I. Preformed thiocyanate content of some foods [in German]. *Hoppe Seylers Z Physiol Chem*. 1958;312:26-30.
- World Health Organization (WHO). Indicators for assessing iodine deficiency disorders and their control through salt iodization. Micronutrient Series, WHO/NUT/94.6, Geneva; 1994. Available from: <https://iris.who.int/handle/10665/70715> [Last accessed on May 28, 2025]
- Das D, Das A. Statistics in Biology and Psychology, Sixth edn. Academic Publishers, Kolkata; 2015.
- Zeltser ME, Aidarkhanov BA, Berezhnaya IM, et al. Iodine deficiency and its clinical manifestations in Kazakhstan. International Council for Control of Iodine Deficiency Disorders (ICCIDD). *IDD Newsletter*. 1992;8(1):5-6.
- Lv S, Wang Y, Xu D, et al. Drinking water contributes to excessive iodine intake among children in Hebei, China. *Eur J Clin Nutr*. 2013;67:961-965. doi:10.1038/ejcn.2013.127
- Richman DM, Benson CB, Doubilet PM, et al. Thyroid nodules in pediatric patients: sonographic characteristics and likelihood of cancer. *Radiol*. 2018;288(2):591-599. doi:10.1148/radiol.2018171170
- Chen W, Ho Y, Wang W, et al. Adverse effects on thyroid of Chinese children exposed to long-term iodine excess: optimal and safe Tolerable Upper Intake Levels of iodine for 7- to 14-y-old children. *Am J Clin Nutr*. 2018;107:780-788. doi:10.1093/ajcn/nqy011
- Tamang B, Khatiwada S, Gelal B, et al. Association of antithyroglobulin antibody with iodine nutrition and thyroid dysfunction in Nepalese children. *Thyroid Res*. 2019;12:1-8. doi:10.1186_s13044-019-0067-z
- Zimmermann MB, Aeberli I, Andersson M, et al. Thyroglobulin is a sensitive measure of both deficient and excess iodine intakes in children and indicates no adverse effects on thyroid function in the UIC range of 100-299 µg/L: a UNICEF/ICCIDD Study Group Report. *J Clin Endocrinol Metab*. 2013;98(3):1271-1280. doi:10.1210/jc.2012-3952
- Food Standards Australia New Zealand (FSANZ). Cyanogenic glycosides in cassava and bamboo shoots: a human health risk assessment. Technical Report Series No. 28; 2004:1-24. Available from: https://www.foodstandards.gov.au/publications/documents/28_Cyanogenic_glycosides.pdf [Last accessed on April 19, 2021]
- Aydiner O, Aydinler EK, Akpınar I, Turan S, Bereket A. Normative data of thyroid volume-ultrasonographic evaluation of 422 subjects aged 0-55 years. *Clin Res Pediatr Endocrinol*. 2015;7(2):98-101. doi:10.4274/jcrpe.1818
- Gautam B, Poudel S, Pokhrel S. Significance of thyroid peroxidase antibodies (anti-TPO antibodies) in patients undergoing thyroid function tests. *J Gandaki Med College-Nepal*. 2024;17(2):111-116. doi:10.3126/jgmc-n.v17i2.71641



32. Lin Z, Lu C, Teng D, et al. Influencing factors and new reference intervals of adult thyroid volume in iodine-sufficient areas of China. *Biological Trace Element Res.* 2023;201:5652-5661. doi:10.1007/s12011-023-03635-x
33. Liu L, Liu L, Li M, et al. Study on association between height, weight, iodine supplementation and thyroid volume. *Br J Nutr.* 2022;127:1358-1366. doi:10.1017/S0007114521002622
34. Chandra AK, Ghosh D, Mukhopadhyay S, Tripathy S. Effect of bamboo shoot, *Bambusa arundinacea* (Retz.) Wild. on thyroid status under conditions of varying iodine intake in rats. *Indian J Exp Biol.* 2004;42:781-786.
35. Dey SK, Brahma BK, Goyal AK. GC-MS profiling of phytocomponents from hydromethanolic fraction of *Bambusa tulda* leaf: a potential candidate for anti-diabetic activity. *Int J Life Sci Res.* 2018;6(3):382-388.
36. Khadem S, Marles RJ. Monocyclic phenolic acids; hydroxy- and polyhydroxybenzoic acids: occurrence and recent bioactivity studies. *Molecules.* 2010;15(11):7985-8005. doi:10.3390/molecules15117985
37. Sarkar D, Chandra AK, Chakraborty A, et al. Effects of bamboo shoots (*Bambusa balcooa*) on thyroid hormone synthesizing regulatory elements at cellular and molecular levels in thyrocytes. *J Ethnopharmacol.* 2020;250:112463. doi:10.1016/j.jep.2019.112463
38. Sarkar D, Chandra AK, Chattopadhyay S, et al. Possible mechanism of bamboo shoots (*Bambusa balcooa*) induced thyroid disruption—an in vitro study. *Hum Exp Toxicol.* 2020;10:960327120958037. doi:10.1177/0960327120958037
39. Ma C, Li D, Yin Y, et al. Establishing thresholds and effects of gender, age, and season for thyroglobulin and thyroid peroxidase antibodies by mining real-world big data. *Clin Biochem.* 2019;74:36-41. doi:10.1016/j.clinbiochem.2019.08.011
40. Singh LH, Chandra AK, Yumnam SD, et al. Thiocyanate in excess develops goiter followed by autoimmune thyroid diseases even after effective salt iodization in a rural community of north east India. *Ecotox Environ Safe.* 2021;208:111711. doi:10.1016/j.ecoenv.2020.111711
41. Miller JK, Moss BR, Swanson EW, Lyke WA. Effect of thyroid status and thiocyanate on absorption and excretion of iodine by cattle. *J Dairy Sci.* 1975;58(4):526-531.
42. Leko MB, Gunjaca I, Pleic N, Zemunik T. Environmental factors affecting thyroid-stimulating hormone and thyroid hormone levels. *Int J Mol Sci.* 2021;22:6521. doi:10.3390/ijms22126521
43. Chen W, Li X, Wu Y, et al. Associations between iodine intake, thyroid volume, and goiter rate in school-aged Chinese children from areas with high iodine drinking water concentrations. *Am J Clin Nutr.* 2017;105:228-233. doi:10.3945/ajcn.116.139725
44. Leung AM, Brent GA. The Influence of Thyroid Hormone on Growth Hormone Secretion and Action (Chapter 4). In: Cohen LE (ed) *Growth Hormone Deficiency: Physiology and Clinical Management*, Springer; 2016. doi:10.1007/978-3-319-28038-7
45. Stanikova D, Krajcovicova L, Lobotkova D, et al. Thyroid hormone levels and BMI-SDS changes in adolescents with obesity. *Front Endocrinol.* 2023;14:1304970. doi:10.3389/fendo.2023.1304970
46. Frohlich E, Wahl R. Thyroid autoimmunity: role of anti-thyroid antibodies in thyroid and extra-thyroidal diseases. *Front Immunol.* 2017;8:521. doi:10.3389/fimmu.2017.00521

Beyond fecal diversion: stoma-related morbidity in a cohort of 318 children

©Abdurrahman Urve Uzun^{*1}, ©Can İhsan Öztörün²¹Department of Pediatric Surgery, Gülhane Training and Research Hospital, University of Health Sciences, Ankara, Türkiye²Department of Pediatric Surgery, Faculty of Medicine, Ankara Yıldırım Beyazıt University, Ankara, Türkiye

Received: 11/09/2026

Accepted: 11/10/2026

Published: 20/10/2026

Cite this article: Uzun AU, Öztörün Cİ. Beyond fecal diversion: stoma-related morbidity in a cohort of 318 children. *Surg Child.* 2026;3(4):149-154. doi:10.51271/SOC-0088

*Corresponding Author: Abdurrahman Urve Uzun, uzunurve@gmail.com

ABSTRACT

Aims: To evaluate complications occurring from intestinal stoma formation to closure in children and to investigate the associations between complication development and demographic and clinical characteristics.

