Research Square

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

Single-stage versus staged reconstruction for high and intermediate anorectal malformations in neonates: a systematic review and meta-analysis

Tuyen Tran
tuyen.tran97@gmail.com
Vinmec Central Park, Vinmec Healthcare System
Chuong Ngoc Thi Dang
Vinmec Central Park, Vinmec Healthcare System
Chieu Minh Vuong
Children's Hospital No.2

Systematic Review

Keywords: Anorectal malformations, Posterior sagittal anorectoplasty, Fecal incontinence, Neonatal surgery, Systematic review, Meta-analysis

Posted Date: May 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9207685/v1

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

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Pediatric Surgery International on July 15th, 2026. See the published version at https://doi.org/10.1007/s00383-026-06545-w.

Abstract

Purpose

The surgical approach to high and intermediate anorectal malformations remains controversial. This systematic review compared continence and infection outcomes between the two approaches.

Methods

MEDLINE, Embase, and Cochrane CENTRAL were searched per a registered protocol (PROSPERO CRD420261291319). Comparative studies of single-stage versus staged repair in neonates with high or intermediate anorectal malformations were included. Risk of bias was assessed using ROBINS-I. Random-effects meta-analysis with prediction intervals was performed.

Results

Four comparative studies (746 patients: 399 single-stage, 347 staged) were included. Pooled analysis showed no significant difference in incontinence outcomes (RR 0.93, 95% CI: 0.52–1.66, p = 0.80; 95% prediction interval: 0.28–3.08) with substantial heterogeneity (I^2=83%). Infection data could not be pooled due to methodological concerns in outcome reporting. All included studies demonstrated moderate to serious risk of bias. GRADE certainty was LOW for incontinence and VERY LOW for infection.

Conclusion

Current evidence does not demonstrate superiority of either approach for bowel continence (LOW certainty). Infection evidence was insufficient for synthesis (VERY LOW). Wide prediction intervals indicate substantial uncertainty; available data cannot confirm or refute inferior outcomes after single-stage repair. Prospective studies with standardized outcomes are needed.

INTRODUCTION

Anorectal malformations (ARM) affect approximately 1 in 5,000 live births and represent a spectrum of congenital anomalies requiring complex surgical reconstruction [1]. The classification of ARM into low, intermediate, and high types, while considered a backward classification, still reflects both anatomic features and surgical management strategies [2]. While low-type malformations such as rectovestibular and perineal fistulas are routinely managed with primary single-stage repair in the neonatal period, the optimal surgical approach for high and intermediate lesions (including rectoprostatic, rectovesical, and rectobulbar fistulas) remains subject to ongoing debate [3–5].

Traditional management of high and intermediate ARM has followed a three-stage approach: initial colostomy creation for fecal diversion, delayed definitive repair (typically posterior sagittal anorectoplasty, PSARP) at 2–3 months of age, and subsequent colostomy closure several months later [1, 5]. This staged approach was developed to allow pelvic structures to grow, facilitate anatomic identification during repair, and reduce the risk of overwhelming infection in neonates [1]. However, this strategy requires multiple operations, prolonged hospital stays, stoma-related morbidity, and substantial caregiver burden, with particular implications for families in resource-limited settings [6–8].

Over the past two decades, several centers have reported experience with primary single-stage neonatal repair for high and intermediate ARM, performed without preceding colostomy [9–12]. These studies argued about reduced treatment burden, elimination of stoma-related complications, improved family quality of life, and cost savings as key advantages. Technical advances including MRI and laparoscopy-assisted anorectal pull-through (LAARP) have further expanded options for minimally invasive single-stage reconstruction [4, 10, 11]. Yet concerns persist that operating on neonatal anatomy may result in inferior long-term functional outcomes, particularly regarding fecal incontinence.

Previous systematic reviews pooled heterogeneous populations including low-type ARM alongside high and intermediate types [3, 4], obscuring the evidence for the population where clinical equipoise genuinely exists. Prior meta-analyses have not adequately addressed substantial heterogeneity in outcome definitions, limiting the validity of pooled estimates.

The objective of this systematic review and meta-analysis was to compare bowel incontinence and early postoperative infection outcomes between primary single-stage definitive repair and traditional staged reconstruction in neonates with high and intermediate anorectal malformations exclusively. By focusing strictly on intermediate or high lesions and employing systematic methodology including PROSPERO

registration and GRADE evidence assessment, we aimed to provide clinically relevant synthesis that directly addresses the population where surgical decision-making remains genuinely uncertain.

METHODS

Protocol and Registration

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [13]. The protocol was registered a priori with PROSPERO (registration number CRD420261291319) on 22 January 2026.

Eligibility Criteria

Studies were eligible for inclusion if they met the following criteria:

Population: Neonates and infants undergoing definitive repair for high or intermediate (non-low, above-levator) anorectal malformations, including: (1) high-type lesions: rectoprostatic fistula, rectovesical fistula, or ARM without fistula with rectal pouch located above the levator ani; and (2) intermediate-type lesions: rectobulbar (rectourethral bulbar) fistula or rectovaginal fistula when explicitly classified as intermediate or high. Studies including low-type lesions (rectovestibular fistula, perineal fistula, cutaneous fistula, rectal stenosis) were excluded unless data were stratified by ARM type and high/intermediate outcomes could be extracted separately.

Intervention

Primary single-stage repair performed in the neonatal period without preceding colostomy, including posterior sagittal anorectoplasty (PSARP), laparoscopy-assisted anorectal pull-through (LAARP), or other definitive reconstructive procedures.

Comparator

Staged reconstruction involving initial colostomy creation followed by delayed definitive repair and subsequent colostomy closure.

Outcomes

Primary outcomes were (1) long-term bowel continence assessed using validated scoring systems or dichotomized functional outcomes, and (2) early postoperative infection including wound infection, pelvic infection, or surgical site infection. Secondary outcomes included constipation, mortality, reoperation rate, length of hospital stay, quality of life, and cost.

Study design

Comparative cohort studies (prospective or retrospective) and randomized controlled trials comparing single-stage versus staged reconstruction were eligible. Case reports, narrative reviews, and studies without comparative data were excluded.

