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.

The Influence of Clinical History on the Accuracy of Fetal Weight Estimation at Term: A Prospective Study

Daniel Malchi
Emek Medical Center
Noga Shachor
Emek Medical Center
Nitzan Dana Sela
Emek Medical Center
Khadeje Seh
Emek Medical Center
Ido Izhaki
University of Haifa
Gali Garmi
Technion-Israel Institute of Technology
Noah Zafran
Technion-Israel Institute of Technology
Adi Dayan-Schwartz
adida@technion.ac.il

Technion-Israel Institute of Technology

Research Article

Keywords: Fetal weight estimation, Obstetric history, ultrasonography, Leopold maneuvers, birthweight, Estimation accuracy

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11389189/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

Objective

To assess whether access to obstetric and medical history affects the accuracy of clinical and ultrasonographic estimated fetal weight (EFW) at term.

Methods

We conducted a prospective paired observational study of singleton pregnancies at 37 + 0 to 42 + 0 weeks. A routine examiner reviewed the complete clinical history and performed clinical and ultrasonographic EFW. A second examiner, blinded to the history, exact gestational age, prior estimates, and patient-provided information, performed both assessments and then repeated them after reviewing the history. Accuracy was quantified using absolute percentage error (APE) relative to measured birthweight. The primary outcome was the within-participant difference in ultrasound APE between the blinded and routine examiners.

Results

Of 183 enrolled participants, 158 had complete data for all six assessments. Mean birthweight was 3,465 ± 413 g. Mean ultrasound APE was 6.11% for the blinded examiner and 6.00% for the routine examiner; the mean paired difference was 0.11 percentage points (95% bootstrap CI − 0.84 to 0.96; P=.808). Mean clinical APE was 8.02% for the blinded examiner and 6.54% for the routine examiner (difference 1.48 percentage points, 95% CI 0.47–2.68; P=.007). In exploratory within-examiner analyses, APE decreased after history review by 0.97 percentage points for clinical estimation and 0.58 percentage points for ultrasonography. Among infants weighing at least 4,000 g, the blinded-versus-routine ultrasound difference was 1.30 percentage points (95% CI − 0.73 to 3.63; P=.275).

Conclusion

In term singleton pregnancies, access to clinical history did not significantly influence ultrasound-based EFW accuracy. Differences observed in clinical and exploratory within-examiner analyses were small and unlikely to be clinically meaningful.

What does this study adds to the clinical work:

Ultrasound fetal weight estimation at term showed no statistically significant difference in accuracy between blinded and routine examiners, with an observed difference in mean absolute error of only 4.4 g. These findings suggest limited additional benefit of clinical history for ultrasound estimation accuracy and support the relative objectivity of the examination.

Introduction

Estimated fetal weight (EFW) informs obstetric counseling and delivery planning, particularly when fetal growth restriction or macrosomia is suspected[1, 2]. Neither abdominal palpation nor ultrasonographic biometry measures fetal weight directly, and substantial estimation errors remain common[3–5]. Whether clinical history materially changes estimation accuracy is therefore relevant to the interpretation of both examinations.

Maternal characteristics, gestational age, prior neonatal birthweights, and maternal impressions of fetal size may provide useful context but may also influence examiner expectations. Moreover, the growing burden of maternal obesity and diabetes further limits the performance of ultrasound-based EFW, as both conditions are associated with decreased accuracy of the estimate [6]. When comparing clinical with sonographic EFW, sonography shows a clear accuracy advantage at birthweight extremes, whereas differences between the two methods are minimal across most other pregnancies[7]. Sonographic EFW is not entirely objective; a study of expected-value bias during ultrasound biometry showed that exposure to on-screen biometric information could influence operators to adjust measurements toward expected gestational-age values[8]. These findings show differences in accuracy, but they do not show how much history changes estimates when the same examiner repeats the assessment.

In routine practice, EFW are almost always performed after reviewing the patient's obstetric and medical history. Knowledge of prior neonatal birthweights and maternal risk factors may influence clinicians' expectations, potentially biasing the fetal weight estimation.

In this prospective study, we assessed how access to patients' obstetric and medical history affects the accuracy of clinical and sonographic EFW.

