

Effect of Pre-labor Amniotomy Following Balloon Induction on Delivery Timing and Infectious Outcomes: A Propensity-Score-Matched Retrospective Cohort Study

Ying Xu

China-Japan Friendship Hospital

Wenhui Li

China-Japan Friendship Hospital

Cuiping Feng

China-Japan Friendship Hospital

Yanting Chen

China-Japan Friendship Hospital

Zhenzhen Zhang

China-Japan Friendship Hospital

Xiaoyu Pan

changpanxiaoyu@163.com

China-Japan Friendship Hospital

Research Article

Keywords: labor induction, balloon catheter, amniotomy, chorioamnionitis, intrapartum fever, propensity-score matching, retrospective cohort study

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10977510/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Balloon induction is a common method for cervical ripening. Most studies treat balloon induction as a single intervention without exploring whether amniotomy should be performed prior to the onset of labor after balloon removal—a controversial decision, as amniotomy may accelerate delivery but also disrupt the membrane barrier and increase the risk of ascending infection.

Methods

This was a single-center retrospective cohort study conducted at a tertiary hospital in China. Participants were grouped according to clinical pathways: the balloon group (no pre-labor amniotomy; $n = 86$) and the pre-labor amniotomy after balloon group ($n = 76$); the primary spontaneous amniotomy ($n = 12$) was included only for descriptive analysis. The primary outcome was documented clinical chorioamnionitis. Key secondary outcomes included intrapartum fever $\geq 38.0^{\circ}\text{C}$; neonatal sepsis, pneumonia, and NICU admission as safety outcomes. The primary analysis used propensity-score matching (1:1 nearest neighbor matching with a 0.2 standard deviation (SD) caliper); matched binary outcomes were analyzed using exact McNemar tests, calculating paired odds ratios (OR) and exact 95% confidence intervals (CI). Pre-specified secondary subgroups excluded controls who went into labor spontaneously within 0–6 h after balloon removal, and re-analysis was performed among women facing this decision.

Results

Among 174 women, 59 pairs completed matching (core covariate balanced). Chorioamnionitis rates: 4/59 (6.8%) vs. 3/59 (5.1%), paired OR 1.33 (95% CI 0.23–9.10; $P = 1.000$); Firth-adjusted OR 1.81 (0.52–6.89). Fever $\geq 38.0^{\circ}\text{C}$: 5/59 vs. 1/59, paired OR 5.00 (0.56–236.49; $P = 0.219$), consistent in direction across methods. Neonatal sepsis and pneumonia: 0/174. NICU admission: 6/59 vs. 0/59 ($P = 0.031$), attenuated after exact matching. No difference in time from induction to delivery post-matching (-0.35 h, $P = 0.098$). In secondary subgroups, time from induction to delivery increased with longer waiting thresholds (from 29.5 to 33.8 h), while the amniotomy pathway remained at 26.9 h; the gap in vaginal delivery within 24 h widened progressively (most stringent threshold: 30.3% vs. 7.5%, OR 5.35, $P = 0.005$); chorioamnionitis remained negative throughout. Intrapartum antibiotics: 93.4% vs. 46.5% ($P < 0.001$).

Conclusions

Among women not yet in labor after balloon removal, pre-labor amniotomy is associated with faster delivery, demonstrating a clear advantage for vaginal delivery within 24 hours, without evidence of increased confirmed infection or postpartum hemorrhage-related risk; signals of intrapartum fever are

consistent in direction but imprecise, requiring validation by larger studies. Based on current evidence, pre-labor amniotomy is a reasonable option for accelerating delivery.

Background

Labor induction is one of the most common obstetric interventions. For women with an unfavorable cervix, balloon catheters (Foley catheter or double-balloon) facilitate cervical ripening through mechanical dilation without increasing the risk of uterine hyperstimulation; the World Health Organization (WHO) recommends balloon catheters for induction, noting their particular utility when uterine hyperstimulation should be avoided or prostaglandins are contraindicated [1].

In this cohort, all women achieved cervical ripening following balloon induction. Once the cervix is ripe and the balloon is removed, a subsequent decision point arises: if the woman has not yet gone into labor, should amniotomy be performed immediately to initiate or accelerate delivery, or should expectant management be adopted, awaiting spontaneous labor, spontaneous rupture, or oxytocin? This decision involves a trade-off between accelerating delivery and disrupting the membrane barrier. Amniotomy can induce or augment contractions, shorten the time from induction to delivery, thereby reducing the window of bacterial exposure post-rupture; conversely, it disrupts the membrane barrier and prolongs the interval from rupture to delivery, potentially increasing opportunities for ascending infection. Previous studies comparing different induction methods and infectious outcomes have largely treated balloon induction as a single intervention or compared balloon with oxytocin or prostaglandins; few have specifically examined the scenario of pre-labor following cervical ripening with a balloon [2].

Such real-world comparative analyses are subject to inherent structural bias: women whose trajectory rapidly progresses to labor after balloon removal have better outcomes, but they can only enter the non-amniotomy pathway, systematically enriching the comparison group with favorable outcomes. If this treatment escalation selection bias (a form of guaranteed-time bias) remains unaddressed, crude comparisons may simultaneously overestimate infection risk and underestimate the efficiency advantage of the amniotomy pathway.

Therefore, our objective was to compare pre-labor amniotomy pathway versus the balloon path (spontaneous rupture or post-labor amniotomy) among women undergoing balloon cervical ripening, evaluating: (1) the primary outcome of documented clinical chorioamnionitis; (2) key secondary outcomes including intrapartum fever (maximum temperature $\geq 38.0^{\circ}\text{C}$), neonatal infection-related outcomes, and delivery efficiency; and (3) the robustness of these associations in pre-specified secondary subgroup analyses restricted to women not in labor after balloon removal—i.e., the actual population facing the amniotomy decision.

Methods

Study Design and Data Sources

This was a single-center retrospective cohort study of women delivering at the China-Japan Friendship Hospital from December 2024 to November 2025. Data were sourced from the hospital's obstetric electronic medical record system and supplemented through structured chart review. The report adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [3]. This study was approved by the Clinical Research Ethics Committee of the hospital (Approval No.: 2026-KY-290, 15 July 2026). Given the retrospective nature of the study, the institutional review board waived the requirement for written informed consent.