No restrictions were applied based on publication date, language, or geographic location.

Information Sources and Search Strategy

We searched MEDLINE (via PubMed) and Cochrane Central Register of Controlled Trials (CENTRAL) from inception through 30 January 2026, and Embase (via Ovid) from inception through 12 March 2026. The search strategy combined terms related to anorectal malformations (including "anorectal malformation*", "imperforate anus", "anal atresia"), neonatal/pediatric surgery ("infant, newborn", "neonat*", "pediatric surgery"), and surgical approaches ("single-stage", "one-stage", "1-stage", "primary repair", "staged", "colostomy"). Medical Subject Headings (MeSH), Emtree terms, and free-text terms were combined using Boolean operators. Population-level exclusion terms (rectovestibular fistula, perineal fistula, cloacal malformation) were applied at the PubMed search stage to improve specificity; all retrieved records were subsequently screened for eligibility at the title/abstract and full-text stages. We additionally searched ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform (ICTRP) for relevant registered trials; no eligible registered trials were identified. The complete search strategy for each database is provided in Online Resource 1. Reference lists of included studies and relevant systematic reviews were manually screened for additional eligible studies.

Study Selection

Titles and abstracts were independently screened by two reviewers (TT and CMV) according to the eligibility criteria. Inter-rater agreement for screening was high (99.6%, 5/1,247 discordant decisions). Discrepancies were resolved through discussion until consensus was reached.

Full-text articles were independently assessed by both reviewers using the same process. Agreement on full-text inclusion decisions was 100% (14/14 concordant decisions). When uncertainties arose regarding eligibility criteria interpretation, these were resolved through discussion with reference to the registered protocol.

Data Extraction

Data were extracted independently by two reviewers (TT and CMV) using a structured, predefined data extraction form. Extracted variables included study characteristics, population details, surgical approaches, outcome definitions, follow-up duration, and risk-of-bias domains. Discrepancies in extracted data were identified and resolved through direct comparison and discussion. For studies reporting multiple follow-up time points, the longest available follow-up was used for functional outcomes.

Risk of Bias Assessment

Risk of bias was assessed independently by two reviewers (TT and CMV) using the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool [14]. Disagreements were resolved through discussion until consensus was achieved. As anticipated, all included observational studies were judged to have at least moderate risk of bias due to confounding by indication and non-randomized design.

Data Synthesis and Statistical Analysis

For dichotomous outcomes, we calculated risk ratios (RR) with 95% confidence intervals (CI) for individual studies using 2×2 contingency tables. We performed random-effects meta-analysis using the DerSimonian-Laird method given anticipated clinical and methodological heterogeneity across studies [15]. Given the small number of included studies (n = 4), pooled estimates were interpreted cautiously and supported by prediction intervals to characterize uncertainty rather than reliance on point estimates or p-values alone. Statistical heterogeneity was assessed using the I^2 statistic and chi-square test, with I^2 values of 25%, 50%, and 75% representing low, moderate, and substantial heterogeneity, respectively [16].

Although heterogeneity was substantial (I^2=83%), quantitative synthesis was performed to estimate the direction and range of effect and to quantify the degree of uncertainty in a rare condition where individual studies are underpowered. We did not interpret the pooled estimate as definitive but as a summary measure reflecting the current fragmented evidence base. To further characterize uncertainty, we calculated 95% prediction intervals, which estimate the expected range of the true effect in a future comparable study [16].

For outcomes where substantial heterogeneity (I^2>75%) was present or where methodological concerns precluded reliable pooling, we conducted narrative synthesis describing individual study findings and limitations. Given the small number of studies (n = 4), subgroup analyses would be statistically underpowered and potentially misleading; therefore, we restricted exploration of heterogeneity to qualitative assessment. We planned sensitivity analyses excluding studies at serious or critical risk of bias and those with specific methodological concerns (e.g., studies restricted to male patients).

All statistical analyses were performed using R with standard meta-analysis calculations. We considered $p < 0.05$ as statistically significant for pooled estimates and $p < 0.10$ for heterogeneity tests, given the limited statistical power of the Q-test with few studies. Given the limited number of included studies, publication bias was not formally assessed [16].

Certainty of Evidence Assessment

We assessed the certainty of evidence for primary outcomes using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework [17]. Evidence from observational studies was initially rated as LOW certainty and downgraded as appropriate for risk of bias, inconsistency, indirectness, and imprecision.

RESULTS

Study Selection

The systematic search identified 1,764 records: 585 from MEDLINE, 1,162 from Embase, and 17 from Cochrane CENTRAL. After removal of 517 duplicates, 1,247 unique records underwent title and abstract screening, of which 1,233 were excluded. Fourteen studies underwent full-

text review, of which ten were excluded: three were single-arm studies without a comparator group, two were systematic reviews or meta-analyses, two used an incorrect comparator (technique rather than staging), one reported an overlapping cohort from the same institution, one included a mixed ARM population without subtype stratification, and one was a staged-only series. Four studies met all eligibility criteria and were included in the systematic review (Fig. 1) [9, 11, 12, 24].

Study Characteristics

The four included studies collectively enrolled a total of 1,341 patients across all reported outcomes; 746 patients contributed data to the primary continence analysis (399 single-stage, 347 staged). Three studies employed open posterior sagittal anorectoplasty (PSARP) [9, 12, 24], while one utilized laparoscopy-assisted anorectal pull-through (LAARP) [11]. All studies reported neonatal repair for high and intermediate anorectal malformations. Study characteristics are summarized in Table 1.

Table 1
Characteristics of studies included in the systematic review

Study (year)	Country	Study design	Study period	Population	ARM classification	Surgical approach	Sample size (SS / ST)	Follow-up (years)	Outcomes reported
Liu et al. (2004)	China	Non-randomized comparative cohort (2 centers)	NR	Neonates with high & intermediate ARM	Wingspread	Open PSARP	65 / 48	Median 6.7	Incontinence, infection, constipation
Gangopadhyay et al. (2006)	India	Historical comparative cohort (single center)	1989–2003	Neonates with high & intermediate ARM	NR	Open PSARP	254 / 200*	≥ 3	Incontinence, infection
Menon et al. (2017)	India	Retrospective comparative cohort (single center)	2006–2016	Male neonates with high & intermediate ARM	NR	Open PSARP	47 / 61	≥ 3	Incontinence, mortality
Zhou et al. (2022)	China	Retrospective propensity score–matched cohort (single center)	2013–2018	Male neonates with high & intermediate ARM	Krickenbeck	LAARP	33 / 38**	Median 4.5	Incontinence, infection, constipation

Abbreviations:

ARM, anorectal malformation; SS, single-stage repair; ST, staged repair; PSARP, posterior sagittal anorectoplasty; LAARP, laparoscopy-assisted anorectal pull-through; NR, not reported.