Methods

Study design and participants

We conducted a prospective paired observational study in the Department of Obstetrics and Gynecology of the Emek Medical Center, between March 2018 and January 2022. Institutional review board approval was obtained in February 2018 (EMC-0119-17), and all participants provided written informed consent.

The study was conducted in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Supplementary STROBE checklist).

Eligible participants were adults with singleton pregnancies at 37 + 0 to 42 + 0 weeks. Exclusions included multiple pregnancy, known uterine or fetal anomaly, oligohydramnios (amniotic fluid index < 5 cm), polyhydramnios (> 25 cm), non-reassuring fetal status, maternal hemodynamic instability, or inability to consent.

After enrollment, participants were excluded for delivery at another institution, an assessment-to-delivery interval of 14 days or more, or subsequently identified ineligibility. Included participants had birthweight and all six EFW measurements recorded.

After consent, a routine obstetrician (Physician A) reviewed the complete medical and obstetric history and performed clinical EFW using Leopold maneuvers, followed by ultrasonographic EFW. A second obstetrician (Physician B) then performed the same two assessments without access to the history, the exact gestational age, the first physician's estimates, or information supplied by the patient. Before the second examiner entered the room, prior ultrasound estimates were removed from the machine display.

Physician B then reviewed the history obtained by Physician A and repeated palpation followed by ultrasonography.

Each participant contributed six estimates: routine clinical and ultrasound, blinded clinical and ultrasound, and post-history clinical and ultrasound. Research assessments did not replace routine care.

All ultrasound examinations were performed using a GE Voluson P6 system and a standard transabdominal probe. The machine-generated EFW was based on the Hadlock 2 formula[9]. Physicians were categorized according to their recorded experience group: fewer than 3 years of residency, 3–6 years of residency, or consultant-level experience.

The primary outcome was ultrasound estimation accuracy between the blinded examiner and the routine examiner.

Accuracy was quantified using absolute percentage error (APE), calculated as |EFW − birthweight| / birthweight × 100; lower APE indicates greater accuracy.

The primary contrast was blinded minus routine ultrasound APE.

The corresponding between-examiner clinical comparison was secondary. All within-examiner comparisons, including clinical and ultrasound estimates before and after history review, were exploratory because repeated assessment by the same clinician is susceptible to recall, anchoring, and fixed-order bias.

Additional descriptive outcomes included estimates within ± 10% and ± 15% of birthweight, differences across all six assessment conditions, composites, physician experience, and maternal impressions compared to previous birth weight. A composite was the of clinical and ultrasound EFW within a condition, with APE calculated against birthweight.

Birthweight extremes analyses used bands of ≤ 2,500, 2,501–3,999, 4,000–4,499, and ≥ 4,500 g. Given sparse extremes, the high-birthweight analysis included all infants ≥ 4,000 g. An additional sensitivity analysis pooled infants ≤ 2,500 or ≥ 4,000 g as combined birthweight extremes. Both analyses used 2,501–3,999 g as the reference group.

Maternal weight gain was delivery minus pre-pregnancy weight.

Maternal impressions of “larger,” “similar,” or “smaller” were compared with the recorded previous neonatal birthweight. Percentage change was (index − previous birthweight) / previous birthweight × 100. Changes from − 10% to + 10%, inclusive, were similar; values above + 10% or below − 10% were larger or smaller. Both an interpretable response and previous birthweight were required.

Statistical analysis

Continuous variables are presented as mean ± standard deviation (SD) or median [interquartile range (IQR)], and categorical variables as n (%).

The primary ultrasound contrast estimated the participant-level paired difference in APE between the blinded and routine examiners, reported in percentage points with a 95% bias-corrected and accelerated bootstrap confidence interval (CI; 40,000 resamples). Paired comparisons were assessed using sign-flip permutation tests based on 300,000 Monte Carlo resamples. The same methods were used for secondary and exploratory paired contrasts.

The single primary ultrasound comparison was evaluated using an unadjusted two-sided P value with $\alpha = 0.05$. The secondary between-examiner clinical comparison is reported with a nominal, unadjusted P value. Holm adjustment was applied separately to the two exploratory clinical/ultrasound before–after contrasts and to the two exploratory composite contrasts, controlling error within each exploratory family. Bootstrap CIs are pointwise and unadjusted. Examiner experience distributions were compared using the Stuart–Maxwell test for paired marginal homogeneity.