Participants and Groups

We included women who underwent induction of labor and delivery at our hospital with complete extraction of key variables; all enrolled women achieved cervical ripening after balloon induction. Two primary comparison groups were defined based on the clinical pathway: the balloon group, comprising pregnant women who received balloon cervical ripening without antepartum amniotomy (including cases of spontaneous rupture of membranes, intrapartum amniotomy, and no rupture of membranes prior to delivery); and the pre-labor amniotomy group, comprising pregnant women who underwent amniotomy after balloon cervical ripening but before the onset of labor. Concurrently, 12 women with favorable cervical conditions underwent primary amniotomy for induction; due to the small sample size and differing case composition, this group was subjected only to descriptive analysis without inferential comparisons. Grouping rules and exclusion criteria were prespecified (Fig. 1).

Exposure Definition

The exposure was pre-labor amniotomy following balloon cervical ripening, determined by the recorded mode of membrane rupture and its timing relative to the onset of labor. The time interval from balloon removal to amniotomy (median 1.17 h, IQR 0.00–24.70 h) was used for timing analyses.

Outcome Definitions

The primary outcome was the clinical diagnosis of chorioamnionitis recorded in discharge summaries or progress notes. A secondary key outcome was an intrapartum peak temperature of at least 38.0°C; a lower threshold of > 37.8°C was used in a sensitivity analysis. Neonatal outcomes included sepsis, pneumonia, admission to the neonatal intensive care unit (NICU), and Apgar scores. Based on the method of original record-keeping, unrecorded fetal sepsis and pneumonia were considered absent, unrecorded NICU status was considered not transferred, and unrecorded Apgar scores were considered 10. Neonatal white blood cell counts and C-reactive protein were tested only when there were clear clinical indications and were reported descriptively only for infants tested for these markers; no between-group comparisons were performed. Time-to-event outcomes included induction-to-labor, induction-to-delivery, and rupture-to-delivery times, together with vaginal delivery within 24 and 48 h post-induction. Process outcomes included the rate of vaginal delivery (including operative vaginal delivery) and the total and post-rupture numbers of vaginal examinations. Mode of delivery and complication outcomes included cesarean section, surgical delivery (cesarean section or operative vaginal delivery), postpartum hemorrhage, and fetal distress. The following covariates were used for adjustment: age, gestational age at delivery, gravidity, nulliparity

(yes/no), pre-pregnancy body mass index (BMI), admission Bishop score, group B streptococcus (GBS) status, diabetes, hypertension, and history of prior cesarean section. Missing pre-pregnancy BMI values were imputed with the median, and missing Bishop scores were imputed with the cohort median (4 points); both variables were accompanied by a missing-indicator variable.

Statistical Analysis

Continuous variables are presented as medians and interquartile ranges; comparisons among three groups used the Kruskal-Wallis test, and pairwise comparisons used the Mann-Whitney U test. Categorical variables are presented as frequencies and percentages; 2x2 contingency tables were analyzed using Fisher's exact test, and larger tables used the Fisher-Freeman-Halton exact test. The primary analysis used propensity-score matching [4]: propensity scores were estimated via logistic regression including the aforementioned covariates (12 terms in total, including missing indicators), and 1:1 nearest-neighbor matching without replacement was performed using 0.2 times the SD of the logit of the propensity score as the caliper width [5]. Oxytocin administration and membrane rupture-related variables reflect post-procedural clinical decisions rather than baseline patient characteristics, and were thus excluded from the propensity score model. Matching quality was assessed using standardized mean differences (SMDs), with values below 0.1 considered indicative of good balance [6]. Binary outcomes in the matched cohort were compared using exact McNemar's test; paired ORs and their 95% CIs were derived from the binomial distribution. Continuous outcomes were compared using the Wilcoxon signed-rank test, and CIs for median differences were estimated via bootstrapping with 2,000 resamplings. In the full cohort, adjusted ORs were estimated using Firth-penalized logistic regression suitable for rare outcomes [7], with group assignment as the exposure variable and age, gestational age, nulliparity, Bishop score, and history of prior cesarean section as covariates; because only three women had missing Bishop scores, models including missingness indicators were unstable, so complete-case results were reported as the primary findings. Sensitivity analyses set the caliper at 0.1 and 0.3 times the SD, applied exact matching within strata of hypertension, and used inverse probability of treatment weighting (IPTW). All tests were two-sided, with $P < 0.05$ considered statistically significant; given the presence of multiple secondary and exploratory outcomes, no multiplicity correction was performed, and these results should be interpreted as exploratory rather than confirmatory.

Women who went into labor immediately upon balloon removal never faced the decision regarding amniotomy, yet they were all entered into the comparison groups and experienced favorable outcomes; thus, crude comparisons would overestimate risk and obscure the benefits of amniotomy. As a sensitivity analysis, we conducted a secondary stratification analysis similar to a landmark approach: using exclusion thresholds of 0, 2, 4, and 6 h post-balloon removal, we excluded comparison-group women who went into labor spontaneously within each threshold (cases going into labor at or before removal were counted under the 0-hour threshold) and repeated the above comparisons among women who actually faced the decision regarding amniotomy. Since grouping was conditional on a post-treatment variable (onset of labor), the results should be interpreted as applicable to the population defined by that condition; findings across all exclusion thresholds are reported.

To distinguish infectious from non-infectious components of intrapartum fever, we performed a matched analysis on the duration of epidural analgesia: up to four fever-free controls were matched to each febrile case within the same group, requiring identical oxytocin status and similar baseline characteristics. For each matched set, the index time was defined as the interval from the onset of confirmed labor in the case to the onset of fever; controls were followed from their own onset of labor to the same elapsed labor time, and cumulative epidural analgesia durations up to the index time were compared between groups. The calendar day of fever onset was taken as the delivery day. For controls, the calendar days of epidural analgesia initiation and cessation were inferred from recorded labor-onset and delivery times. Interruptions in epidural infusion could not be obtained and were assumed to be zero duration. Controls who delivered prior to the index time were excluded from the primary version of this exploratory analysis. Group comparisons used the Mann-Whitney U test, and paired comparisons used the Wilcoxon signed-rank test. All analyses were conducted in Python (version 3.12).

Results

Cohort Characteristics

Among 174 women, 86 followed the balloon induction pathway, 76 followed the pre-labor amniotomy pathway, and 12 had primary amniotomy. Except for slightly lower gestational age ($P = 0.037$), higher pre-pregnancy BMI ($P = 0.025$), a higher proportion of hypertension (22.4% vs. 9.3%, $P = 0.029$), and fewer GBS-positive screenings (0/76 vs. 4/85, $P = 0.123$) in the amniotomy group, major baseline characteristics were balanced between groups; Bishop scores at admission were identical (median 4 in both, $P = 0.658$) (Table 1). The primary amniotomy group had higher Bishop scores (median 6), with 5/12 developing chorioamnionitis, 5/12 experiencing fever, and 12/12 receiving intrapartum antibiotic treatment; this group is presented descriptively only.