Methods: This single-center retrospective study included 318 pediatric patients who underwent intestinal stoma formation between January 1, 2019, and December 31, 2025. Demographic and clinical characteristics and complications occurring before stoma closure were evaluated. Groups were compared using Pearson's Chi-square test and Welch's t-test. Variables associated with complication development were examined using multivariable logistic regression analysis. Collinearity between gestational age and birth weight was assessed using Pearson correlation and variance inflation factors (VIF), and three multivariable models were compared.

Results: At least one complication was identified in 105 patients (33.0%). Peristomal dermatitis occurred in 54 (17.0%), stoma prolapse in 22 (6.9%), wound dehiscence in 20 (6.3%), bleeding in six (1.9%), and adhesive/band-related ileus in three patients (0.9%). The complication rate was 38.2% in the emergency surgery group and 21.4% in the elective group ($p=0.003$). Patients who developed complications had a lower gestational age (37.8 ± 4.2 vs 39.1 ± 4.4 weeks; $p=0.011$) and lower birth weight (2360 ± 480 vs 2510 ± 310 g; $p=0.004$). Eight patients with prolapse underwent surgical intervention, whereas 14 were managed with close observation. Gestational age and birth weight were strongly correlated ($r=0.79$; VIF 2.62). In the model including both variables, emergency surgery (odds ratio [OR] 2.42; 95% confidence interval [CI] 1.38-4.27) and birth weight (OR 0.88 per 100 g; 95% CI 0.79-0.98) remained associated with complications, whereas gestational age did not (OR 1.02 per week; 95% CI 0.93-1.11).

Conclusion: Stoma-related morbidity affected approximately one-third of children during the diversion period, underscoring that stoma formation represents a clinically relevant phase of care rather than a neutral bridge to definitive surgery. Emergency surgery and lower birth weight were independently associated with complication development, whereas the association of gestational age was not independent of birth weight. These findings support structured surveillance of stoma function, skin integrity, growth, and nutritional status throughout the period of diversion, particularly in higher-risk children.

Keywords: Child, enterostomy, colostomy, ileostomy, postoperative complications

INTRODUCTION

Intestinal stoma formation plays an important role in the surgical management of congenital and acquired gastrointestinal disorders in children. Fecal diversion is used to achieve bowel decompression, protect a distal bowel segment or anastomosis, and facilitate staged surgical treatment. Nevertheless, from the time of stoma formation until restoration of intestinal continuity, the stoma itself constitutes an additional source of morbidity. Peristomal skin problems, prolapse, stenosis, retraction, and complications requiring reoperation have all been reported in children.^{1,2}

In patients with anorectal malformations in particular, stoma location and configuration may influence both stoma care and subsequent reconstruction.³

Reported complication rates in pediatric studies vary widely because of differences in patient characteristics, follow-up duration, and complication definitions. Some studies have evaluated only major complications requiring intervention, whereas others have also included events managed conservatively. Vogel et al.⁴ separately examined



major morbidity after stoma formation and closure in young children, whereas Tan and Deogracias⁵ focused on major complications in a large pediatric population. Neonatal series that also include minor events provide a broader picture of stoma-related morbidity.⁶ Accordingly, the assessment period and outcome definitions must be clearly specified when complication rates are interpreted.

Gestational age, birth weight, and clinical condition at the time of surgery are patient-related factors that may be associated with complication development. Associations have been reported between lower birth weight and gestational age and complications in infants, and between lower preoperative weight and complications in neonates with necrotizing enterocolitis.^{7,8} However, these associations have not been consistent in studies using reoperation as the outcome.⁹ In addition, growth may be adversely affected while an enterostomy remains in situ in neonates.¹⁰

The aim of this study was to evaluate complications occurring until stoma closure or the end of available follow-up in 318 children who underwent intestinal stoma formation and to investigate associations between complication development and demographic and clinical characteristics.

METHODS

Ethics

The study was approved by the Clinical Researches Ethics Committee of Ankara Bilkent City Hospital (Date: 15.06.2026, Decision No: E2-26-368) and was conducted in accordance with the principles of the Declaration of Helsinki. Owing to the retrospective nature of the study, the requirement for individual informed consent was waived by the ethics committee.

Study Design and Patient Population

This single-center retrospective study included 318 pediatric patients who underwent intestinal stoma formation between January 1, 2019, and December 31, 2025. Gender, gestational age, birth weight, presence of comorbidity, urgency of surgery, stoma type, and complication development were evaluated.

Comorbidity was defined as the presence of cardiac, renal, neurologic, or other associated anomalies documented in the medical records, excluding the primary disease that prompted stoma formation. Surgical procedures were classified as emergency or elective. Stomas were analyzed in two groups: colostomies and small-bowel stomas; ileostomies, jejunostomies, and duodenostomies were included in the small-bowel stoma group. Patients with anomalies affecting more than one organ system were counted in each applicable descriptive category; for comparative analyses, comorbidity was coded as present or absent. The small-bowel stoma group was formed because of the small number of patients in the individual subgroups.

Assessment of Complications

The primary outcome was the development of at least one stoma-related complication or related surgical wound complication during the period from stoma formation to closure. For patients whose stoma had not been closed, follow-

up was censored at the date of the last clinical record within the study period; for patients who died before closure, follow-up was censored at the date of death. Surgical complications related to or occurring after stoma closure were excluded. Dermatitis and severe dermatitis were analyzed within the same category.

Management of prolapse and episodes of adhesive/band-related ileus occurring before stoma closure were recorded. Anthropometric course and nutritional support were evaluated as secondary descriptive outcomes.

Weight-for-age measurements were obtained retrospectively from routine clinical records. Weight-for-age z scores were calculated using the WHO Child Growth Standards.¹¹ Enteral nutritional support was provided to patients with a z score below -2. Measurement dates were not standardized, and anthropometric findings were analyzed descriptively.

Statistical Analysis

Categorical variables were presented as number and percentage, and continuous variables as mean±standard deviation. Patients with and without complications were compared using Pearson's chi-square test and Welch's t-test.

Variables associated with complication development were evaluated using multivariable binary logistic regression analysis. Regression results were reported as odds ratios (ORs), 95% confidence intervals (CIs), and p values. All tests were two-sided, and $p < 0.05$ was considered statistically significant. Gestational age was entered per one-week and birth weight per 100-g increase. Collinearity was assessed using the Pearson correlation coefficient and VIF (>5 considered problematic). Three multivariable logistic regression models were fitted: model 1, urgency+gestational age+birth weight; model 2, urgency+gestational age; model 3, urgency+birth weight. Model fit was compared using the Akaike information criterion (AIC). Analyses were performed using Python (SciPy, statsmodels).

RESULTS

Patient Characteristics

Of the 318 patients included in the study, 186 (58.5%) were boys and 132 (41.5%) were girls. Stoma formation was performed under emergency conditions in 220 patients (69.2%) and electively in 98 (30.8%). A colostomy was present in 143 patients (45.0%), whereas 175 (55.0%) had a small-bowel stoma. Comorbidities were recorded in 88 patients (27.7%). Within the small-bowel stoma group, 128 patients had an ileostomy (73.1%), 42 a jejunostomy (24.0%), and 5 a duodenostomy (2.9%). Among patients with comorbidities, cardiovascular anomalies were present in 52 (16.4% of the cohort), renal in 23 (7.2%), and neurological in 20 (6.3%); categories were not mutually exclusive.

At least one complication developed in 105 patients (33.0%), whereas no complication was documented during the available follow-up in 213 (67.0%). Demographic and clinical characteristics according to complication status are presented in **Table 1**.