Birthweight subgroup analyses used exact paired sign-flip tests and pointwise BCa bootstrap CIs based on 40,000 resamples. Effect modification was assessed by comparing participant-level APE contrasts between birthweight groups using Welch tests.

For continuous and binary predictors Spearman correlation, Mann–Whitney U, or Kruskal–Wallis tests were used as appropriate, with Benjamini–Hochberg false-discovery-rate (FDR) adjustment across all tests. Maternal directional association and categorical agreement were assessed using Spearman correlation and unweighted Cohen’s kappa, respectively.

Analyses were performed in Python using SciPy, statsmodels, NumPy, and pandas.

Sample-size

The target mean paired difference in absolute percentage error (APE) was 2. Assuming a standard deviation of the paired differences of 7.5[10] percentage points, 90% power, a two-sided α of 0.05, and allowing for 20% loss, the recruitment target was 183 women.

Results

Of 183 enrolled participants, 25 were excluded because delivery occurred elsewhere (n = 11), the assessment-to-delivery interval was above 14 days (n = 5), a fetal anomaly was identified (n = 5), polyhydramnios was present (n = 3), or oligohydramnios was present (n = 1). The analytic cohort included 158 participants with complete data for all six original assessments (Fig. 1).

Mean maternal age was 29.8 ± 5.3 years and pre-pregnancy body mass index (BMI) was 25.0 ± 4.9 kg/m². Gestational age averaged 39.7 ± 1.0 weeks at assessment and 40.2 ± 1.1 weeks at delivery; the median interval was 2 [1–5] days. Mean birthweight was 3,465 ± 413 g; one infant weighed < 2,500 g and 16 weighed ≥ 4,000 g. Confirmed diabetes was recorded in 15/146 interpretable records (10.3%) (Table 1).

Table 1
Participant, neonatal, and examiner characteristics

Characteristic	Value
Maternal and obstetric characteristics	
Maternal age, years	29.8 ± 5.3
Pre-pregnancy body mass index, kg/m²	25.0 ± 4.9
Maternal height, cm	162.7 ± 6.2
Recorded parity, median [IQR]	2 [1–3]
Gestational weight gain, kg	13.3 ± 6.1
Gestational age at assessment, weeks	39.7 ± 1.0
Gestational age at delivery, weeks	40.2 ± 1.1
Assessment-to-delivery interval, days	2 [1–5]
Maternal comorbidity	45 (28.5%)
Pregnancy complication	26 (16.5%)
Previous fetal growth restriction	14 (8.9%)
Mean previous neonatal birthweight, g	3212 ± 410
Diabetes status and recorded ethnicity	
Confirmed gestational/pregestational diabetes (n = 146)	15 (10.3%)
Diabetes screening not performed	10 (6.3%)
Jewish ethnicity	106 (67.1%)
Arab ethnicity	50 (31.6%)
Other recorded ethnicity	2 (1.3%)
Neonatal and delivery characteristics	
Neonatal birthweight, g	3465 ± 413
Birthweight < 2,500 g	1 (0.6%)
Birthweight ≥ 4,000 g	16 (10.1%)
Male neonate	88 (55.7%)
Spontaneous vaginal delivery	127 (80.4%)

Values are mean ± standard deviation, median [interquartile range], or n (%).

Characteristic	Value
Maternal and obstetric characteristics	
Vacuum-assisted vaginal delivery	11 (7.0%)
Cesarean delivery, including failed vacuum	20 (12.7%)
Shoulder dystocia	1 (0.6%)
Physician experience	
Routine examiner: <3 years	85 (53.8%)
Routine examiner: 3–6 years	28 (17.7%)
Routine examiner: consultant	45 (28.5%)
Second examiner: <3 years	61 (38.6%)
Second examiner: 3–6 years	38 (24.1%)
Second examiner: consultant	59 (37.3%)
Values are mean ± standard deviation, median [interquartile range], or n (%).	

Experience distributions differed between examiners: residents with < 3 years performed 85/158 Physician A assessments (53.8%) and 61/158 Physician B assessments (38.6%); consultants performed 45 (28.5%) and 59 (37.3%), respectively (Stuart–Maxwell P = 0.026).