Table 1
Baseline Characteristics by Group

Variable	Type	Balloon pathway	Pre-labor amniotomy	Primary amniotomy (descriptive)	P (3 groups)	P (balloon vs amniotomy)
Age, years	Continuous, median (IQR)	32.50 (30.00–34.75)	33.00 (31.00–35.00)	32.50 (31.00–36.75)	0.422	0.289
GA at delivery, days	Continuous, median (IQR)	282.00 (280.00–285.00)	281.00 (276.00–284.00)	281.00 (277.50–283.25)	0.083	0.037
Pre-pregnancy BMI, kg/m2	Continuous, median (IQR)	20.96 (19.26–23.66)	22.06 (20.41–24.29)	20.96 (20.77–23.47)	0.068	0.025
Admission Bishop score	Continuous, median (IQR)	4.00 (3.00–5.00)	4.00 (3.00–5.00)	6.00 (5.50–7.00)	0.001	0.658
Gravidity	Continuous, median (IQR)	1.00 (1.00–2.00)	1.00 (1.00–2.00)	2.00 (1.00–2.00)	0.549	0.693
Parity before delivery	Continuous, median (IQR)	0.00 (0.00–0.00)	0.00 (0.00–0.00)	0.50 (0.00–1.00)	0.066	0.514
Nulliparous	Binary, n/N (%)	66/86 (76.7%)	62/76 (81.6%)	6/12 (50.0%)	0.068	P = 0.563
GBS positive	Binary, n/N (%)	4/85 (4.7%)	0/76 (0.0%)	0/12 (0.0%)	0.173	P = 0.123
Diabetes	Binary, n/N (%)	23/86 (26.7%)	24/76 (31.6%)	3/12 (25.0%)	0.807	P = 0.603
Hypertension	Binary, n/N (%)	8/86 (9.3%)	17/76 (22.4%)	0/12 (0.0%)	0.024	P = 0.029
Thyroid disorder	Binary, n/N (%)	6/86 (7.0%)	8/76 (10.5%)	4/12 (33.3%)	0.038	P = 0.577
Prior cesarean section	Binary, n/N (%)	1/86 (1.2%)	0/76 (0.0%)	0/12 (0.0%)	1.0	P = 1.000

Propensity-Score Matching Quality

The propensity score model (12 covariates) converged in the pooled data of 162 women (McFadden pseudo R-squared 0.086). One-to-one nearest-neighbor matching (0.2 SD caliper, without replacement) formed 59 pairs; 17 treated and 27 control women remained unmatched. The maximum absolute SMD before matching was 0.361; after matching, |SMD| for all core covariates (age, gestational age, BMI, Bishop score, nulliparity, GBS, prior cesarean) was ≤ 0.064 . Residual imbalance was limited to sparse indicators

(Bishop missing indicator – 0.184, three women in total; BMI missing indicator – 0.153; hypertension – 0.121), examined further by exact matching within the hypertension stratum (61 pairs).

Primary Outcome: Clinical Chorioamnionitis

In matched pairs, chorioamnionitis rates were 4/59 (6.8%) vs. 3/59 (5.1%), with a paired OR of 1.33 (95% CI 0.23–9.10), exact McNemar P value of 1.000, and paired risk difference of 0.02 (95% CI -0.07–0.10).

Unmatched data: 6/76 (7.9%) vs. 5/86 (5.8%), Fisher P value of 0.758; Firth-adjusted OR (complete Bishop scores) was 1.81 (95% CI 0.52–6.89; P = 0.353); inverse probability weighting showed 7.9% vs. 5.2%. All methods consistently indicated no difference, but CIs were wide (Tables 2–3).

Table 2
Outcomes by Group (Unmatched)

Variable	Type	Balloon pathway	Pre-labor amniotomy	Primary amniotomy (descriptive)	P (3 groups)	P (balloon vs amniotomy)
Clinical chorioamnionitis (primary outcome)	Binary, n/N (%)	5/86 (5.8%)	6/76 (7.9%)	5/12 (41.7%)	0.003	OR = 0.74, P = 0.758
Intrapartum fever (max temp > = 38.0 C)	Binary, n/N (%)	2/86 (2.3%)	7/76 (9.2%)	5/12 (41.7%)	0.0	OR = 0.25, P = 0.092
Intrapartum fever (max temp > 37.8 C, sensitivity definition)	Binary, n/N (%)	5/86 (5.8%)	12/76 (15.8%)	5/12 (41.7%)	0.002	OR = 0.37, P = 0.077
Intrapartum antibiotics	Binary, n/N (%)	40/86 (46.5%)	71/76 (93.4%)	12/12 (100.0%)	0.0	OR = 0.50, P = 0.007
Vaginal delivery (incl. forceps)	Binary, n/N (%)	75/86 (87.2%)	56/76 (73.7%)	9/12 (75.0%)	0.074	
Spontaneous vaginal delivery	Binary, n/N (%)	66/86 (76.7%)	46/76 (60.5%)	7/12 (58.3%)	0.053	
Vaginal delivery within 24 h	Binary, n/N (%)	38/86 (44.2%)	23/76 (30.3%)	8/12 (66.7%)	0.026	
Vaginal delivery within 48 h	Binary, n/N (%)	68/86 (79.1%)	41/76 (53.9%)	9/12 (75.0%)	0.002	
Cesarean section	Binary, n/N (%)	11/86 (12.8%)	20/76 (26.3%)	3/12 (25.0%)	0.074	
Operative delivery (cesarean or forceps)	Binary, n/N (%)	20/86 (23.3%)	30/76 (39.5%)	5/12 (41.7%)	0.053	
Postpartum hemorrhage	Binary, n/N (%)	9/86 (10.5%)	15/76 (19.7%)	2/12 (16.7%)	0.215	
Fetal distress	Binary, n/N (%)	10/86 (11.6%)	10/76 (13.2%)	5/12 (41.7%)	0.035	
Neonatal sepsis (blank = none)	Binary, n/N (%)	0/86 (0.0%)	0/76 (0.0%)	0/12 (0.0%)	1.0	
Neonatal pneumonia (blank = none)	Binary, n/N (%)	0/86 (0.0%)	0/76 (0.0%)	0/12 (0.0%)	1.0	
NICU admission (blank = none)	Binary, n/N (%)	3/86 (3.5%)	7/76 (9.2%)	1/12 (8.3%)	0.258	OR = 0.38, P = 0.200