**Table 1.** Demographic and clinical characteristics of patients with and without complications

Variable	Complication present (n=105)	No complication (n=213)	P
Gender, male/female, n	63/42	123/90	0.701
Gestational age, weeks, mean±SD	37.8±4.2	39.1±4.4	0.011
Birth weight, g, mean±SD	2360±480	2510±310	0.004
Comorbidity, present/absent, n	25/80	63/150	0.280
Emergency stoma formation, n (%)	84 (38.2)	136 (61.8)	0.003
Elective stoma formation, n (%)	21 (21.4)	77 (78.6)	-
Colostomy, n (%)	53 (37.1)	90 (62.9)	0.166
Small-bowel stoma, n (%)	52 (29.7)	123 (70.3)	-

SD: Standard deviation. Percentages are row percentages within the relevant surgical urgency or stoma category. Pearson's Chi-square test was used for categorical variables and Welch's t-test for continuous variables. A single p value is reported for each comparison of surgical urgency and stoma type. The small-bowel stoma group includes ileostomy, jejunostomy, and duodenostomy.

Factors Associated with Complication Development

Complications developed in 84 of 220 patients (38.2%) who underwent emergency stoma formation and in 21 of 98 patients (21.4%) who underwent elective stoma formation. Emergency surgery was associated with higher unadjusted odds of complication development (OR, 2.26; 95% CI, 1.30-3.94; $p=0.003$).

Mean gestational age was 37.8±4.2 weeks in the complication group and 39.1±4.4 weeks in the group without complications. The mean between-group difference was -1.30 weeks (95% CI, -2.30 to -0.30; $p=0.011$). Mean birth weight was 2360±480 g and 2510±310 g, respectively, corresponding to a mean difference of -150 g (95% CI, -251.64 to -48.36; $p=0.004$).

No statistically significant association was identified between gender and complication development (male/female OR, 1.10; 95% CI, 0.68-1.77; $p=0.701$). Similarly, the presence of comorbidity was not significantly associated with complications (OR, 0.74; 95% CI, 0.43-1.27; $p=0.280$).

The complication rate was 37.1% (53/143) among patients with colostomies and 29.7% (52/175) among those with small-bowel stomas. This difference was not statistically significant (unadjusted OR, 1.39; 95% CI, 0.87-2.23; $p=0.166$).

Gestational age and birth weight were strongly correlated (Pearson $r=0.79$, $p<0.001$). In model 1, VIFs were 2.62 for both gestational age and birth weight and 1.00 for urgency. In model 1, emergency surgery (OR 2.42; 95% CI 1.38-4.27; $p=0.002$) and birth weight (OR 0.88 per 100 g; 95% CI

0.79-0.98; $p=0.016$) were independently associated with complications, whereas gestational age was not (OR 1.02 per week; 95% CI 0.93-1.11; $p=0.708$). In model 2, gestational age was associated with complications (OR 0.93; 95% CI 0.89-0.99; $p=0.012$), and in model 3 birth weight was associated (OR 0.89 per 100 g; 95% CI 0.84-0.95; $p<0.001$). Emergency surgery remained associated in all models (OR 2.30-2.42). Model 3 had the lowest AIC (388.3 vs 390.2 for model 1 and 394.1 for model 2) (Table 2). Crude complication rates within the small-bowel stoma group were 22/128 for ileostomies (17.2%), 25/42 for jejunostomies (59.5%), and 5/5 for duodenostomies.

Distribution of Complications

Peristomal dermatitis was recorded in 54 patients (17.0%), stoma prolapse in 22 (6.9%), wound dehiscence in 20 (6.3%), bleeding in six (1.9%), and adhesive/band-related ileus in three (0.9%) (Table 3).

Table 3. Complications recorded before stoma closure

Complication	Number of patients, n	Entire cohort, %
At least one complication	105	33.0
Peristomal dermatitis	54	17.0
Stoma prolapse	22	6.9
Wound dehiscence	20	6.3
Bleeding	6	1.9
Adhesive/band-related ileus	3	0.9

Eight of the 22 patients with prolapse (36.4%) underwent surgical intervention. The remaining 14 patients (63.6%), in whom intestinal content continued to pass through the stoma, were managed with close observation without additional surgery for prolapse; stoma closure was performed at the earliest clinically appropriate time. Three patients with adhesive/band-related ileus underwent adhesiolysis before stoma closure.

Review of follow-up records for weight-for-age suggested a less favorable pattern of weight gain among children with small-bowel stomas, whereas age-appropriate weight was generally maintained in patients with colostomies. Enteral nutritional support was provided to patients with a weight-for-age z score below -2. Among patients assessed after stoma closure, an improvement in weight gain was observed. Because measurements were not obtained at standardized time points, these observations do not establish stoma-specific growth trajectories or a causal effect of stoma closure on growth.

Table 2. Multivariable logistic regression analysis of factors associated with complication development

Variable	Model 1	Model 2	Model 3
Emergency vs elective surgery, OR (95% CI); p	2.42 (1.38-4.27); 0.002	2.30 (1.32-4.03); 0.004	2.41 (1.37-4.25); 0.002
Gestational age (per week), OR (95% CI); p	1.02 (0.93-1.11); 0.708	0.93 (0.89-0.99); 0.012	-
Birth weight (per 100 g), OR (95% CI); p	0.88 (0.79-0.98); 0.016	-	0.89 (0.84-0.95); <0.001
AIC	390.2	394.1	388.3

OR: Odds ratio, CI: Confidence interval, AIC: Akaike information criterion, VIF: Variance inflation factor. Model 1: Urgency+gestational age+birth weight, Model 2: Urgency+gestational age, Model 3: Urgency+birth weight. Pearson r (gestational age-birth weight)=0.79; VIF 2.62 for both variables in model 1.



DISCUSSION

In this study, at least one complication was recorded while the stoma remained in situ in 105 of 318 children (33.0%) who underwent intestinal stoma formation. Peristomal dermatitis was the most common complication, followed by prolapse and wound dehiscence. Emergency stoma formation and lower birth weight were independently associated with complication development, whereas lower gestational age was not independent of birth weight. Approximately one-third of patients with prolapse required surgical intervention.

Comparison of complication rates across pediatric stoma studies requires caution because of differences in patient populations and outcome definitions. Steinau et al.¹ reported a complication rate of 38.2% after enterostomy formation. Although this is close to the 33.0% rate observed in our cohort, the studies differ in age distribution, primary disease, and follow-up characteristics. Sarac reported¹² greater morbidity in a smaller pediatric series, with skin complications being particularly prominent. In the study by Massenga et al.,¹³ which included both children and adults, the complication rate in the pediatric subgroup was 26.7%. Only the separately reported pediatric results from that study are appropriate for comparison with our cohort.

The severity threshold used to define a complication also directly influences reported rates. Vogel et al.⁴ reported major morbidity rates of 27% after stoma formation and 23% after closure in young children. In the large pediatric series by Tan and Deogracias,⁵ the rate of major complications was 9.8%. These rates are not directly equivalent to the outcome used in our study, which also included conservatively managed events such as dermatitis. The neonatal enterostomy study by Wolf et al.⁶ likewise demonstrated that inclusion of minor events broadens the estimated overall burden of morbidity. Therefore, differences in complication rates between centers may reflect not only surgical outcomes but also how events are defined and documented.

The outcome in our study was restricted to the period from stoma formation to stoma closure. Events related to the closure procedure or occurring after closure were not included. In studies encompassing both stoma formation and closure, such as the study by Alghamdi et al.,¹⁴ wound-related and intestinal obstruction outcomes should be interpreted with this distinction in mind. The study by O'Connor and Sawin¹⁵ in premature infants also demonstrated substantial morbidity; however, assessment among survivors and inclusion of closure-related events limit direct comparison with our study.