* Larger sample sizes were reported for infection outcomes in this study, suggesting cumulative counting across procedures.

** Larger matched cohorts were reported for infection outcomes (42 / 42).

Liu et al. (2004) conducted a non-randomized comparative cohort study across two centers in China, enrolling 113 patients (65 single-stage, 48 staged) with high and intermediate ARM classified using Wingspread criteria [24]. Continence was assessed using a modified Wingspread score (excellent, good, fair, poor) at median 6.7-year follow-up, with excellent and good outcomes combined for dichotomous analysis. The study reported moderate risk of bias due to non-randomized allocation, though loss to follow-up was acceptable at 23.6% (35/148 patients).

Gangopadhyay et al. (2006) reported a historical comparative cohort from a single center in India spanning 1989–2003 [9]. The study included 454 patients for continence analysis (254 single-stage, 200 staged), with a larger cohort (490 single-stage, 458 staged) reported for infection outcomes. All patients underwent open PSARP for high and intermediate ARM. Continence was evaluated using Kelly score at minimum 3-year follow-up, with "good" outcomes compared against "fair" or "poor" results. Serious risk of bias was identified due to historical control design and substantial attrition bias, with 56.3% loss to follow-up in the staged group (258/458 patients). The discrepancy in sample sizes between outcome analyses raised concerns about denominator definitions and potential selective outcome reporting.

Menon et al. (2017) conducted a retrospective comparative cohort at a single center in India from 2006–2016, enrolling 108 male patients (47 single-stage, 61 staged) with high and intermediate ARM managed with open PSARP [12]. Continence assessment used Kelly score at

minimum 3-year follow-up. The study showed a serious risk of bias due to restriction to male patients (limiting generalizability), high attrition (24.8% loss to follow-up, 40/161 patients), with an additional 13 deaths (8.1%), yielding 108 patients with evaluable continence outcomes, and retrospective design with potential unmeasured confounding.

Zhou et al. (2022) reported a propensity-score matched cohort from a single center in China (2013–2018), including 71 male patients with high and intermediate ARM classified using Krickenbeck criteria [11]. Single-stage patients (n = 33 for continence analysis, n = 42 for infection analysis) underwent neonatal one-stage LAARP, while staged patients (n = 38 for continence, n = 42 for infection) received initial colostomy followed by delayed LAARP and subsequent closure. Bowel function was assessed using the Bowel Function Score at median 4.5-year follow-up, with "normal" and "good" outcomes combined for dichotomous analysis. Moderate risk of bias was identified attributable to retrospective design, though propensity-score matching mitigated confounding from measured covariates. Loss to follow-up was 15.5% (13/84 patients). This was the most contemporary evidence using modern laparoscopic techniques.

Risk of Bias Assessment

Overall risk of bias was moderate in two studies and serious in two studies (Fig. 2). Major concerns included non-randomized treatment allocation, confounding by indication, attrition bias, and heterogeneity in outcome definitions. Blinding of outcome assessment was not feasible due to the surgical nature of interventions and long-term follow-up requirements.

Bowel Incontinence (Long-term Outcome)

All four studies reported long-term bowel incontinence outcomes, although definitions varied across studies (Table 2). Continence assessment instruments and follow-up durations are detailed in Table 2.

Table 2
Long-term bowel incontinence outcomes by study

Study (year)	Outcome definition	Single-stage repair	Staged repair	Effect estimate (RR, 95% CI)	Direction of effect
Liu et al. (2004)	Modified Wingspread (poor/fair vs good/excellent)	30 / 65 (46.2%)	20 / 48 (41.7%)	1.11 (0.72–1.70)	No significant difference
Gangopadhyay et al. (2006)	Kelly score (fair/poor vs good)	81 / 254 (31.9%)	110 / 200 (55.0%)	0.58 (0.47–0.72)	Lower incontinence with single-stage
Menon et al. (2017)	Kelly score (fair/poor vs good)	21 / 47 (44.7%)	16 / 61 (26.2%)	1.70 (1.01–2.89)	Higher incontinence with single-stage
Zhou et al. (2022)	Bowel Function Score (poor vs normal/good)	3 / 33 (9.1%)	6 / 38 (15.8%)	0.58 (0.16–2.12)	No significant difference
Abbreviations:					
RR, risk ratio (single-stage vs staged); CI, confidence interval.					
Higher RR indicates higher incontinence following single-stage repair.					

Individual study results were inconsistent in both magnitude and direction of effect. Liu et al. reported similar incontinence rates between single-stage and staged repair (RR 1.11, 95% CI: 0.72–1.70). Gangopadhyay et al. reported lower incontinence following single-stage repair compared with staged reconstruction (RR 0.58, 95% CI: 0.47–0.72), whereas Menon et al. found higher incontinence following single-stage repair (RR 1.70, 95% CI: 1.01–2.89). Zhou et al. also reported lower incontinence rates following single-stage repair, although the difference was not statistically significant (RR 0.58, 95% CI: 0.16–2.12).

Pooled analysis of all four studies (746 patients) showed no significant difference in long-term bowel incontinence between single-stage and staged repair (RR 0.93, 95% CI: 0.52–1.66; p = 0.80), with substantial heterogeneity (I^2=83.3%) and a wide 95% prediction interval (0.28–3.08) (Fig. 3). The prediction interval indicates that in a future comparable study, the true effect could range from substantial benefit (RR = 0.28) to substantial harm (RR = 3.08), reflecting the high degree of uncertainty in the current evidence base

Sensitivity analysis excluding Menon et al. yielded similar results (RR 0.75, 95% CI: 0.44–1.26; I^2=71.7%). Further exclusion of both Menon et al. and Gangopadhyay et al. demonstrated no significant difference between approaches (RR 1.04, 95% CI: 0.69–1.56; I^2=0.0%). Although

heterogeneity appeared absent, this estimate is based on only two studies and should be interpreted with caution, as I^2 cannot be reliably estimated with so few observations. Individual study results are summarized in Table 2.