Variable	Type	Balloon pathway	Pre-labor amniotomy	Primary amniotomy (descriptive)	P (3 groups)	P (balloon vs amniotomy)
Apgar 1 min < 7 (blank = 10)	Binary, n/N (%)	0/86 (0.0%)	0/76 (0.0%)	0/12 (0.0%)	1.0	
Apgar 5 min < 7 (blank = 10)	Binary, n/N (%)	0/86 (0.0%)	0/76 (0.0%)	0/12 (0.0%)	1.0	
Induction-to-labor, h	Continuous, median (IQR)	20.08 (9.24–24.98)	19.50 (17.48–41.00)	1.07 (0.98–1.38)		
Induction-to-delivery, h	Continuous, median (IQR)	25.30 (16.78–33.50)	26.92 (21.68–49.02)	8.34 (3.93–12.54)		
Rupture-to-delivery, h	Continuous, median (IQR)	2.88 (1.31–6.45)	8.39 (5.81–11.75)	8.34 (3.93–12.54)		
Total vaginal examinations	Continuous, median (IQR)	6.00 (5.00–8.00)	6.00 (5.00–8.00)	6.50 (5.00–8.50)		
Post-rupture vaginal examinations	Continuous, median (IQR)	3.00 (2.00–5.00)	5.00 (4.00–6.00)	5.00 (3.75–6.25)		
Neonatal WBC (measured only)	Continuous, median (IQR)	23.34 (16.56–26.11)	25.20 (19.83–28.38)	20.37 (17.45–21.23)		
Neonatal CRP (measured only)	Continuous, median (IQR)	2.50 (1.22–7.15)	4.52 (2.50–5.69)	2.50 (2.50–4.03)		
Cord blood pH (measured only)	Continuous, median (IQR)	7.27 (7.24–7.30)	7.22 (7.20–7.26)	7.28 (7.25–7.30)		
Birth weight, g	Continuous, median (IQR)	3400.00 (3092.50–3637.50)	3357.50 (3040.00–3552.50)	3337.50 (3107.50–3650.00)		
Removal-to-amniotomy interval, h	Continuous, median (IQR)	7.82 (3.32–11.96)	1.17 (0.00–24.70)			

Table 3
Matched-Cohort Paired Outcomes (59 Pairs)

Outcome	Pre-labor amniotomy	Balloon pathway	Paired OR (95% CI)	P
Clinical chorioamnionitis (primary outcome)	4/59	3/59	1.33 (0.23–9.10)	1.0
Intrapartum fever (max temp > = 38.0 C)	5/59	1/59	5.00 (0.56–236.49)	0.219
Intrapartum fever (max temp > 37.8 C, sensitivity definition)	10/59	3/59	4.50 (0.93–42.80)	0.065
Intrapartum antibiotics	55/59	27/59	15.00 (3.81–129.54)	0.0
Vaginal delivery (incl. forceps)	43/59	49/59	0.57 (0.21–1.46)	0.286
Vaginal delivery within 24 h	18/59	26/59	0.53 (0.21–1.26)	0.169
Vaginal delivery within 48 h	31/59	44/59	0.38 (0.15–0.90)	0.024
Cesarean section	16/59	10/59	1.75 (0.69–4.81)	0.286
Postpartum hemorrhage	14/59	5/59	2.80 (0.95–9.93)	0.064
Fetal distress	8/59	4/59	2.33 (0.53–13.98)	0.344
NICU admission (blank = none)	6/59	0/59	inf (1.18–inf)	0.031
Neonatal sepsis (blank = none)	0/59	0/59	inf (nan-inf)	1.0
Neonatal pneumonia (blank = none)	0/59	0/59	inf (nan-inf)	1.0
Induction-to-labor, h (median difference)	19.50	20.92	-1.42 (-2.75–11.08)	0.062
Induction-to-delivery, h (median difference)	26.32	26.67	-0.35 (-5.75–13.93)	0.098
Rupture-to-delivery, h (median difference)	8.23	3.20	5.03 (2.54–4.87)	0.0
Total vaginal examinations (median difference)	6.00	6.00	0.00 (-1.00–1.00)	0.71
Post-rupture vaginal examinations (median difference)	5.00	3.00	2.00 (0.00–2.00)	0.003

Key Secondary Outcome: Intrapartum Fever

In matched pairs, the incidence of highest intrapartum temperature $\geq 38.0^{\circ}\text{C}$ was 5/59 (8.5%) vs. 1/59 (1.7%), with a paired OR of 5.00 (95% CI 0.56–236.49; $P = 0.219$). Unmatched data: 2.3% vs. 9.2% ($P = 0.092$); Firth-adjusted OR was 6.66 (95% CI 1.39–64.69; $P = 0.016$); inverse probability weighting showed 2.0% vs. 9.2%. Matching, Firth correction, weighting, and caliper variants all showed consistent direction, but with only 9 fevers occurring in the entire cohort, estimates were imprecise.

Intrapartum antibiotic use was 71/76 (93.4%) vs. 40/86 (46.5%) ($P < 0.001$).

Intrapartum Fever and Epidural Duration (Matched Exploration)

In the matched sample comprising 9 febrile cases and 36 afebrile controls, index-time comparison anchored at each woman's confirmed labor onset (controls observed until the same elapsed labor time as the case's labor-to-fever interval, tau 5.2–9.6 h) showed that cumulative epidural analgesia duration for febrile cases was median 5.2 h (IQR 4.8–7.0) vs. 3.6 h (2.3–4.8) for controls ($n = 22$; Mann-Whitney $P = 0.006$); within-group paired differences were median + 3.2 h, positive in 7 of 8 comparable pairs (Wilcoxon signed-rank test $P = 0.021$); 11 controls delivered prior to the index time and exited. Total-duration comparison yielded consistent results (6.8 h vs. 4.8 h, $P = 0.005$; paired difference + 3.0 h, positive in 7/9 sets, $P = 0.033$). Median time from analgesia onset to fever occurrence was 5.2 hours, with 8/9 cases occurring ≥ 4 hours later (Supplementary Table I).

Neonatal Safety Domain

Among all 174 neonates, the incidence of sepsis was 0 and pneumonia was 0 (one-sided 95% CI upper limit approximately 1.7%); the incidence of Apgar scores below 7 at 1 or 5 minutes was 0. NICU admission: unmatched was 3/86 (3.5%) vs. 7/76 (9.2%) ($P = 0.192$); matched was 6/59 vs. 0/59 (paired OR infinite, 95% CI 1.18–infinity; $P = 0.031$); exact matching for hypertension was 6/61 vs. 1/61 ($P = 0.063$), with a weakened signal. Neonatal WBC and CRP were measured in approximately 30% of neonates, with similar medians among those measured (descriptive; Table 2).