The inclusion of patients whose stomas had not yet been closed or who died before closure also warrants consideration. In these patients, follow-up was censored at the last clinical record or at death. Accordingly, the rate reported in our study represents the frequency of complications documented over observation periods of varying duration; it does not represent the cumulative probability that every patient would experience a complication by the time of stoma closure.

The complication rate was higher among patients who underwent emergency stoma formation than among those undergoing elective procedures. The nature of the underlying disease requiring urgent intervention, physiological instability, and perioperative inflammatory burden may have contributed to this association. Because these features could not be evaluated in sufficient detail in our study, the direct effect of emergency surgery cannot be disentangled from the effects of the primary disease and overall clinical condition.

Xu et al.¹⁶ reported univariable associations between nutritional status, inflammation, and incisional complications in children with very-early-onset inflammatory bowel disease caused by defects in interleukin-10 signaling; however, no statistically significant independent factor was identified in multivariable analysis. Although findings from this specific patient group cannot be directly generalized to the broader pediatric stoma population, they underscore the need to consider surgical urgency together with disease severity. The association observed in our study should not be interpreted as suggesting that clinically necessary emergency surgery should be delayed.

The lower gestational age and birth weight observed among patients who developed complications are consistent with the findings of Bethell et al. in infants with enterostomies.⁷ Aguayo et al.⁸ also reported associations between lower gestational age, lower preoperative weight, and stoma complications in neonates with necrotizing enterocolitis. However, birth weight and weight at the time of surgery are distinct clinical measures and should not be used interchangeably.

Bælum et al.⁹ found no significant difference in birth weight or gestational age between infants with necrotizing enterocolitis who did and did not require reoperation. This discrepancy may be related to patient selection, sample size, and the use of reoperation as the outcome. In our study, gestational age and birth weight were strongly correlated ($r=0.79$). Although VIF values (2.62) were below the conventional threshold, this correlation limits separation of their independent effects. Gestational age lost its association after adjustment for birth weight, whereas birth weight remained associated, and the model with birth weight alone had the lowest AIC. Birth weight may therefore capture most of the maturity-related risk in this cohort, but this should be interpreted cautiously.

No statistically significant difference in complication frequency was observed between the colostomy and small-bowel stoma groups. However, this finding does not establish that the two stoma groups have equivalent complication profiles. Vogel et al.⁴ identified ileostomy as being associated with major morbidity, whereas the prediction model developed by Zeng et al.¹⁷ in infants identified small-bowel stoma as a variable associated with complication development. Differences in age, primary disease, complication severity, and analytical methods may explain the divergence between those studies and our findings.

In our study, the small-bowel stoma group included ileostomies, jejunostomies, and duodenostomies, which limits evaluation of differences related to the anatomical level



of the stoma. In addition, comparisons based on broad stoma categories do not adequately capture stoma location and technical configuration.

Ileostomies (73.1%), jejunostomies (24.0%) and duodenostomies (2.9%) may differ in absorptive capacity and fluid/electrolyte losses. Crude complication rates differed across subtypes (17.2%, 59.5% and 5/5, respectively), so the nonsignificant colostomy–small-bowel comparison may mask subtype-specific patterns.

Cardiovascular anomalies were the most frequently documented associated abnormalities, followed by renal and neurological anomalies. The composite binary comorbidity variable was not significantly associated with complications; this does not exclude subtype-specific risks.

The large colostomy series by Çiğdem et al.,² the experience of Peña et al.³ in anorectal malformations, and the more recent study by Klein et al.¹⁸ highlight the importance of anatomical and technical features of colostomy formation. In light of these studies, details such as the bowel segment used, loop versus divided stoma configuration, and the relationship of the stoma to the planned definitive operation should be recorded in a standardized manner. Prolapse and skin complications were also prominent in the study by Hamdi et al.¹⁹ in patients with anorectal malformations. Nevertheless, such disease-specific series should not be used as a universal benchmark for all pediatric patients with stomas.

Peristomal dermatitis was recorded in 54 patients (17.0%) and was the most frequent complication. This finding demonstrates that stoma-related morbidity is not limited to events requiring reoperation. Structured monitoring of skin findings and stoma-care problems may help distinguish transient events from persistent care-related difficulties. Our study was not designed to compare the effectiveness of specific stoma-care products or nursing interventions.

Prolapse was recorded in 22 patients (6.9%), eight of whom required surgical intervention. The 14 patients in whom intestinal content continued to pass through the stoma were closely observed, and stoma closure was undertaken at the earliest clinically appropriate time. This approach illustrates that prolapse management was individualized according to clinical findings and bowel function. Johnson's²⁰ systematic review and retrospective series highlighted the range of treatment strategies used in children and the limited comparative evidence available. Because treatment allocation was not randomized in our study, the relative superiority of conservative and surgical approaches cannot be determined.

Three patients (0.9%) underwent surgery for adhesive/band-related ileus before stoma closure. This finding indicates that morbidity while a stoma remains in situ may extend to the need for repeat abdominal surgery. Adhesion-related obstruction occurring after closure represents a different outcome and was therefore not included in the primary complication assessment in our study.

Coletta et al.²¹ reported their experience with tube enterostomy in selected neonates across three centers.

Such studies provide a rationale for evaluating alternative techniques in specific clinical settings. However, small and highly selected series should not be regarded as evidence of superiority over conventional stoma techniques. The heterogeneous and retrospective nature of our cohort likewise precludes determination of an optimal technique.

The burden of living with a stoma is not limited to local and surgical complications. Chong et al.¹⁰ demonstrated adverse growth trajectories from stoma formation to closure in neonates with enterostomies after necrotizing enterocolitis. Bethell et al.⁷ also evaluated growth outcomes in infants with enterostomies. In the descriptive clinical observations included in our study, children with small-bowel stomas appeared to have a less favorable anthropometric course, and enteral nutritional support was provided to those with measurements below -2 standard deviations. The observation that patients evaluated after closure reached age-appropriate anthropometric values supports continued growth monitoring after restoration of intestinal continuity. However, without standardized measures, predefined assessment time points, and the number of patients evaluated at each time point, neither a stoma-type-specific difference in growth nor a causal effect of stoma closure on growth can be established. Post-closure growth observations should therefore be interpreted separately from pre-closure surgical complications.

Limitations

The single-center retrospective design of this study is susceptible to underdocumentation of minor complications, particularly those managed outside the surgical unit. Differences in follow-up duration, as well as follow-up ending at death or at the last clinical record before stoma closure, may have influenced complication ascertainment. The use of broad anatomical stoma categories and the binary classification of comorbidity may not fully capture the clinical heterogeneity of the cohort. Ileostomies, jejunostomies and duodenostomies were combined, and comorbidity was analyzed as a binary variable. Gestational age and birth weight were highly correlated, limiting separation of their independent effects. Anthropometric measurements were not obtained at standardized time points; therefore, changes in weight or weight-for-age z scores cannot be attributed directly to the stoma, and conclusions about stoma-specific growth trajectories should be made with caution.

CONCLUSION

Intestinal stoma formation in children should be regarded not merely as a temporary step toward definitive reconstruction but as a distinct period of clinically meaningful morbidity. Approximately one-third of patients experienced at least one complication while the stoma remained in situ; peristomal dermatitis predominated, and a substantial proportion of prolapses required operative management. The associations with emergency surgery and lower birth weight identify clinically vulnerable groups who may warrant closer surveillance. A structured, risk-adapted follow-up pathway integrating assessment of stoma function, skin integrity, growth, and nutritional status may help recognize complications earlier and optimize care throughout the



diversion period. Prospective studies using standardized complication definitions and longitudinal nutritional assessment are needed to clarify modifiable risk factors and determine whether targeted stoma-care strategies can reduce morbidity.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was approved by the Clinical Researches Ethics Committee of Ankara Bilkent City Hospital (Date: 15.06.2026, Decision No: E2-26-368).