Mean ultrasound APE was 6.11% for the blinded examiner and 6.00% for the routine examiner. The primary blinded-minus-routine difference was 0.11 percentage points (95% bootstrap CI, − 0.84 to 0.96; two-sided permutation P = 0.808) (Table 2).

Table 2
Paired differences in estimation accuracy

Paired contrast	n	Mean ΔAPE [95% CI], percentage points	P Value	Holm-adjusted P
Primary ultrasound comparison				
Ultrasound: blinded − routine	158	0.11 [-0.84, 0.96]	0.808	—
Secondary clinical comparison				
Clinical: blinded − routine	158	1.48 [0.47, 2.68]	0.007	—
Exploratory within-examiner comparisons				
Clinical: after − before history	158	-0.97 [-2.18, -0.52]	< 0.001	< 0.001
Ultrasound: after − before history	158	-0.58 [-1.09, -0.09]	0.024	0.024
Exploratory composite comparisons				
Composite: after − before history	158	-0.68 [-1.10, -0.32]	< 0.001	0.001
Composite: blinded − routine	158	0.63 [-0.23, 1.42]	0.140	0.140

Differences are within-participant mean changes in absolute percentage error (APE), in percentage points. Negative after-minus-before values indicate improved accuracy. Intervals are pointwise 95% bias-corrected and accelerated bootstrap CIs (40,000 resamples), without multiplicity adjustment.

Holm adjustment applies only within the two exploratory before–after contrasts and, separately, the two exploratory composite contrasts.

Mean clinical APE was 8.02% for the blinded examiner and 6.54% for the routine examiner (difference, 1.48 percentage points; 95% bootstrap CI, 0.47–2.68; nominal permutation P = 0.007) (Table 2).

Across all six assessment conditions, an exploratory Friedman test showed variation in APE ($\chi^2 = 35.30$; $P < 0.001$). Mean APE was 6.54% for routine clinical, 6.00% for routine ultrasound, 8.02% for blinded clinical, 6.11% for blinded ultrasound, 7.05% for post-history clinical, and 5.53% for post-history ultrasound assessment (Table 3; Fig. 2). The post-history ultrasound ranking is exploratory and does not alter the nonsignificant primary between-examiner ultrasound result.

Table 3
Accuracy of fetal weight estimation by assessment condition

Assessment	Mean APE ± SD, %	Median APE [IQR], %	Within ± 10%	Within ± 15%
Routine examiner; history available				
Clinical palpation	6.54 ± 4.88	5.73 [2.84–9.89]	120 (75.9%)	151 (95.6%)
Ultrasonography	6.00 ± 4.75	5.28 [2.33–8.42]	133 (84.2%)	152 (96.2%)
Clinical–ultrasound composite	6.07 ± 4.63	5.29 [2.55–8.58]	128 (81.0%)	154 (97.5%)
Second examiner; before history				
Clinical palpation	8.02 ± 6.95	6.76 [3.15–10.67]	115 (72.8%)	135 (85.4%)
Ultrasonography	6.11 ± 4.66	5.28 [2.15–9.29]	124 (78.5%)	151 (95.6%)
Clinical–ultrasound composite	6.70 ± 5.05	5.69 [2.72–9.33]	121 (76.6%)	146 (92.4%)
Second examiner; after history				
Clinical palpation	7.05 ± 5.23	6.37 [2.93–9.31]	127 (80.4%)	142 (89.9%)
Ultrasonography	5.53 ± 4.27	5.09 [2.13–7.77]	134 (84.8%)	151 (95.6%)
Clinical–ultrasound composite	6.02 ± 4.40	5.07 [2.63–8.01]	130 (82.3%)	149 (94.3%)

All rows are based on 158 paired participants. Threshold columns indicate estimates within the stated percentage of actual birthweight.

APE = |estimated fetal weight − measured birthweight| / measured birthweight × 100.

SD, standard deviation; IQR, interquartile range.

A composite is the arithmetic mean of clinical and ultrasound estimates, with its APE recalculated against birthweight.