Final Delivery Mode and Intrapartum Complications (Exploratory Analysis)

In matched pairs, operative delivery (cesarean or forceps) occurred in 24/59 (40.7%) vs. 14/59 (23.7%), with a paired OR of 2.25 (95% CI 0.93–5.98; $P = 0.076$); postpartum hemorrhage rates were 14/59 (23.7%) vs. 5/59 (8.5%), with a paired OR of 2.80 (95% CI 0.95–9.93; $P = 0.064$); spontaneous vaginal delivery rates were 35/59 vs. 45/59 (paired OR 0.44, 95% CI 0.17–1.07; $P = 0.076$); and cesarean section rates were 16/59 vs. 10/59 ($P = 0.286$). Unmatched data: operative delivery rate was 39.5% vs. 23.3% ($P = 0.028$), cesarean section rate was 26.3% vs. 12.8% ($P = 0.044$); median blood loss was 400 ml (IQR 200–500) vs. 300 ml (200–400) ($P = 0.026$), and postpartum hemorrhage rate presented as a binary variable was 19.7% vs. 10.5% ($P = 0.122$). Regarding indications for cesarean section, arrest-spectrum indications (arrest of descent, cephalopelvic disproportion, failed induction) were 11/76 vs. 4/86 ($P = 0.054$) and infection-related indications (chorioamnionitis or fever) were 4/76 vs. 1/86 ($P = 0.187$). There were no significant differences in forceps use (13.2% vs 10.5%, $P = 0.632$), fetal distress (13.2% vs 11.6%, $P = 0.814$), or length of stay (median 6 days for both, $P = 0.553$) (Supplementary Table G).

In the secondary subgroup (those facing the decision), these differences attenuated significantly: operative delivery rate was 28.6% vs. 39.5% at N = 0 (P = 0.211), narrowing to 32.5% vs. 39.5% at N = 6 (P = 0.546); P-values for median blood loss difference ranged from 0.088 to 0.145; P-values for arrest-spectrum cesarean indications ranged from 0.171 to 0.573. After re-matching within the secondary cohort, the paired OR for operative delivery was 1.36 (95% CI 0.59–3.28; N = 0, 50 pairs) and 1.33 (0.52–3.58; N = 2, 52 pairs), the paired OR for postpartum hemorrhage was 2.00 (0.62–7.46) and 1.60 (0.46–6.22), and the paired OR for spontaneous vaginal delivery was 0.73 and 0.75—all not statistically significant. Among the comparison group, only 2 of 20 operative deliveries occurred in rapid stratification (2/23 at N < = 0) (Supplementary Table H). These delivery outcomes are post-hoc exploratory analyses and have not undergone multiplicity correction.

Delivery Timing and Process Outcomes

In matched pairs, induction-to-delivery time did not differ (median 26.3 vs. 26.7 h; median difference – 0.35 h, 95% CI -5.75–13.93; P = 0.098); rupture-to-delivery was longer in the amniotomy pathway (+ 5.03 h, P < 0.0001); post-rupture vaginal examinations were more frequent (median 5 vs 3, + 2; P = 0.003); total examination counts were identical (median 6 in both).

Secondary grouping analysis (Table 4) excluded comparison-group women who entered labor spontaneously shortly after balloon removal: among 86 balloon-pathway women, 23 (26.7%) were already in labor at or before balloon removal, 28 within 2 h, and 46 within 6 h. After excluding these women, who never faced the amniotomy decision:

Table 4
Secondary Grouping Analysis by Exclusion Threshold

Exclusion threshold	Outcome	Pre-labor amniotomy (n = 76)	Waiting pathway	Unmatched OR	Unmatched P	Firth-adjusted OR (95% CI), P
0 h	Clinical chorioamnionitis	6/76 (7.9%)	5/63 (7.9%)	0.99	1.0	1.37 (0.39–5.24), P = 0.627
0 h	Intrapartum fever > = 38.0 C	7/76 (9.2%)	2/63 (3.2%)	3.09	0.183	5.13 (1.06–50.01), P = 0.041
0 h	Vaginal delivery within 24 h	23/76 (30.3%)	17/63 (27.0%)	1.17	0.71	
0 h	NICU admission	7/76 (9.2%)	2/63 (3.2%)	3.09	0.183	
0 h	Induction-to-delivery, h (median difference)	26.92 (21.68–49.02)	29.48 (23.59–36.38)	-2.57 h	0.783	
0 h	Induction-to-delivery median difference (PSM, 50 pairs)	27.26	29.62	-2.36 h	0.487	
2 h	Clinical chorioamnionitis	6/76 (7.9%)	5/58 (8.6%)	0.91	1.0	1.31 (0.37–5.03), P = 0.677
2 h	Intrapartum fever > = 38.0 C	7/76 (9.2%)	2/58 (3.4%)	2.84	0.299	4.97 (1.02–48.60), P = 0.046
2 h	Vaginal delivery within 24 h	23/76 (30.3%)	12/58 (20.7%)	1.66	0.238	
2 h	NICU admission	7/76 (9.2%)	1/58 (1.7%)	5.78	0.137	
2 h	Induction-to-delivery, h (median difference)	26.92 (21.68–49.02)	29.78 (25.25–37.52)	-2.87 h	0.302	
2 h	Induction-to-delivery median difference (PSM, 52 pairs)	27.33	29.78	-2.45 h	0.633	
4 h	Clinical chorioamnionitis	6/76 (7.9%)	5/51 (9.8%)	0.79	0.754	1.13 (0.32–

Exclusion threshold	Outcome	Pre-labor amniotomy (n = 76)	Waiting pathway	Unmatched OR	Unmatched P	Firth-adjusted OR (95% CI), P
						4.38), P = 0.849
4 h	Intrapartum fever ≥ 38.0 C	7/76 (9.2%)	2/51 (3.9%)	2.49	0.313	4.33 (0.89–42.40), P = 0.072
4 h	Vaginal delivery within 24 h	23/76 (30.3%)	6/51 (11.8%)	3.25	0.018	
4 h	NICU admission	7/76 (9.2%)	1/51 (2.0%)	5.07	0.143	
4 h	Induction-to-delivery, h (median difference)	26.92 (21.68–49.02)	31.40 (27.43–41.46)	-4.48 h	0.058	
4 h	Induction-to-delivery median difference (PSM, 49 pairs)	27.28	31.6	-4.32 h	0.511	
6 h	Clinical chorioamnionitis	6/76 (7.9%)	3/40 (7.5%)	1.06	1.0	1.67 (0.39–9.62), P = 0.502
6 h	Intrapartum fever ≥ 38.0 C	7/76 (9.2%)	2/40 (5.0%)	1.93	0.717	3.50 (0.71–34.53), P = 0.132
6 h	Vaginal delivery within 24 h	23/76 (30.3%)	3/40 (7.5%)	5.35	0.005	
6 h	NICU admission	7/76 (9.2%)	0/40 (0.0%)	inf	0.094	
6 h	Induction-to-delivery, h (median difference)	26.92 (21.68–49.02)	33.83 (28.73–48.75)	-6.91 h	0.008	
6 h	Induction-to-delivery median difference (PSM, 38 pairs)	27.26	34.67	-7.41 h	0.296	