Informed Consent

This retrospective study used pre-existing anonymized patient data. No additional intervention was performed, and there was no direct patient contact. The study was approved by the Ethics Committee, and the requirement for written informed consent was waived by the ethics committee.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Conceptualization: AUU, CİÖ; Design: AUU, CİÖ; Supervision: AUU, CİÖ; Materials: AUU, CİÖ; Data Collection and/or Processing: AUU, CİÖ; Analysis and/or Interpretation: AUU, CİÖ; Literature Review: AUU, CİÖ; Writing: AUU, CİÖ; Critical Review: AUU, CİÖ.

REFERENCES

- Steinau G, Ruhl KM, Hörnchen H, Schumpelick V. Enterostomy complications in infancy and childhood. *Langenbecks Arch Surg.* 2001; 386(5):346-349. doi:10.1007/s004230100243
- Çiğdem MK, Onen A, Duran H, Öztürk H, Otçu S. The mechanical complications of colostomy in infants and children: analysis of 473 cases of a single center. *Pediatr Surg Int.* 2006;22:671-676. doi:10.1007/s00383-006-1718-4
- Peña A, Migotto-Krieger M, Levitt MA. Colostomy in anorectal malformations: a procedure with serious but preventable complications. *J Pediatr Surg.* 2006;41(4):748-756. doi:10.1016/j.jpedsurg.2005.12.021
- Vogel I, Eeftinck Schattenkerk LD, Venema E, et al. Major stoma related morbidity in young children following stoma formation and closure: a retrospective cohort study. *J Pediatr Surg.* 2022;57:402-406. doi:10.1016/j.jpedsurg.2021.11.021
- Tan AYE, Deogracias JMC. Major stoma complications in pediatric patients in a tertiary hospital in a low-middle-income country: a retrospective cohort study. *Pediatr Surg Int.* 2024;40:208. doi:10.1007/s00383-024-05791-0
- Wolf L, Gfroerer S, Fiegel H, Rolle U. Complications of newborn enterostomies. *World J Clin Cases.* 2018;6(16):1101-1110. doi:10.12998/wjcc.v6.i16.1101
- Bethell G, Kenny S, Corbett H. Enterostomy-related complications and growth following reversal in infants. *Arch Dis Child Fetal Neonatal Ed.* 2017;102(3):F230-F234. doi:10.1136/archdischild-2016-311126
- Aguayo P, Fraser JD, Sharp S, St Peter SD, Ostlie DJ. Stomal complications in the newborn with necrotizing enterocolitis. *J Surg Res.* 2009;157(2):275-278. doi:10.1016/j.jss.2009.06.005
- Bælum JK, Rasmussen L, Qvist N, Ellebæk MB. Enterostomy complications in necrotizing enterocolitis (NEC) surgery, a retrospective chart review at Odense University Hospital. *BMC Pediatr.* 2019;19:110. doi:10.1186/s12887-019-1488-5
- Chong C, van Druten J, Briars G, et al. Neonates living with enterostomy following necrotizing enterocolitis are at high risk of becoming severely underweight. *Eur J Pediatr.* 2019;178(12):1875-1881. doi:10.1007/s00431-019-03440-6
- WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. Geneva: World Health Organization; 2006.
- Sarac F. Intestinal stoma types, indications and complications in pediatric patients: a three-year single-center experience with necrotizing enterocolitis as the leading indication. *Med Bull Haseki.* 2025;63(5):281-287. doi:10.4274/haseki.galenos.2025.80299
- Massenga A, Chibwae A, Nuri AA, et al. Indications for and complications of intestinal stomas in the children and adults at a tertiary care hospital in a resource-limited setting: a Tanzanian experience. *BMC Gastroenterol.* 2019;19:157. doi:10.1186/s12876-019-1070-5
- Alghamdi HA, Alqahtani MSM, Asiri HMM, et al. Investigating colostomy-related morbidity in children following stoma formation and closure in a tertiary hospital, Abha, Saudi Arabia: a retrospective cohort study 2024. *BMC Pediatr.* 2024;24:630. doi:10.1186/s12887-024-05089-z
- O'Connor A, Sawin RS. High morbidity of enterostomy and its closure in premature infants with necrotizing enterocolitis. *Arch Surg.* 1998; 133(8):875-880. doi:10.1001/archsurg.133.8.875
- Xu X, Zhou Y, Tan Z, et al. Risk factors for stoma and incision complications of enterostomy in children with very early-onset inflammatory bowel disease: a prospective cohort study. *Eur J Pediatr.* 2025;184:146. doi:10.1007/s00431-024-05952-2
- Zeng Z, Hu Y, Hou S, Yang R, Shi Z, Li X. Risk factors for complications after infantile enterostomy: development of a clinical prediction model. *Front Public Health.* 2025;13:1566789. doi:10.3389/fpubh.2025.1566789
- Klein T, Guarnizo-Diaz GS, Semaan A, Boemers TM, Vahdani RM. Colostomy complications in patients with anorectal malformations. *Langenbecks Arch Surg.* 2025;410:210. doi:10.1007/s00423-025-03795-5
- Hamdi MA, Mohamed WM, Al-Ani UFT. The colostomy complications in anorectal malformation: a retrospective study. *J Coloproctol.* 2021; 41(3):286-288. doi:10.1055/s-0041-1730259
- Johnson P. Intestinal stoma prolapse and surgical treatments of this condition in children: a systematic review and a retrospective study. *Surg Sci.* 2016;7:400-426. doi:10.4236/ss.2016.79057
- Coletta R, Zulli A, O'Shea K, Mussi E, Bianchi A, Morabito A. Minimizing enterostomy complication in neonates, lessons learnt from three European tertiary centres. *Children (Basel).* 2022;9(2):162. doi:10.3390/children9020162

Pediatric supernumerary testis: preservation, orchiopexy, or orchiectomy? A contemporary review of diagnosis and surgical decision-making

✉Nikolaos Baltogiannis^{*1}, ✉Fotini Fili¹, ✉Athanasios Tsangos¹, ✉Melina-Eleni Baltogianni²

¹Department of Pediatric Surgery, Children's Hospital of Athens Medical Group, Athens, Greece

²Department of Gynecology and Obstetrics, General Hospital "Agios Andreas", Patras, Greece

Received: 23/08/2026

Accepted: 29/09/2026

Published: 20/10/2026

Cite this article: Baltogiannis N, Fili F, Tsangos A, Baltogianni ME. Pediatric supernumerary testis: preservation, orchiopexy, or orchiectomy? A contemporary review of diagnosis and surgical decision-making. *Surg Child*. 2026;3(4):155-160. doi:10.51271/SOC-0089

*Corresponding Author: Nikolaos Baltogiannis, baltnick@gmail.com

ABSTRACT

Polyorchidism is a rare congenital anomaly in which more than two testes are present. In children, its importance lies less in its rarity than in the surgical dilemma created when an additional gonad is identified during examination, imaging, or surgery for cryptorchidism, hernia, torsion, or an inguinoscrotal mass. This review examines the contemporary evidence on pediatric supernumerary testes and proposes a practical framework for deciding when a viable gonad should be preserved, fixed, or removed. A focused narrative review was undertaken using key historical publications, systematic reviews, contemporary pediatric pooled analyses, and selected case reports relevant to embryology, classification, imaging, diagnostic pitfalls, torsion, malignancy, reproductive potential, and surgical management. The available evidence supports a preservation-oriented approach for a viable, morphologically reassuring supernumerary testis when surveillance is feasible. Orchiopexy is appropriate when a viable gonad is malpositioned or has a demonstrable fixation abnormality, whereas orchiectomy remains appropriate for irreversible ischemia, severe atrophy or involution, grossly abnormal tissue, or genuine oncological suspicion. The absolute malignant risk of a normal pediatric supernumerary testis remains uncertain and should not be inferred directly from historical case-report series. Management should therefore be individualized according to viability, morphology, position, vascular and ductal anatomy, associated cryptorchidism, torsion risk, oncological concern, and feasibility of follow-up. The presence of a supernumerary testis alone is not an indication for orchiectomy.