Comparing the routine examiner with the second examiner before history, accuracy within ± 10% was 120/158 (75.9%) versus 115/158 (72.8%) for clinical palpation, 133/158 (84.2%) versus 124/158 (78.5%) for ultrasonography, and 128/158 (81.0%) versus 121/158 (76.6%) for the clinical–ultrasound composite. At ± 15%, the corresponding counts were 151/158 (95.6%) versus 135/158 (85.4%), 152/158 (96.2%) versus 151/158 (95.6%), and 154/158 (97.5%) versus 146/158 (92.4%), respectively (Table 2).

Exploratory within-examiner comparisons

In exploratory analyses within Physician B, clinical APE decreased from 8.02% to 7.05% (mean change, − 0.97 percentage points; 95% bootstrap CI, − 2.18 to − 0.52; Holm-adjusted P < 0.001), and ultrasound APE from 6.11% to 5.53% (− 0.58 percentage points; 95% CI, − 1.09 to − 0.09; adjusted P = 0.024) (Table 2). Mean absolute error decreased by 34.7 g clinically and 20.7 g ultrasonographically. Despite the small P values, these findings remain exploratory because the same clinician performed both estimates in a fixed sequence.

Exploratory analyses by birthweight

The ≤ 2,500-g and ≥ 4,500-g strata contained one participant each; 15 weighed 4,000–4,499 g (Supplementary Table S1). Among all 16 infants ≥ 4,000 g, mean ultrasound APE was 6.64% for blinded versus 5.34% for routine assessment (difference, 1.30 percentage points; 95% CI, − 0.73 to 3.63; exact P = 0.275). The corresponding difference was − 0.04 points among 141 infants weighing 2,501–3,999 g (interaction P = 0.296). No subgroup or interaction test remained significant after Holm adjustment (Supplementary Table S2).

Pooling the low- and high-birthweight groups yielded 17 infants with combined extremes versus 141 in the reference group. Mean ultrasound APE was 6.93% for blinded versus 5.54% for routine assessment (difference, 1.38 percentage points; 95% CI, − 0.53 to 3.57; exact P = 0.220). The contrast differed from that in the reference group by 1.42 points (95% CI, − 1.04 to 3.88; interaction P = 0.244). No combined-extremes contrast or interaction remained significant after correction (Supplementary Table S3).

Physician experience and predictors

Patient analysis found no association between examiner experience and APE (Supplementary Table S4).

Eight variables showed nominal associations with APE in exploratory univariable analyses; however, none remained statistically significant after false discovery rate (FDR) correction for multiple comparisons (Supplementary Table S5). The strongest nominal association was observed between pregnancy weight gain and blinded clinical APE, with higher corrected weight gain associated with lower APE (Spearman ρ=−0.237; n = 155; P = 0.003). However, this association did not remain statistically significant after false discovery rate correction (q = 0.289).

Post-history clinical weight gain, birthweight, parity, height, and previous birthweight associations also failed FDR correction.

Maternal impression of fetal size

Of 78 directional maternal responses, 77 had a recorded previous birthweight. Maternal impression correlated with actual percentage change (ρ = 0.472; P < 0.001), but exact categorical agreement was 43/77 (55.8%; κ = 0.302) using the ± 10% threshold (Supplementary Table S6).

Discussion

In this prospective study, we evaluated the accuracy of EFW by obstetricians using both clinical examination and ultrasound. Assessments were performed by routine examiners with access to patients' medical and obstetric histories and by independent examiners blinded to this information. This design allowed us to assess the influence of clinical information on estimation accuracy.

Our primary finding was that ultrasound estimation accuracy did not differ significantly between routine and blinded examiners. Although clinical estimation was associated with significantly greater error among blinded examiners, the absolute difference was small and unlikely to be clinically meaningful.

Interpretation and previous evidence

Our findings suggest that access to maternal medical and obstetric history offers no clinically meaningful advantage for ultrasound fetal weight estimation at term in this cohort. This supports the relative objectivity of sonographic estimation. Beyond the absence of a statistically significant difference, the observed effect was negligible in absolute magnitude: mean APE differed by only 0.11 percentage points, and the difference in mean absolute error was 4.4 g (209.0 g for the blinded examiner versus 204.6 g for the routine examiner).