Early Postoperative Infection

Three studies reported postoperative wound infection outcomes (Online Resource 3). However, substantial methodological limitations precluded quantitative synthesis.

Two studies reported infection data with unclear outcome denominators. Gangopadhyay et al. reported infection sample sizes that differed markedly from those used for continence analysis, suggesting cumulative counting of infections across multiple surgical stages rather than per definitive repair. Liu et al. similarly reported infection events likely accumulated across staged procedures, limiting interpretability of group comparisons.

Only Zhou et al. provided clearly defined per-procedure infection rates for the definitive repair. In this propensity score–matched cohort, wound infection occurred in 0 of 42 patients (0%) undergoing single-stage repair and in 2 of 42 patients (4.8%) undergoing staged repair ($p = 0.49$).

Given inconsistent reporting, unclear denominators, and high risk of bias in infection outcome assessment, no meta-analysis was performed for postoperative infection. Individual study findings are summarized in Online Resource 3.

Secondary Outcomes

Constipation

Two studies reported postoperative constipation outcomes. Liu et al. reported similar constipation rates between single-stage and staged repair (29/65 vs 23/48). Zhou et al. also reported comparable rates (10/33 vs 15/38). Due to limited data and heterogeneous outcome definitions, no quantitative synthesis was performed.

Mortality

Mortality outcomes were inconsistently reported. Liu et al. and Zhou et al. reported no deaths in either treatment group. Gangopadhyay et al. reported deaths in the single-stage group only, while staged mortality was not reported. Menon et al. reported deaths in both groups with markedly different rates. The absence of staged-group mortality data raises the possibility of selective outcome reporting and contributes to the serious risk of bias rating for this study. Given inconsistent reporting and unclear attribution of deaths, mortality outcomes were not pooled.

Reoperation, quality of life, and cost

These clinically relevant outcomes were not systematically reported across included studies and could not be analyzed.

Certainty of Evidence

The certainty of evidence was graded as LOW for bowel incontinence outcomes and VERY LOW for postoperative infection outcomes using the GRADE framework (Table 3).

Table 3
Summary of Findings and Certainty of Evidence (GRADE)

Outcome	No. of studies (patients)	Study design	Relative effect (RR, 95% CI)	Absolute effect*	Certainty of evidence	Explanation for certainty rating
Bowel incontinence (long-term)	4 (746)	Observational cohorts	RR 0.93 (0.52–1.66)	No clear difference between groups	**Low**	Downgraded for serious risk of bias (non-randomized designs, confounding by indication, attrition) and serious inconsistency (I^2 = 83%, opposing directions of effect)
Early postoperative infection	3 (—)**	Observational cohorts	Not pooled	Not estimable	**Very low**	Downgraded for very serious risk of bias (unclear denominators, cumulative event counting) and very serious imprecision (only one study with reliable per-procedure data)
Constipation	2 (184)	Observational cohorts	Not pooled	Similar rates reported	**Very low**	Downgraded for risk of bias, inconsistency in outcome definitions, and imprecision
Mortality	4 (746)	Observational cohorts	Not pooled	Inconsistent reporting	**Very low**	Downgraded for serious risk of bias, inconsistency, and imprecision
Reoperation / Quality of life / Cost	—	—	Not reported	Not available	Not assessable	Outcomes not systematically reported
* Absolute effects not calculated due to heterogeneity of outcome definitions and inconsistent denominators.						
** Infection outcomes were reported in three studies, but denominators were inconsistent and not comparable across studies; therefore, no pooled estimate was calculated.						
GRADE explanatory notes						
• All included studies were observational, starting at LOW certainty by GRADE methodology.						
• For bowel incontinence, evidence was downgraded two levels:						
o Risk of bias (noted but not separately downgraded—already captured in LOW starting point): non-randomized allocation, confounding by indication, substantial loss to follow-up.						
o Serious inconsistency: high statistical heterogeneity (I^2=83%) with opposing directions of effect across studies.						
• Imprecision was not further downgraded for incontinence, as total sample size was moderate and the primary limitation was inconsistency rather than sparse data.						
• For infection outcomes, certainty was downgraded to VERY LOW due to fundamental uncertainty in outcome definition and denominators, precluding valid quantitative synthesis.						

For bowel incontinence, starting at LOW for observational design (inherently accounting for non-randomized allocation and baseline confounding), evidence was downgraded one additional level for serious inconsistency (I^2=83%, opposing effect directions), yielding a final certainty of LOW.

Evidence for postoperative infection was downgraded for very serious risk of bias (unclear denominators, inconsistent reporting) and very serious imprecision (only one study with interpretable per-procedure data).

DISCUSSION

This systematic review identified four comparative studies enrolling 746 patients to evaluate single-stage versus staged reconstruction for high and intermediate anorectal malformations. Pooled analysis demonstrated no statistically significant difference in long-term bowel incontinence (RR 0.93, 95% CI: 0.52–1.66, p = 0.80), though substantial heterogeneity (I^2=83%) and wide prediction intervals (0.28–3.08) indicate considerable uncertainty. Early postoperative infection could not be reliably synthesized due to inconsistent outcome definitions and reporting across studies. The certainty of evidence was rated as LOW for continence and VERY LOW for postoperative infection.

A dominant assumption in pediatric surgery has been that primary neonatal repair compromises long-term continence because of technical challenges related to small pelvic dimensions, immature sphincter musculature, and limited visualization of the muscle complex [1, 5].

Despite the persistence of this belief, comparative evidence has never conclusively demonstrated worse functional outcomes following single-stage repair. The present meta-analysis aligns with this observation: incontinence occurs with comparable frequency after both approaches, though the wide confidence and prediction intervals preclude definitive conclusions in either direction. The absence of a detected difference should not be interpreted as proof of equivalence.