Keywords: Polyorchidism, supernumerary testis, cryptorchidism, orchiopexy, orchiectomy, testicular torsion, ultrasonography

INTRODUCTION

Polyorchidism is among the rarest congenital anomalies of the male gonad. It is defined by the presence of more than two testes, most commonly three, although more complex forms have occasionally been described. Despite its rarity, the anomaly presents a familiar surgical problem: whether tissue that appears testicular but is anatomically unusual should be preserved, repositioned, or removed.

The evidence available to answer this question is inherently imperfect. Most of the historical literature consists of individual case reports and small case series. Bergholz and Wenke¹ synthesized 140 histologically confirmed cases in their meta-analysis, while Balawender et al.² later incorporated contemporary imaging-based diagnoses and demonstrated the increasing role of conservative management. More recently, Fang et al.³ analyzed a two-center pediatric cohort and pooled 113 children with 127 supernumerary testes.

These studies improve our understanding of presentation and management, but they do not eliminate the limitations created by rarity, publication bias, incomplete follow-up, and changing diagnostic practice.

This limitation should shape the way the literature is interpreted. Percentages derived from published cases should not automatically be treated as population risks, and management preferences should not be mistaken for evidence-based mandates.

The additional testis may be identified as an asymptomatic scrotal structure, during evaluation of an inguinal mass, or unexpectedly during surgery for cryptorchidism, inguinal hernia, or testicular torsion. Published pediatric experience illustrates this spectrum. Baltogiannis et al.⁴ described triorchidism associated with cryptorchidism encountered



during orchiopexy. A later report by Baltogiannis et al.⁵ directly discussed conservative versus surgical management. In 2017, the same group reported an atypically positioned supernumerary testis presenting as an inguinal mass after previous surgery for cryptorchidism.⁶ These observations were separated by more than a decade, but all raised the same operative question: what is the value of the additional gonad, and what is gained—or lost—by removing it?

Contemporary management has moved away from routine orchiectomy. This shift is reasonable, but it should not be converted into a new dogma. A healthy, vascularized, morphologically normal testis is not equivalent to a severely atrophic intra-abdominal gonad, and neither is equivalent to a nonviable torsed testis or a lesion suspicious for neoplasia. For this reason, the management question is better framed as preserve, fix, or remove rather than simply conservative versus surgical treatment.

LITERATURE SEARCH AND SELECTION

For this focused narrative review, PubMed-indexed publications were searched through September 2025 using the terms “polyorchidism,” “supernumerary testis,” and “triorchidism,” combined with terms related to children, imaging, torsion, malignancy, orchiopexy, and orchiectomy. Reference lists of relevant reviews were also examined for pertinent older reports. Publications were selected for their relevance to pediatric presentation, anatomical classification, diagnostic imaging, complications, and surgical decision-making. Historical reports, major pooled reviews, the recent pediatric analysis, and illustrative case reports were included. No formal systematic screening, risk-of-bias assessment, or meta-analysis was performed for this narrative review.

EMBRYOLOGY AND ANATOMICAL DEVELOPMENT

The precise embryologic mechanism of polyorchidism remains uncertain. The most widely accepted explanation involves abnormal division of the genital ridge before completion of gonadal differentiation, although the timing and plane of division may vary. This can result in different relationships between the supernumerary testis, epididymis, and vas deferens.

The anatomical variation is clinically more important than the embryological theory itself. A supernumerary testis may have its own epididymis and vas deferens, may share one or both structures with the adjacent testis, or may lack a draining vas. Nocks⁷ described polyorchidism with equal-

sized testes and normal spermatogenesis and proposed an embryologic mechanism based on division of the genital ridge. That report remains relevant because it demonstrates that an additional testis cannot automatically be considered functionless.

Nevertheless, the presence of apparently normal testicular tissue does not establish an independent fertility contribution. The true reproductive value of most supernumerary testes remains uncertain, particularly in children, where long-term functional data are sparse. Anatomy can suggest potential function, but it should not be made to prove more than it actually does.

CLASSIFICATION: WHAT IS USEFUL AT THE OPERATING TABLE?

Singer, Donaldson, and Jackson⁸ developed a functional classification based on anatomical organization and the relationship of the testes to their drainage pathways. Bergholz et al.⁹ later emphasized whether the additional gonad was drained by a vas deferens and proposed an anatomical classification that remains clinically useful.

A viable gonad connected to an epididymis and vas deferens has a plausible reproductive role. An undrained gonad is less likely to contribute to fertility. However, classification should not dictate treatment mechanically. A gonad with apparently favorable ductal anatomy may still be severely atrophic or ischemic, while an undrained gonad may still contain viable endocrine tissue. Shared vascular or ductal anatomy may also make attempts at separation more dangerous than leaving both testes intact.

For the pediatric surgeon, three anatomical questions may therefore be more useful than the label itself: Is the supernumerary testis viable and well vascularized? Does it have an epididymal or vasal connection suggesting functional anatomy? Can it be left or placed in a position in which it can be safely examined and followed? The principal classification systems and their practical surgical relevance are summarized in Table.

CLINICAL PRESENTATION AND ASSOCIATED ANOMALIES

Most children do not present because polyorchidism is already known. A painless scrotal structure is common, but contemporary pediatric data show substantial variation. Fang et al.³ documented asymptomatic presentations as well as cryptorchidism, inguinal hernia, and torsion. Balawender et

Table. Anatomical classifications of polyorchidism and practical surgical relevance			
System	Basis	Key categories	Practical surgical relevance
Singer et al.	Functional/anatomical relationship	Anatomical configuration and drainage pattern linked to potential reproductive function.	Introduced the principle that management should reflect anatomy and potential function rather than numerical redundancy.
Bergholz et al.	Drainage by vas deferens and relation of epididymal/vasal structures	Type A: draining vas deferens; type B: no draining vas deferens; type A subgroups describe shared or separate epididymal/vasal anatomy.	Useful for describing ductal anatomy and potential reproductive relevance; should not be used as an automatic treatment rule.
Operative assessment	Direct anatomy and tissue quality	Viability, morphology, vascular pedicle, epididymis, vas deferens, position, and mobility.	Determines whether preservation, orchidopexy, or excision is technically safe in the individual child.



al.² found 76% of supernumerary testes to be scrotal and 24% extrascrotal; among extrascrotal testes, most were inguinal.

The left side predominates in the major pooled series.^{1,3} The reason for this distribution remains uncertain and should not be overinterpreted. Associated abnormalities matter mainly because they often determine how and when the diagnosis is made.

The 2005 report by Baltogiannis et al.⁴ illustrates intraoperative discovery during orchiopexy for cryptorchidism. Their 2017 report illustrates a different problem: after previous orchiopexy, an inguinal mass was ultimately found to represent additional testicular tissue.⁶ The lesson is not that every inguinal mass in a child with cryptorchidism represents polyorchidism, but that abnormal gonadal descent should lower the threshold for considering atypical testicular anatomy.

DIAGNOSTIC PITFALLS AND IMAGING

Ultrasonography with color Doppler is the first-line imaging modality. In an uncomplicated intrascrotal case, a homogeneous oval structure with echogenicity and vascularity similar to the normal testis may be sufficiently characteristic to avoid exploratory surgery. Chung and Yao¹⁰ demonstrated this classic pattern and the value of high-resolution sonography.

Diagnostic confidence falls when the supernumerary testis is small, extrascrotal, distant from the orthotopic testis, or associated with other urogenital abnormalities. Fang et al.³ reported an ultrasonographic misdiagnosis rate of 22.22% in their pooled pediatric analysis, with misdiagnosis associated with younger age, extrascrotal location, greater distance from the orthotopic testis, and concomitant anomalies. The atypical inguinal presentation described by Baltogiannis et al.⁶ reflects the same practical difficulty.