- The median induction-to-delivery time lengthened with the threshold: 29.5 h (N = 0), 29.8 h (N = 2), 31.4 h (N = 4), 33.8 h (N = 6), while the amniotomy pathway remained constant at 26.9 h; unmatched median differences moved monotonically from - 2.57 h (P = 0.783) to -6.91 h (P = 0.008), with renewed matching yielding - 2.36, -2.45, -4.32, and - 7.41 h.

- Vaginal delivery within 24 h: the amniotomy pathway was stable at 30.3% (23/76) while the waiting pathway decreased from 27.0% (N = 0) to 7.5% (N = 6); OR rose monotonically from 1.17 (P = 0.710) to 5.35 (P = 0.005).
- Vaginal delivery within 48 h was still more common in the waiting pathway at N = 0 (73.0% vs. 53.9%, P = 0.023), indicating that most women in the waiting pathway also delivered vaginally within 48 h. The amniotomy advantage is concentrated on vaginal delivery within 24 h, rather than overall vaginal delivery rates, which showed no significant difference.
- Chorioamnionitis remained null at every threshold (unmatched OR 0.79–1.06; Firth-adjusted OR 1.13–1.67, all P > 0.5); fever ORs attenuated with the threshold (unmatched 3.09 to 1.93; Firth 5.13 to 3.50) without reversing. All 2 cases of fever and all 5 cases of chorioamnionitis (at N ≤ 4 h) in the comparison group occurred among the waiting women rather than the rapid progressors.

Timing Subgroups (Exploratory Analysis)

Within the amniotomy pathway (n = 76), removal-to-amniotomy intervals were < 2 h in 42 women (median induction-to-delivery 23.2 h; fever 3/42; chorioamnionitis 5/42) and > 4 h in 30 women (52.4 h; fever 4/30; chorioamnionitis 1/30); four women were between 2 and 4 h. Groups defined by interval necessarily experienced different waiting times, so differences existed prior to group assignment; this study does not provide causal interpretation (Additional files).

Sensitivity Analysis

Calipers of 0.1 and 0.3 SD, inverse probability weighting, and exact matching within hypertension strata preserved the direction of the primary and key secondary results (Additional files). In Firth models, including the Bishop missingness indicator, the indicator (three women) showed sparse-data artifacts with OR > 16; complete-Bishop-case models are therefore primary.

Discussion

Among women in our cohort who received balloon cervical ripening, we used propensity-score matching as the primary method to compare maternal infection, neonatal safety, and delivery efficiency between pre-labor amniotomy and the balloon pathway, and addressed the inherent treatment escalation selection bias in this comparison through pre-specified subgroup analyses. Three findings were particularly prominent. First, among women truly facing the amniotomy decision, pre-labor amniotomy was associated with faster delivery: the advantage of completing vaginal delivery within 24 h and the reduction in induction-to-delivery time emerged monotonically as exclusion thresholds tightened, whereas apparent lack of efficiency advantage only existed in full-cohort comparisons containing rapid self-progressors. Second, confirmed infectious outcomes were robustly null: chorioamnionitis did not differ under any population definition or method (matched paired OR 1.33; secondary-grouping ORs 0.79–1.67), and no neonatal sepsis or pneumonia was found across the full cohort. Third, the numerically higher operative delivery and postpartum hemorrhage signals in the fully matched cohort attenuated to non-significance after renewed matching in the decision-facing population (paired ORs 1.33–1.36 and 1.60–2.00), indicating they were

largely driven by favorable delivery trajectories of rapid self-progressors diluting the comparison group. The one unresolved signal is intrapartum fever: consistent in direction across all methods (paired OR 5.00; Firth-adjusted ORs 3.50–6.66), but based on very few events (nine cohort-wide), attenuating with population definition, and a non-specific proxy outcome.

Comparison with the literature. Reported rates of chorioamnionitis after balloon ripening vary widely [2]; direct prospective comparisons of pre-labor amniotomy following balloon ripening are rarely reported. Our chorioamnionitis rate of approximately 5–10% falls within the reported range. The associations of fever with rupture-to-delivery time and post-rupture examination frequency are compatible with the ascending-infection hypothesis [8, 9] and with the established association of prolonged labor with fever [10].

Boundaries of interpretation. Intrapartum fever is the only signal awaiting clarification, framed by three points. First, intrapartum antibiotic use was 93.4% in the amniotomy pathway vs. 46.5%, indicating systematically intensified monitoring and intervention; if antibiotics suppressed fever or diagnosed infection, the observed difference in fever would be underestimated rather than overestimated. Second, NICU admission was more frequent in matched pairs (6/59 vs. 0/59) but attenuated under exact matching, with no confirmed sepsis or pneumonia; the most plausible explanation is the cascade of monitoring following membrane rupture rather than neonatal infection. Third, intrapartum temperature $\geq 38.0^{\circ}\text{C}$ is a proxy outcome whose non-infectious etiologies are particularly prominent in our data (next section). In summary, the fever signal does not negate the overall benefit profile of pre-labor amniotomy for the decision-making population, but should not be ignored; its clinical significance requires prospective adjudication.

Non-infectious explanations for intrapartum fever. Non-infectious causes play an important role in the differential diagnosis of intrapartum fever, most notably fever associated with neuraxial analgesia: randomized data show that epidural analgesia is independently associated with intrapartum maternal fever [11], thought to cause fever via altered thermoregulation and sterile inflammatory responses such as elevated interleukin-6, with risk increasing with duration of analgesia, typically appearing several hours after initiation [12]. In our cohort, nearly all delivering women received epidural analgesia; the above matched exploration showed that at the same point on the labor time axis, women with fever had accumulated epidural exposure approximately 3 hours longer than matched afebrile controls (paired difference + 3.2 h, $P = 0.021$). Since labor timing was aligned, this gradient cannot be explained by longer labor in febrile women, supporting an association between duration of epidural exposure itself and intrapartum fever; additionally, 11 rapid-delivery controls withdrew before their index time point, making this estimate conservative in direction. If analgesia-related fever constitutes a substantial portion of intrapartum fever, then the fever signal in the amniotomy pathway may be predominantly non-infectious rather than reflecting rising infection after rupture—corroborated by the null chorioamnionitis finding across all methods and zero confirmed neonatal infection. Limitations: days inferred from anchors, pause minutes not abstracted (underestimating both arms alike), and only 9 cases—constituting exploratory evidence (Supplementary Table I).