Interpretation in Clinical Context

The pooled estimate masks clinically important variation that merits attention. Gangopadhyay et al. reported substantially lower incontinence following single-stage repair (RR 0.58), while Menon et al. found the opposite (RR 1.70). These opposing findings likely reflect differences in patient selection rather than true differences in treatment effect. The decision between single-stage and staged repair is driven by anatomic factors assessed intraoperatively and on preoperative imaging—including sacral ratio, presacral air position on cross-table prone lateral radiograph (CTPL), fistula level, presence of associated spinal anomalies, and the surgeon's assessment of whether adequate muscle complex visualization can be achieved without colostomy [5, 12]. These factors produce confounding by indication unresolvable through retrospective analysis.

Single-stage repair requires reconstruction through a neonatal pelvis approximately one-third the size encountered during delayed repair. Proponents cite avascular, well-defined tissue planes facilitating dissection [9, 24]; opponents warn of increased risk to the muscle complex and urethra in the diminutive operative field [1, 5]. Neither position has been definitively supported by comparative data.

Liu (2004) employed open PSARP, while Zhou (2022) used laparoscopy-assisted anorectoplasty from a high-volume center; comparing outcomes across these eras conflates the staging decision with advances in visualization and perioperative care.

Limitations of the Evidence Base

Three structural issues explain why clinical uncertainty persists despite two decades of comparative data.

First, the absence of outcome standardization has fragmented the evidence base. The modified Wingspread classification, Kelly score, and Bowel Function Score have different domains and scoring thresholds with no published cross-validation. Pooling these instruments is arithmetically possible but clinically problematic, as the direction of cross-instrument misclassification is unpredictable. Future studies must adopt a single validated instrument to enable meaningful comparison.

Second, temporal evolution of surgical technique confounds the comparison. In the included studies, the staged cohorts were frequently drawn from earlier time periods (most explicitly in Gangopadhyay et al., where staged patients were treated from 1989–1996 and single-stage patients from 1996–2003). This historical control design means that differences between approaches are inseparable from concurrent advances in perioperative care and operative technique.

Third, confounding by indication is nearly absolute. The staging decision reflects a composite intraoperative judgment based on CTPL findings, fistula level, associated anomalies, and surgeon confidence—factors incompletely captured in retrospective data. Propensity score matching, as employed by Zhou et al., addresses measured confounders but cannot account for unmeasured clinical assessment. These three structural issues require fundamentally different study designs to address.

Hartford et al. (2022) and Miscia et al. (2024) pooled low-type ARM alongside high and intermediate lesions [3, 4], inflating the single-stage denominator with patients never considered for staged repair [1, 25]. When restricted to the anatomically appropriate population, the apparent advantage disappears. Additionally, previous meta-analyses pooled different continence instruments without acknowledging inferential limitations, and infection outcomes were not scrutinized—two of three studies likely used cumulative event counting across procedures, rendering per-patient rates incomparable.

Implications for Practice and Research

A randomized controlled trial is unlikely to be realized and may not be appropriate for this question. Randomization would require surgeons to perform single-stage repair on anatomy they judge unsuitable, or impose colostomy on neonates they believe could safely undergo definitive repair—violating equipoise at the individual level. The rarity of the condition further limits feasibility; even high-volume centers enroll fewer than 10–15 eligible patients per year.

We propose three components for interpretable comparative evidence. Prospective international registry enrollment capturing standardized anatomic characterization, operative documentation (including surgeon's rationale for staging decision), and blinded outcome assessment using a prespecified validated instrument. The ARM-Net consortium provides existing infrastructure. Advanced observational designs applied to registry data, including instrumental variable analysis and regression discontinuity designs, can address unmeasured confounding that

propensity score methods cannot eliminate. Infection outcomes should be reported using established surgical site infection definitions with procedure-specific denominators and defined surveillance windows.

The choice between single-stage and staged reconstruction should remain individualized based on anatomic complexity, surgeon experience, institutional resources, and family preferences. The present analysis found no evidence that single-stage repair results in worse long-term continence—but certainty is low, and the prediction interval encompasses both substantial benefit and harm.

In resource-limited settings where stoma care resources are constrained and loss to follow-up is high [6–8], single-stage repair offers reduced treatment burden and improved completion rates—Gangopadhyay et al. reported that only 43.7% of staged patients completed all three operations [9]. Conversely, centers with established staged protocols may reasonably continue current practice. Shared decision-making should explicitly acknowledge uncertainty; framing single-stage repair as “riskier” or staged as the “safer default” is not supported by available evidence.

Strengths and Limitations

This review was prospectively registered and restricted to anatomically appropriate populations, avoiding dilution by low-type malformations that compromised previous syntheses. Systematic risk-of-bias and certainty-of-evidence assessments were applied, and unreliable infection data were not pooled—preserving interpretive integrity. Independent dual review with high inter-rater agreement reduced selection bias.

The primary limitation is the quality of available evidence. All studies were observational with non-randomized allocation, subject to confounding by indication and heterogeneous outcome definitions. Two studies demonstrated serious risk of bias due to historical control designs and substantial attrition. Our search encompassed three major biomedical databases (MEDLINE, Embase, CENTRAL) and two trial registries; while additional regional indices were not searched, the comprehensive coverage of MEDLINE and Embase for surgical literature makes it unlikely that eligible studies were missed. Two of four included studies enrolled only male patients, limiting generalizability to females with high/intermediate ARM. A planned secondary analysis of single-arm case series was not performed due to insufficient homogeneity across studies. These limitations reflect deficiencies of the primary literature rather than shortcomings of the review methodology.

CONCLUSIONS

Current observational evidence does not demonstrate superiority of either single-stage or staged reconstruction for bowel continence outcomes (certainty of evidence: LOW, GRADE). Evidence for postoperative infection outcomes was insufficient for quantitative synthesis; rates varied substantially across studies (range: 0–33%) and no consistent direction of effect was identified (certainty of evidence: VERY LOW, GRADE). The wide confidence and prediction intervals indicate substantial uncertainty in these estimates. Despite long-standing assumptions regarding neonatal single-stage repair, available data are insufficient to confirm or refute concerns about inferior functional outcomes. Given low-certainty evidence and methodological limitations, treatment should remain individualized, and prospective registry-based studies with standardized outcome assessment are required to reduce uncertainty.