Magnetic resonance imaging may add value when ultrasonography is equivocal, when the gonad is extrascrotal, or when anatomical relationships or a focal lesion are not adequately characterized. Yalçinkaya et al.¹¹ showed the complementary role of sonography and MRI, while Danrad et al.¹² suggested that imaging can help define epididymal and ductal anatomy relevant to potential reproductive function. MRI should not, however, be performed simply because polyorchidism is rare. In a child with an unequivocally normal intrascrotal supernumerary testis, additional imaging is difficult to justify if it will not change management.

TORSION: A REAL RISK, BUT NOT A UNIVERSAL INDICATION FOR FIXATION

Torsion is an important acute complication reported in association with polyorchidism. Ferro and Iacobelli¹³ described two children with duplicated scrotal testes and bell-clapper deformity and performed orchidopexy of testes judged to be at risk. Their report provides direct operative evidence that abnormal fixation may coexist with gonadal duplication.

That observation should not be extended beyond what it proves. Published torsion cases are subject to reporting bias, and no study has established the absolute incidence of torsion in an unselected pediatric population with supernumerary testes. Prophylactic exploration and fixation of every incidentally diagnosed intrascrotal supernumerary testis therefore cannot be supported.

Fixation is more defensible when there is a specific anatomical problem: a viable cryptorchid testis requiring orchidopexy, a gonad that has detorsed and recovered perfusion, a bell-clapper deformity, or marked abnormal mobility encountered at surgery.

MALIGNANCY: WHAT DO WE REALLY KNOW?

Concern regarding malignancy has historically been a major argument for orchiectomy. Bergholz and Wenke¹ identified neoplasms among published histologically confirmed cases and noted that malignant tumors with known location occurred predominantly in cryptorchid supernumerary testes. This association deserves attention, but the denominator deserves equal attention.

Historical percentages derived from collections of published cases are not prospective lifetime risks. Rare tumors and unusual complications are more likely to be reported than uncomplicated children who remain well. The resulting publication bias makes it unsafe to present a single historical percentage to families as the expected cancer risk of a normal intrascrotal supernumerary testis.

Contemporary pediatric evidence supports caution against routine prophylactic removal. In the pooled analysis by Fang et al.,³ none of 53 postoperative pathology specimens was malignant. This does not prove lifelong oncological safety: the patients were children, follow-up was limited, and specimens came from a selected surgical population. The more defensible conclusion is that malignancy is possible, cryptorchid or abnormal location increases concern, but the absolute risk in an otherwise normal pediatric supernumerary testis remains uncertain.

REPRODUCTIVE AND ENDOCRINE POTENTIAL

Some supernumerary testes have normal size, echotexture, vascularity, epididymal anatomy, and vasal drainage. Normal spermatogenesis has been documented in selected cases.⁷ Danrad et al.¹² further suggested that imaging-defined epididymal and vasal anatomy may help infer potential reproductive relevance.

These observations make routine removal of normal tissue difficult to justify. They do not establish that retaining an additional testis improves future fertility, and there are no meaningful comparative pediatric studies demonstrating such an outcome. The same limitation applies to endocrine function. The argument for preservation should therefore remain modest and defensible: viable normal gonadal tissue should not be sacrificed without a valid clinical reason.



SURGICAL DECISION-MAKING: PRESERVE, FIX, OR REMOVE?

Earlier reports by Baltogiannis and colleagues⁴⁻⁶ span the evolution of this debate: triorchidism associated with cryptorchidism in 2005, direct discussion of conservative versus surgical management in 2007, and an atypical extrascrotal presentation illustrating diagnostic uncertainty in 2017. Viewed together, these reports reinforce a point that remains relevant: the decision should be made for the individual gonad, not for the diagnosis of polyorchidism.

Preservation

Preservation is favored when the supernumerary testis is viable, well perfused, morphologically reassuring, free of a suspicious focal lesion, scrotal or capable of safe scrotal placement, and suitable for reliable surveillance. A recognizable epididymal or vasal connection further supports preservation because it suggests potential reproductive anatomy.^{8,9,12}

Observation is particularly reasonable for an uncomplicated intrascrotal testis diagnosed confidently on imaging. Khedis et al.¹⁴ advocated a selective management strategy rather than routine excision. The systematic review by Balawender et al.² similarly documented observation in 46% of patients.

Orchiopexy

A viable testis should be fixed when there is a specific correctable anatomical problem. Examples include cryptorchid or ectopic position when tension-free orchidopexy is feasible, torsion in which perfusion is restored, bell-clapper deformity, or clearly abnormal mobility. Fixation should preserve rather than compromise vascular supply.

Orchiectomy

Orchiectomy remains appropriate for irreversible ischemia or necrosis, severe atrophy or involution, grossly dysplastic testicular tissue, a suspicious intratesticular lesion, or an abnormal extrascrotal gonad that cannot be safely positioned or meaningfully followed. Extrascrotal location alone should not automatically mean removal. A viable inguinal testis that can be safely brought into the scrotum is different from a severely atrophic intra-abdominal gonad.

SHARED VASCULAR AND DUCTAL ANATOMY: PROTECT THE BETTER TESTIS

Supernumerary and orthotopic testes may share an epididymis, vas deferens, vascular pedicle, or combinations of these structures. Bergholz et al.⁹ reported a child in whom a supernumerary testis had its own epididymis and deferent duct but shared the blood supply with the ipsilateral distal testis. Such anatomy makes the risk of injuring the better-developed gonad a central operative consideration.

From a surgical standpoint, removal of an additional testis is not justified if the dissection required to remove it creates avoidable vascular or ductal risk to a normal dominant testis. This is not a formal guideline derived from comparative trials; it is a surgical principle. When separation appears

hazardous and neither gonad is suspicious, preservation may be safer than anatomical perfection.

POLYORCHIDISM ASSOCIATED WITH CRYPTORCHIDISM

Cryptorchidism changes the risk-benefit calculation because it complicates surveillance and is independently associated with impaired germ-cell development and testicular malignancy. Mathur et al.¹⁵ and other pediatric reports have documented the association between polyorchidism and cryptorchidism. The report by Baltogiannis et al.⁴ illustrates the classic situation in which triorchidism is discovered during orchiopexy.

At exploration, each gonad should be assessed independently for size, consistency, perfusion, epididymal anatomy, vas deferens, vascular pedicle, mobility, and the feasibility of tension-free scrotal placement. The presence of two ipsilateral gonads does not mean that one must be removed.

A viable testis that can be safely positioned in the scrotum may reasonably be preserved and fixed. A severely atrophic or nonviable gonad may reasonably be excised. If either preservation or removal threatens the vascular or ductal anatomy of the better-developed gonad, protection of the dominant testis takes priority.

FOLLOW-UP AFTER CONSERVATIVE MANAGEMENT

The literature frequently recommends clinical and ultrasonographic follow-up, but no validated surveillance schedule exists for children with polyorchidism. Published intervals reflect local practice rather than evidence that one schedule improves outcomes.

For a normal intrascrotal supernumerary testis, follow-up can reasonably include periodic clinical examination and ultrasound to document growth, morphology, vascularity, and absence of a new focal lesion. Long-term conservative follow-up has been described in selected patients with reassuring imaging.^{16,17} As the child matures, education regarding a new mass, persistent pain, or change in testicular consistency is reasonable. Routine serial tumor markers in an asymptomatic child with normal imaging are not supported by evidence, and repeated MRI in an uncomplicated scrotal case is difficult to justify.

PROPOSED PRESERVE-FIX-REMOVE FRAMEWORK

The proposed algorithm begins with confirmation that the structure represents testicular tissue and assessment of viability. A viable, reassuring, scrotal supernumerary testis is preserved and observed. A viable testis that is cryptorchid, ectopic, or demonstrably at risk because of abnormal fixation is preserved and fixed when this can be achieved safely. Irreversibly ischemic, severely atrophic, involuted, or genuinely suspicious tissue may be removed.