Interpreting the efficiency findings. The null result in the fully matched cohort and the monotonic shortening in secondary analyses are not contradictory: approximately one-quarter of women in the comparison group had already entered labor upon balloon removal, and half did so within 6 h; these women experienced rapid labor progression and never faced the amniotomy decision. The balloon removal operation systematically extended the induction time while leaving the amniotomy pathway unchanged, demonstrating that simple comparisons overestimate the speed of expectant management. Because the secondary grouping is conditioned on post-treatment status (labor status), this constitutes a milestone conditional analysis; indication confounding and guarantee-time components cannot be excluded, and the 48-hour vaginal delivery advantage of the waiting pathway at $N = 0$ indicates that amniotomy's benefit is chiefly faster onset rather than a better final mode of delivery.

Clinical meaning of a shorter induction. For women who did not enter labor after balloon removal, shorter induction time is more than just an efficiency metric. Prolonged induction entails sustained contraction discomfort, restricted mobility from infusions and fetal monitoring, sleep deprivation, and fatigue, potentially increasing feelings of anxiety and loss of control. If the difference in vaginal delivery within 24 h could be prospectively replicated even at the most stringent threshold (30.3% vs. 7.5%), its significance would extend beyond bed turnover to include shortened waiting periods during which women endure discomfort, with potential to improve birth experiences and early postpartum recovery. We note that this study did not measure patient-reported outcomes (pain, fatigue, birth experience, satisfaction), and hospital stays were identical between groups (median 6 days); whether shorter induction translates into measurable recovery benefits requires confirmation through patient-reported outcome studies. This potential benefit must still be weighed against the signal of fever, which continues to rise in the same direction; it must also be balanced against the operative delivery and postpartum hemorrhage signals, which have significantly attenuated among those facing the decision (below).

Exploratory signals in final delivery outcomes and their attenuation. In the full cohort, operative delivery was numerically higher in the pre-labor amniotomy pathway: 39.5% vs. 23.3% unmatched ($P = 0.028$), paired OR 2.25 in matched pairs (95% CI 0.93–5.98; $P = 0.076$); median blood loss was 400 vs. 300 ml ($P = 0.026$), with postpartum hemorrhage paired OR 2.80 (95% CI 0.95–9.93; $P = 0.064$). However, in the secondary subgroup, these differences attenuated significantly: operative delivery rates were 28.6–32.5% vs. 39.5% (all $P > 0.211$), with renewed-matching paired ORs falling to 1.33–1.36; postpartum hemorrhage paired ORs to 1.60–2.00; and the spontaneous vaginal delivery paired OR rising from 0.44 to 0.73–0.75—all without statistical significance. Only 2 of 20 operative deliveries in the comparison group occurred in the rapid layer (2/23 at $N < 0$), indicating that the full-cohort operative delivery and hemorrhage excesses share the same explanation as the null efficiency finding: both outcomes were artificially attenuated by the favorable delivery trajectories of patients with rapid spontaneous labor in the control group, rather than caused by pre-labor amniotomy itself. Arrest-spectrum cesarean indications (11 vs. 4) also lost statistical significance in the decision-facing population ($P = 0.171$ –0.573); indication confounding may have contributed, but infection-related indications (4 vs. 1) remain consistent with the fever signal, so downstream effects of fever or suspected-infection cascades cannot be excluded. Operative delivery and postpartum hemorrhage should therefore be read as weak signals primarily explained by population

definitions; their direction did not reverse in any analysis, and complete exclusion requires larger sample sizes. Forceps rates, fetal distress, and hospital stay did not differ under any population definition.

A person-centered perspective. Pathway selection should also incorporate women's preferences. For women with unfavorable cervical conditions, balloon placement itself is an uncomfortable and uncertain experience; the decision following balloon removal—"should we rupture the membranes now?"—is well-suited for shared decision-making. Our study suggests that for women desiring earlier termination of induction, pre-labor amniotomy offers a faster route to vaginal delivery without clear evidence of increased confirmed infection, dystocia, or hemorrhage; residual uncertainty regarding the fever signal indicates that this decision should remain individualized and thoroughly discussed, balancing women's own preferences for earlier induction termination against minimal intervention. Future research should incorporate patient-reported outcomes (birth experience, sense of control, satisfaction, postpartum fatigue, breastfeeding initiation) to more comprehensively reflect how both pathways influence women's birth experiences.

Limitations. This was a single-center retrospective study with limited sample size (9 cases of fever, 11 cases of chorioamnionitis, 11 NICU admissions), resulting in wide CIs. Thresholds for secondary subgroups were derived from data; despite monotonic gradients and consistent directional agreement across methods, this framework remains an exploratory sensitivity analysis rather than confirmatory evidence. The substantial disparity in intrapartum antibiotic use serves as both a confounder and an inseparable component of the intervention pathway in this design. Intrapartum fever served as a proxy outcome; neonatal laboratory markers were assessed only when clinically indicated. In the exploration of matched epidural duration, days were inferred from anchor points, pause minutes were not extracted, and only 9 cases were available, providing only exploratory evidence for fever differential diagnosis. Multiplicity was not adjusted. Generalizability is limited by single-center practice patterns.

Conclusions

Among women who did not enter labor after balloon removal—those actually facing the choice between amniotomy and expectant management—pre-labor amniotomy was associated with faster delivery and demonstrated a distinct advantage for vaginal delivery within 24 h. Confirmed infection outcomes (chorioamnionitis, neonatal sepsis, and pneumonia) were not increased, and there was no clear evidence of elevated postpartum-hemorrhage-related risk in this population, where operative delivery and hemorrhage signals attenuated markedly. The only signal that was consistent in direction but imprecise is intrapartum fever, whose clinical meaning awaits larger, preferably prospective studies. Based on current evidence, pre-labor amniotomy is a reasonable option for accelerating labor among women who have not entered labor after balloon removal and seek to shorten induction time.

Declarations

Ethics approval and consent to participate:

- This study was approved by the Ethics Committee of China-Japan Friendship Hospital (approval No. 2026-KY-290, 15 July 2026), with a waiver of informed consent granted for the retrospective design. The study

was conducted under the approved protocol “A retrospective study of pregnancy-related factors (e.g., induction methods) and adverse pregnancy outcomes (e.g., infection)”.