Although the secondary clinical comparison was statistically significant, the difference in mean absolute error was also modest, at 53.5 g. These small observed differences suggest limited clinical relevance.

For within-examiner analyses, memory of the first estimate, anchoring, repeated measurement, and the fixed sequence create a high potential for bias. A statistically significant before–after change cannot distinguish an effect of history from these influences.

In addition, although nominally significant, these differences were small and unlikely to be clinically meaningful, with exploratory reductions in mean absolute error of only approximately 35 g clinically and 21 g ultrasonographically.

Previous studies report variable relative accuracy of clinical and ultrasound EFW, with dependence on case mix, formula, timing, and technique[3, 4, 11, 12]. Heresco and colleagues found the clearest sonographic advantage at birthweight extremes in 14,400 term pregnancies, while most other differences were small [7]. That study compared modalities rather than access to history. Our ≤ 2,500 and ≥ 4,000-g group analysis found no statistically detectable modification of history-related accuracy differences. However, only 17 infants met that threshold, leaving clinically important differences unresolved.

Ultrasonography derives EFW from biometric measurements, but acquisition and caliper placement remain operator dependent. Expected-value bias during growth scans has been directly documented[8].

Reported associations of EFW accuracy with habitus, diabetes, fetal size, and examiner experience vary across populations[5, 6, 13–17]. Exploratory analyses identified several nominal associations with APE, including gestational weight gain, birthweight, parity, and maternal height, some of which are consistent with previous reports linking EFW accuracy to maternal and fetal characteristics. However, these associations were inconsistent and none survived FDR correction. Patient-clustered modeling also showed no association between physician seniority and APE. Thus, our findings do not identify consistent patient- or examiner-related determinants of EFW accuracy, while not excluding smaller effects in larger cohorts.

Maternal impressions carried directional information but showed only modest categorical agreement. They may provide context, although these data do not support their use as an independent replacement for clinical or ultrasound assessment.

Strengths and limitations

Strengths of this study include its prospective enrollment, complete availability of all six original estimates for every analyzed participant, a common birthweight reference for all assessments, and a primary comparison between two examiners, with the second examiner initially blinded to clinical history.

Several limitations should be considered. In the secondary within-examiner analysis, the second physician performed the post-history assessment only after completing the blinded assessment; therefore, the effects of history review cannot be separated from repetition, learning, memory, anchoring, or changes in measurement technique. In addition, the direct routine-versus-blinded comparisons involved different examiners, whose experience distributions differed, although physician seniority was not associated with APE in the analysis. Finally, the cohort included very few infants at the extremes of birthweight, and the single-center term singleton population excluded pregnancies with abnormal amniotic fluid volume or fetal anomalies, which may limit generalizability.

Clinical implications

The absence of a statistically significant difference between routine and blinded ultrasound assessments suggests that ultrasound-based EFW is relatively independent of clinical history, supporting its objectivity in clinical practice. Although some secondary and exploratory comparisons demonstrated statistically significant differences, their small absolute magnitude raises questions about their clinical relevance. These findings suggest that access to clinical history offers limited additional value for improving EFW accuracy, particularly when ultrasound is used.

Conclusions

In term singleton pregnancies, access to clinical history did not significantly influence ultrasound-based EFW accuracy. While modest differences were observed in clinical estimation and exploratory within-examiner comparisons, these were unlikely to be clinically meaningful. Overall, our findings support

ultrasound as a relatively objective method of fetal weight estimation, while acknowledging the potential for operator-dependent measurement variability.

Declarations

The authors report no conflict of interest.

Funding: none

IRB: Emek Medical Center (EMC-0119-17, approved February 2018)

Data availability: Access to participant-level data is subject to applicable institutional, ethical, and privacy requirements and may be discussed with the corresponding author.

Author contribution

All authors contributed to the study design (**Conceptualization and Methodology**). D.M., N.D.S., K.S., G.G., and N.Z. contributed to **Investigation**. A.D.S. was responsible for **Data curation**. I.I. performed the **Formal analysis**. A.D.S., D.M., and N.D.S. contributed to **Writing – original draft**. N.Z. contributed to **Writing – review & editing**.