Shared vascular or ductal anatomy is a separate decision point. If removal of the supernumerary gonad risks the dominant ipsilateral testis, protection of the better testis takes priority. **Figure** presents this sequence as the proposed Preserve-Fix-Remove framework.

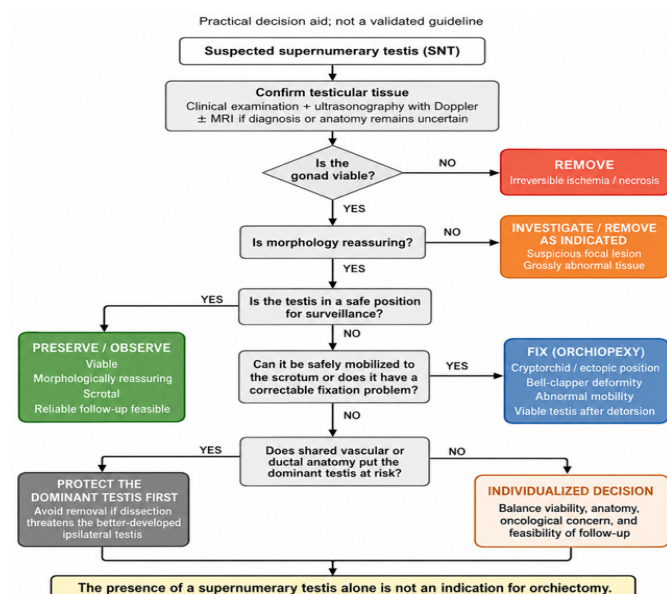


Figure. Proposed Preserve-Fix-Remove framework for pediatric supernumerary testes. The algorithm is intended as a practical synthesis of the available evidence and surgical principles rather than a validated clinical guideline

MRI: Magnetic resonance imaging

LIMITATIONS OF THE AVAILABLE EVIDENCE

There are no randomized studies, prospective comparative management trials, or large population-based cohorts of children with polyorchidism. Most evidence arises from case reports and retrospective pooled analyses collected over decades during which imaging, pathology, orchidopexy techniques, and attitudes toward gonadal preservation have changed. Exceptionally unusual anatomical variants, including reports of more than three testes, further illustrate the heterogeneity of the literature.¹⁸

Publication bias is unavoidable: malignancy, torsion, unusual anatomy, and operative surprises are more likely to be reported than uncomplicated long-term observation. Long-term fertility and endocrine outcomes are poorly documented. The proposed Preserve-Fix-Remove framework should therefore be interpreted as a practical decision aid, not as a validated guideline.

CONCLUSION

Polyorchidism in childhood is rare, but the management problem it creates is familiar to the pediatric surgeon: whether anatomical abnormality justifies removal of potentially useful tissue. The available evidence does not support orchiectomy merely because a supernumerary testis is present. Management should be individualized according to viability, morphology, position, vascular and ductal anatomy, associated cryptorchidism, torsion risk, oncological suspicion, and feasibility of surveillance.

A normal intrascrotal supernumerary testis can generally be preserved and followed. A viable malpositioned or abnormally mobile testis may benefit from orchidopexy. A nonviable, severely atrophic, grossly abnormal, or genuinely suspicious gonad may appropriately be removed. Above all, surgery should not create more harm than the anomaly itself.

ETHICAL DECLARATIONS

Peer Review Process

This review was externally peer-reviewed.

Conflict of Interest

The authors declare no conflicts of interest.

Financial Disclosure

No financial support was received for the preparation or publication of this article.

Author Contributions

Conceptualization: NB; Design: NB, FF; Supervision: NB; Literature Assessment: NB, FF; Analysis and/or Interpretation: NB, FF, AT, MEB; Literature Review: NB, FF, AT, MEB; Writing: NB, FF, AT, MEB; Critical Review: NB, FF, AT, MEB; Final Approval: All Authors; Accountability: All Authors.

REFERENCES

- Bergholz R, Wenke K. Polyorchidism: a meta-analysis. *J Urol*. 2009; 182(5):2422-2427. doi:10.1016/j.juro.2009.07.063
- Balawender K, Wawrzyniak A, Kobos J, et al. Polyorchidism: an up-to-date systematic review. *J Clin Med*. 2023;12(2):649. doi:10.3390/jcm12020649
- Fang H, Wang X, Yu Y, et al. Beyond triorchidism: diagnostic pitfalls and an evidence-driven algorithm for pediatric supernumerary testes. *Eur J Med Res*. 2026;31:333. doi:10.1186/s40001-026-03938-4
- Baltogiannis N, Mavridis G, Keramidis D. Triorchidism associated with cryptorchidism. *Pediatr Med Chir*. 2005;27(1-2):104-105.
- Baltogiannis N, Papandreou E, Faviou E, Christopoulos-Geroulanos G. Triorchidism: conservative or surgical management? *Review Series Urology*. 2007;(3-4):17-19.
- Baltogiannis N, Fili F, Faviou E. Polyorchidism. Atypical presentations of supernumerary testicle can lead to misdiagnosis. *BAOJ Pediat*. 2017; 3:043.
- Nocks BN. Polyorchidism with normal spermatogenesis and equal sized testes: a theory of embryological development. *J Urol*. 1978;120(5):638-640. doi:10.1016/S0022-5347(17)57308-5
- Singer BR, Donaldson JG, Jackson DS. Polyorchidism: functional classification and management strategy. *Urology*. 1992;39(4):384-388. doi:10.1016/0090-4295(92)90220-Q
- Bergholz R, Koch B, Spieker T, Lohse K. Polyorchidism: a case report and classification. *J Pediatr Surg*. 2007;42(11):1933-1935. doi:10.1016/j.jpedsurg.2007.07.031
- Chung TJ, Yao WJ. Sonographic features of polyorchidism. *J Clin Ultrasound*. 2002;30(2):106-108. doi:10.1002/jcu.10031
- Yalçinkaya S, Şahin C, Şahin AF. Polyorchidism: sonographic and magnetic resonance imaging findings. *Can Urol Assoc J*. 2011;5(5):E84-86. doi:10.5489/cuaj.10077
- Danrad R, Ashker L, Smith W. Polyorchidism: imaging may denote reproductive potential of accessory testicle. *Pediatr Radiol*. 2004;34(6): 492-494. doi:10.1007/s00247-003-1120-5
- Ferro F, Iacobelli B. Polyorchidism and torsion. A lesson from 2 cases. *J Pediatr Surg*. 2005;40(10):1662-1664. doi:10.1016/j.jpedsurg.2005.06.029
- Khedis M, Nohra J, Dierickx L, et al. Polyorchidism: presentation of 2 cases, review of the literature and a new management strategy. *Urol Int*. 2008;80(1):98-101. doi:10.1159/000111738



15. Mathur P, Prabhu K, Khamesra HL. Polyorchidism revisited. *Pediatr Surg Int.* 2002;18(5-6):449-450. doi:10.1007/s00383-002-0765-8
16. Artul S, Habib G. Polyorchidism: two case reports and a review of the literature. *J Med Case Rep.* 2014;8:464. doi:10.1186/1752-1947-8-464
17. Savas M, Yeni E, Ciftci H, Cece H, Topal U, Utangac MM. Polyorchidism: a three-case report and review of the literature. *Andrologia.* 2010;42(1):57-61. doi:10.1111/j.1439-0272.2009.00957.x
18. Zahirian Moghadam T, Mohseni Rad H, Zandian H, Hosseinkhani A. Five testicles in the genital area of a thirteen-month-old baby: a case report. *BMC Urol.* 2020;20:128. doi:10.1186/s12894-020-00681-8