Declarations

ACKNOWLEDGMENTS

The authors acknowledge the use of artificial intelligence–based tools to assist with language editing and structured verification support. All methodological decisions, analyses, and interpretations were made by the authors, who take full responsibility for the content and conclusions of this work. No external funding was received.

AUTHOR CONTRIBUTIONS

Tuyen Tran: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft, Writing – review & editing.

Chuong Ngoc Thi Dang: Writing – original draft, Writing – review & editing.

Chieu Vuong Minh: Investigation (independent screening, data extraction, risk-of-bias assessment), Writing – review & editing.

All authors approved the final manuscript and are accountable for all aspects of the work.

DATA AVAILABILITY

All data extracted for this systematic review are available in the supplementary materials. The search strategy and full-text articles assessed for eligibility are available from the corresponding author upon reasonable request.

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Funding No funds, grants, or other support was received.

Ethics approval Not applicable. This study is a systematic review of published literature and did not involve human participants, their data, or biological material.

Protocol registration: PROSPERO CRD420261291319

References

1. Levitt MA, Peña A (2007) Anorectal malformations. Orphanet J Rare Dis 2:33. https://doi.org/10.1186/1750-1172-2-33
2. Holschneider A, Hutson J, Peña A, Beket E, Chatterjee S, Coran A et al (2005) Preliminary report on the International Conference for the Development of Standards for the Treatment of Anorectal Malformations. J Pediatr Surg 40:1521–1526. https://doi.org/10.1016/j.jpedsurg.2005.08.002
3. Hartford L, Brisighelli G, Gabler T, Westgarth-Taylor C (2022) Single-stage procedures for anorectal malformations: a systematic review and meta-analysis. J Pediatr Surg 57:75–84. https://doi.org/10.1016/j.jpedsurg.2021.12.024
4. Miscia ME, Lauriti G, Di Renzo D, Cascini V, Lisi G (2024) Short and long-term outcomes of PSARP versus LAARP and single versus staged repair for infants with high-type anorectal malformations: a systematic review and meta-analysis. Child (Basel) 11:376. https://doi.org/10.3390/children11030376
5. Bischoff A, Bealer J, Peña A (2017) Controversies in anorectal malformations. Lancet Child Adolesc Health 1:323–330. https://doi.org/10.1016/S2352-4642(17)30026-3
6. Chandramouli B, Srinivasan K, Jagdish S, Ananthakrishnan N (2004) Morbidity and mortality of colostomy and its closure in children. J Pediatr Surg 39:596–599. https://doi.org/10.1016/j.jpedsurg.2003.12.016
7. Patwardhan N, Kiely EM, Drake DP, Spitz L, Pierro A (2001) Colostomy for anorectal anomalies: high incidence of complications. J Pediatr Surg 36:795–798. https://doi.org/10.1053/jpsu.2001.22963
8. Pena A, Migotto-Krieger M, Levitt MA (2006) Colostomy in anorectal malformations: a procedure with serious but preventable complications. J Pediatr Surg 41:748–756. https://doi.org/10.1016/j.jpedsurg.2005.12.021
9. Gangopadhyay AN, Gopal SC, Sharma S, Gupta DK, Sharma SP, Mohan TV (2006) Management of anorectal malformations in Varanasi, India: a long-term review of single and three stage procedures. Pediatr Surg Int 22:169–172. https://doi.org/10.1007/s00383-005-1567-6
10. Xiao H, Chen L, Ren X-H, Huang R, Diao M, Li L (2018) One-stage laparoscopic-assisted anorectoplasty for neonates with anorectal malformation and recto-prostatic or recto-bulbar fistula according to the Krickenbeck classification. J Laparoendosc Adv Surg Tech A 28:1029–1034. https://doi.org/10.1089/lap.2017.0690
11. Zhou Y, Ye S, Yang C, Gao R, Ming A-X, Diao M et al (2022) One-stage versus staged laparoscopic-assisted anorectoplasty for high-and intermediate anorectal malformations: medium-term outcomes from a propensity score matching analysis. J Pediatr Surg 57:555–560. https://doi.org/10.1016/j.jpedsurg.2022.01.018
12. Menon P, Rao KLN, Sinha AK, Lokesha K, Samujh R, Mahajan JK et al (2017) Anorectal malformations in males: pros and cons of neonatal versus staged reconstruction for high and intermediate varieties. J Indian Assoc Pediatr Surg 22:83–86. https://doi.org/10.4103/jiaps.JIAPS_15_17
13. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372:n71. https://doi.org/10.1136/bmj.n71
14. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M et al (2016) ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 355:i4919. https://doi.org/10.1136/bmj.i4919
15. DerSimonian R, Laird N (2015) Meta-analysis in clinical trials revisited. Contemp Clin Trials 45:139–145. https://doi.org/10.1016/j.cct.2015.09.002
16. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ et al (2019) Cochrane handbook for systematic reviews of interventions, 2nd edn. Wiley, Chichester
17. GRADE Working Group (n d.) GRADE handbook. https://book.gradepro.org/guideline/overview-of-the-grade-approach. Accessed 9 Feb 2026