Consent for publication:

- Not applicable. This manuscript contains no individually identifiable data or images in any form.

Competing interests:

- The authors declare that they have no competing interests.

Funding:

The authors declare that this study received no specific funding from public, commercial, or non-profit funding agencies.

Author Contribution

Authors’ contributions: Ying Xu: study execution, data curation and statistical analysis, and writing of the original draft. Wenhui Li, Cuiping Feng, Yanting Chen, and Zhenzhen Zhang: data curation and chart review verification. Xiaoyu Pan: study design and supervision, interpretation of results, and manuscript revision. All authors read and approved the final manuscript.

Acknowledgements:

- Not applicable.

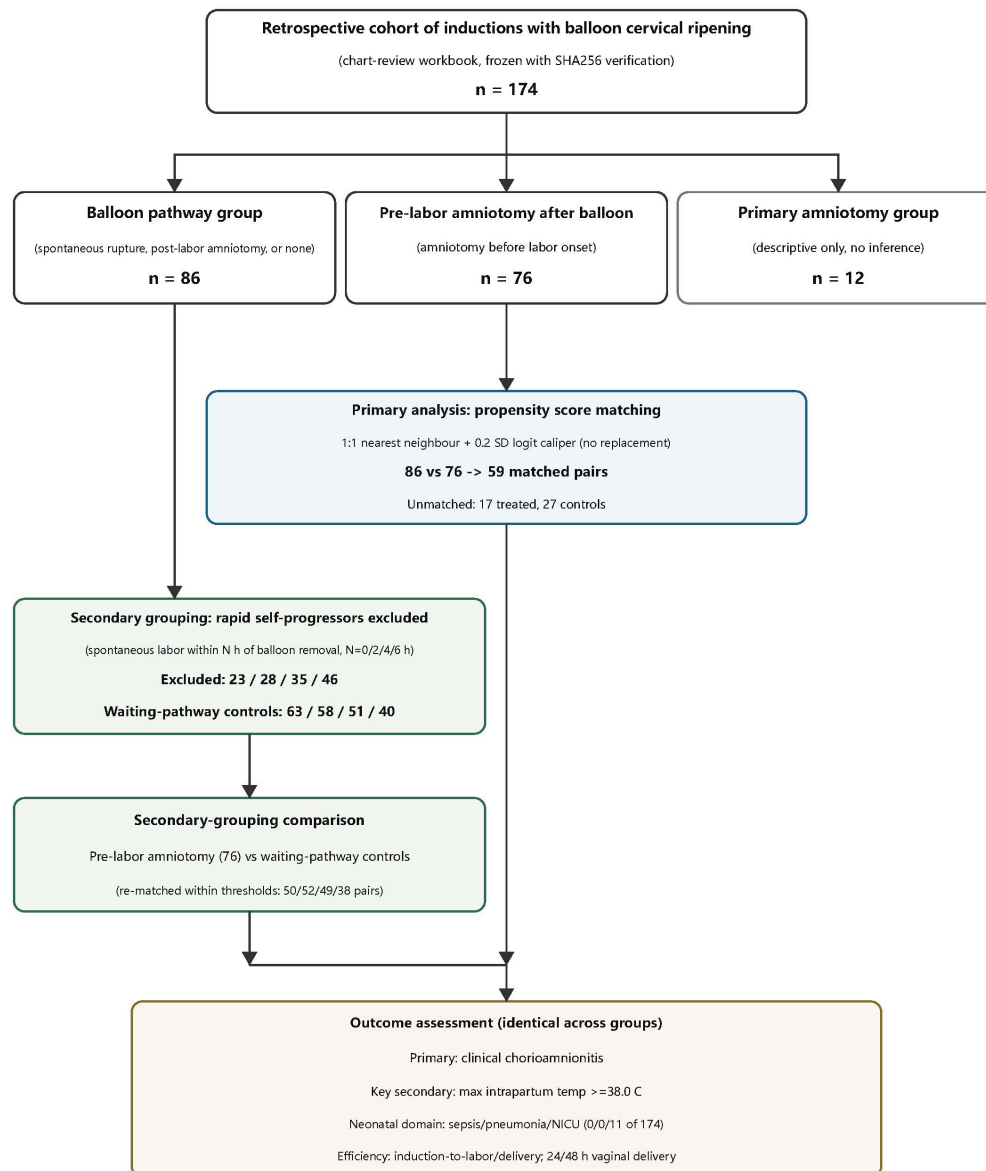
References

1. World Health Organization. WHO recommendations on mechanical methods for induction of labour. Geneva: World Health Organization; 2022.
2. de Vaan MDT, ten Eikelder MLG, Jozwiak M, Palmer KR, Davies-Tuck M, Bloemenkamp KWM, Mol BWJ, Bouvain M. Mechanical methods for induction of labour. Cochrane Database Syst Rev. 2023;2023(3):CD001233.
3. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453–7.
4. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70(1):41–55.

5. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat*. 2011;10(2):150–61.
6. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med*. 2009;28(25):3083–107.
7. Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80(1):27–38.
8. Newton ER, Prihoda TJ, Gibbs RS. Logistic regression analysis of risk factors for intra-amniotic infection. *Obstet Gynecol*. 1989;73(4):571–5.
9. Gluck O, Mizrachi Y, Ganer Herman H, Bar J, Kovo M, Weiner E. The correlation between the number of vaginal examinations during active labor and febrile morbidity, a retrospective cohort study. *BMC Pregnancy Childbirth*. 2020;20(1):246.
10. Maayan-Metzger A, Mazkereth R, Shani A, Kuint J. Risk factors for maternal intrapartum fever and short-term neonatal outcome. *Fetal Pediatr Pathol*. 2006;25(3):169–77.
11. Philip J, Alexander JM, Sharma SK, Leveno KJ, McIntire DD, Wiley J. Epidural analgesia during labor and maternal fever. *Anesthesiology*. 1999;90(5):1271–5.
12. Goetzl L, Evans T, Rivers J, Suresh MS, Lieberman E. Elevated maternal and fetal serum interleukin-6 levels are associated with epidural fever. *Am J Obstet Gynecol*. 2002;187(4):834–8.

Figures

Fig. 1 Cohort flow and primary analysis populations (n=174)



Note: rapid self-progressors were in labor at or before balloon removal (N=0 h includes labor preceding removal); these women never faced the amniotomy-versus-waiting decision.

Figure 1

See image above for figure legend

Supplementary Files

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

- [SUPPORTINGINFORMATION.docx](#)