18. Diao M, Li L, Ye M, Guan K-P, Wei Y-D, Cheng W (2016) Congenital anomaly rectified at birth: one-stage single-incision laparoscopic-assisted anorectoplasty for newborns with anorectal malformations and recto-urethral fistula. Surg Endosc 30:5156–5164. https://doi.org/10.1007/s00464-016-4841-x
19. Gangopadhyay AN, Pandey V (2017) Controversy of single versus staged management of anorectal malformations. Indian J Pediatr 84:636–642. https://doi.org/10.1007/s12098-017-2373-6
20. Nagdeve NG, Bhingare PD, Naik HR (2011) Neonatal posterior sagittal anorectoplasty for a subset of males with high anorectal malformations. J Indian Assoc Pediatr Surg 16:126–128. https://doi.org/10.4103/0971-9261.86863
21. Osifo OD, Osagie TO, Udefiagbon EO (2014) Outcome of primary posterior sagittal anorectoplasty of high anorectal malformation in well selected neonates. Niger J Clin Pract 17:1–5. https://doi.org/10.4103/1119-3077.122821
22. Vick LR, Gosche JR, Boulanger SC, Islam S (2007) Primary laparoscopic repair of high imperforate anus in neonatal males. J Pediatr Surg 42:1877–1881. https://doi.org/10.1016/j.jpedsurg.2007.07.014
23. Adeniran JO, Abdur-Rahman L (2005) One-stage correction of intermediate imperforate anus in males. Pediatr Surg Int 21:88–90. https://doi.org/10.1007/s00383-004-1211-x
24. Liu G, Yuan J, Geng J, Wang C, Li T (2004) The treatment of high and intermediate anorectal malformations: one stage or three procedures? J Pediatr Surg 39:1466–1471. https://doi.org/10.1016/j.jpedsurg.2004.06.021
25. Amanollahi O, Ketabchian S (2016) One-stage vs. three-stage repair in anorectal malformation with rectovestibular fistula. Afr J Paediatr Surg 13:20–25. https://doi.org/10.4103/0189-6725.181702

Figures

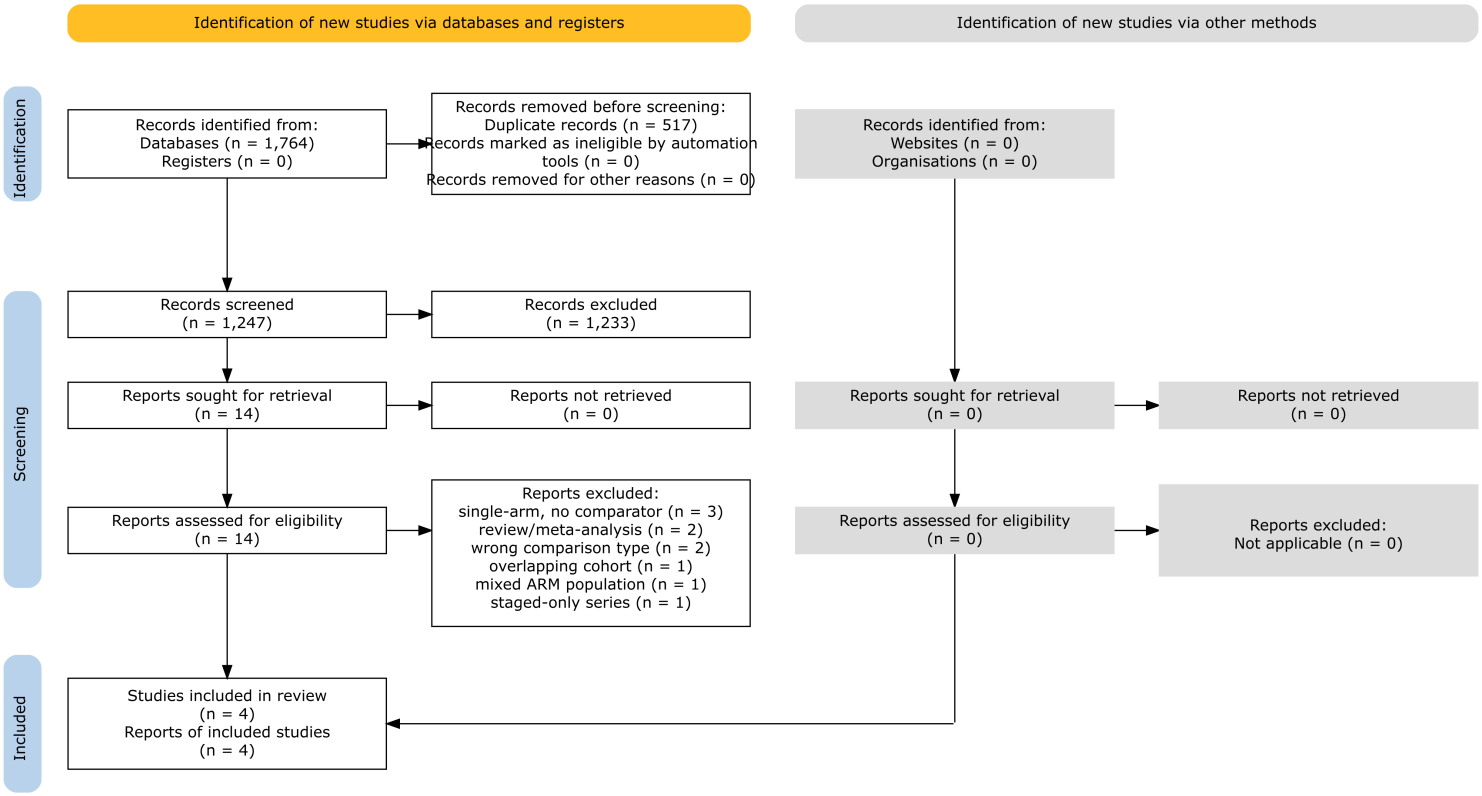


Figure 1

PRISMA 2020 flow diagram of study identification, screening, eligibility assessment, and inclusion

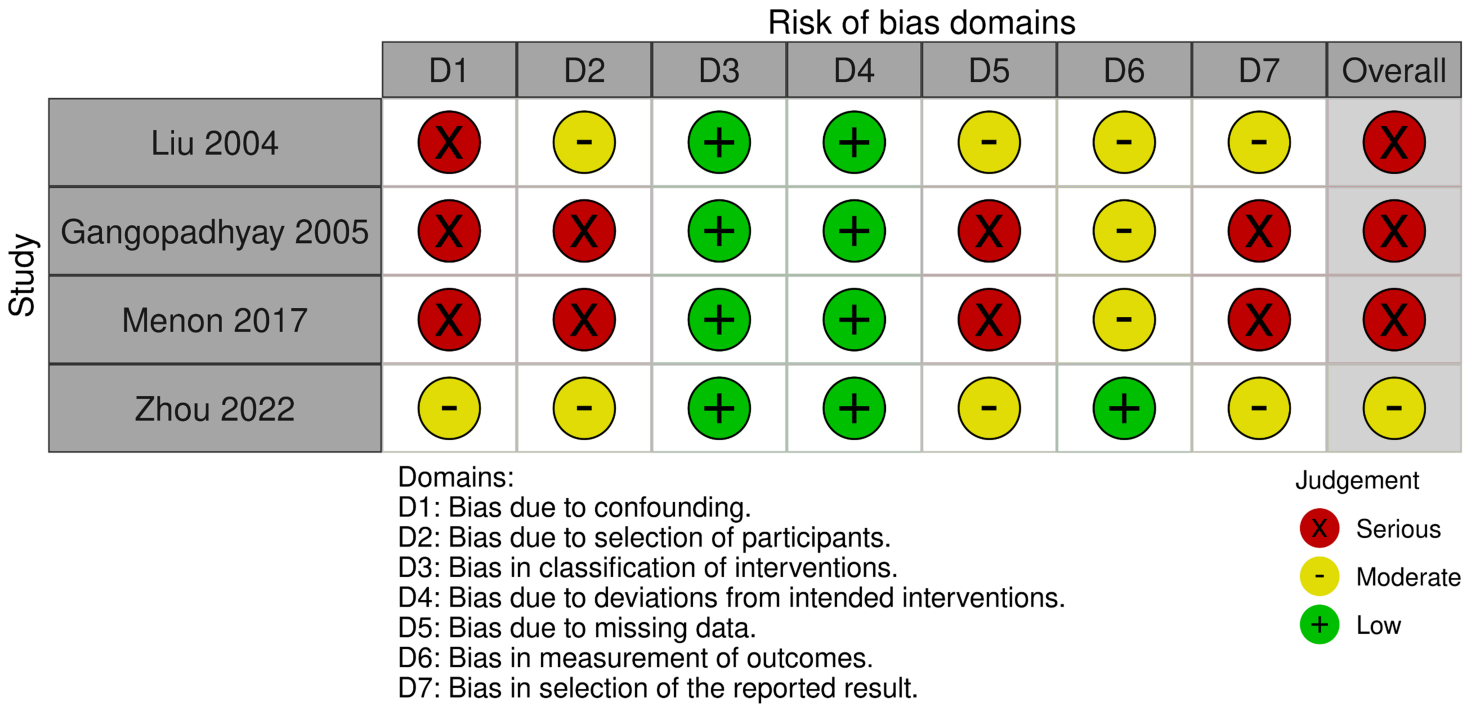


Figure 2

Risk of bias assessment of included studies using the ROBINS-I tool (traffic light plot)

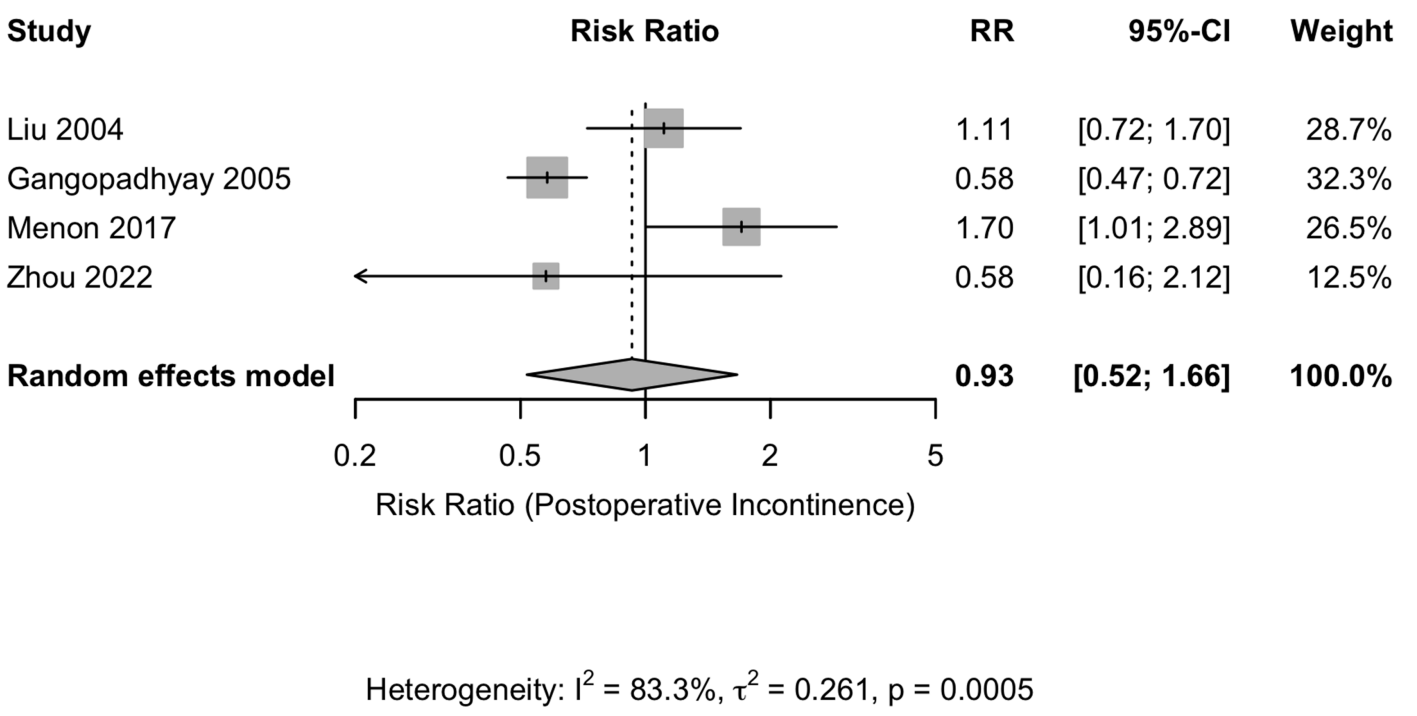


Figure 3

Forest plot comparing long-term bowel continence between single-stage and staged reconstruction (random-effects meta-analysis)

Supplementary Files

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

- ELECTRONICSUPPLEMENTARYMATERIALlegends.docx
- Supplementary.docx
- PRISMA2020Checklist.docx