References

1. (2020) Macrosomia: ACOG Practice Bulletin, Number 216. Obstet Gynecol 135:e18–e35. https://doi.org/10.1097/AOG.0000000000003606
2. (2021) Fetal Growth Restriction: ACOG Practice Bulletin, Number 227. Obstet Gynecol 137:e16–e28. https://doi.org/10.1097/AOG.0000000000004251
3. Preyer O, Husslein H, Concin N, et al (2019) Fetal weight estimation at term - ultrasound versus clinical examination with Leopold's manoeuvres: a prospective blinded observational study. BMC Pregnancy Childbirth 19:122. https://doi.org/10.1186/s12884-019-2251-5
4. Goto E (2017) Comparing the accuracy of maternal, clinical, and ultrasound estimations to predict birthweight: a meta-analysis. Acta Obstet Gynecol Scand 96:1289–1299. https://doi.org/10.1111/aogs.13208
5. Milner J, Arezina J (2018) The accuracy of ultrasound estimation of fetal weight in comparison to birth weight: A systematic review. Ultrasound Leeds Engl 26:32–41. https://doi.org/10.1177/1742271X17732807
6. Dittkrist L, Vetterlein J, Henrich W, et al (2022) Percent error of ultrasound examination to estimate fetal weight at term in different categories of birth weight with focus on maternal diabetes and obesity. BMC Pregnancy Childbirth 22:241. https://doi.org/10.1186/s12884-022-04519-z
7. Heresco L, Yehuda CB, Cohen Rappaport R, et al (2026) The role of sonographic estimated fetal weight compared to clinical assessment: superior accuracy only in extreme fetal weight cases. BMC

Pregnancy Childbirth 26:504. https://doi.org/10.1186/s12884-026-08871-2

8. Drukker L, Droste R, Chatelain P, et al (2020) Expected-value bias in routine third-trimester growth scans. Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol 55:375–382. https://doi.org/10.1002/uog.21929
9. Hadlock FP, Harrist RB, Carpenter RJ, et al (1984) Sonographic estimation of fetal weight. The value of femur length in addition to head and abdomen measurements. Radiology 150:535–540. https://doi.org/10.1148/radiology.150.2.6691115
10. Njoku C, Emechebe C, Odusolu P, et al (2014) Determination of Accuracy of Fetal Weight Using Ultrasound and Clinical Fetal Weight Estimations in Calabar South, South Nigeria. Int Sch Res Not 2014:970973. https://doi.org/10.1155/2014/970973
11. Dudley NJ (2005) A systematic review of the ultrasound estimation of fetal weight. Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol 25:80–89. https://doi.org/10.1002/uog.1751
12. Lindström L, Cnattingius S, Axelsson O, Granfors M (2023) Accuracy and precision of sonographic fetal weight estimation in Sweden. Acta Obstet Gynecol Scand 102:699–707. https://doi.org/10.1111/aogs.14554
13. Kesrouani A, Atallah C, AbouJaoude R, et al (2017) Accuracy of clinical fetal weight estimation by Midwives. BMC Pregnancy Childbirth 17:59. https://doi.org/10.1186/s12884-017-1242-7
14. Mgbafulu CC, Ajah LO, Umeora OUJ, et al (2019) Estimation of fetal weight: a comparison of clinical and sonographic methods. J Obstet Gynaecol J Inst Obstet Gynaecol 39:639–646. https://doi.org/10.1080/01443615.2019.1571567
15. Peregrine E, O'Brien P, Jauniaux E (2007) Clinical and ultrasound estimation of birth weight prior to induction of labor at term. Ultrasound Obstet Gynecol Off J Int Soc Ultrasound Obstet Gynecol 29:304–309. https://doi.org/10.1002/uog.3949
16. Siemer J, Egger N, Hart N, et al (2008) Fetal weight estimation by ultrasound: comparison of 11 different formulae and examiners with differing skill levels. Ultraschall Med 29:159–164. https://doi.org/10.1055/s-2007-963165
17. Daykan Y, Shavit M, Yagur Y, et al (2022) Fetal weight estimation in tall women: is ultrasound more accurate than clinical assessment? A prospective trial. Arch Gynecol Obstet 305:567–572. https://doi.org/10.1007/s00404-021-06177-6

Figures

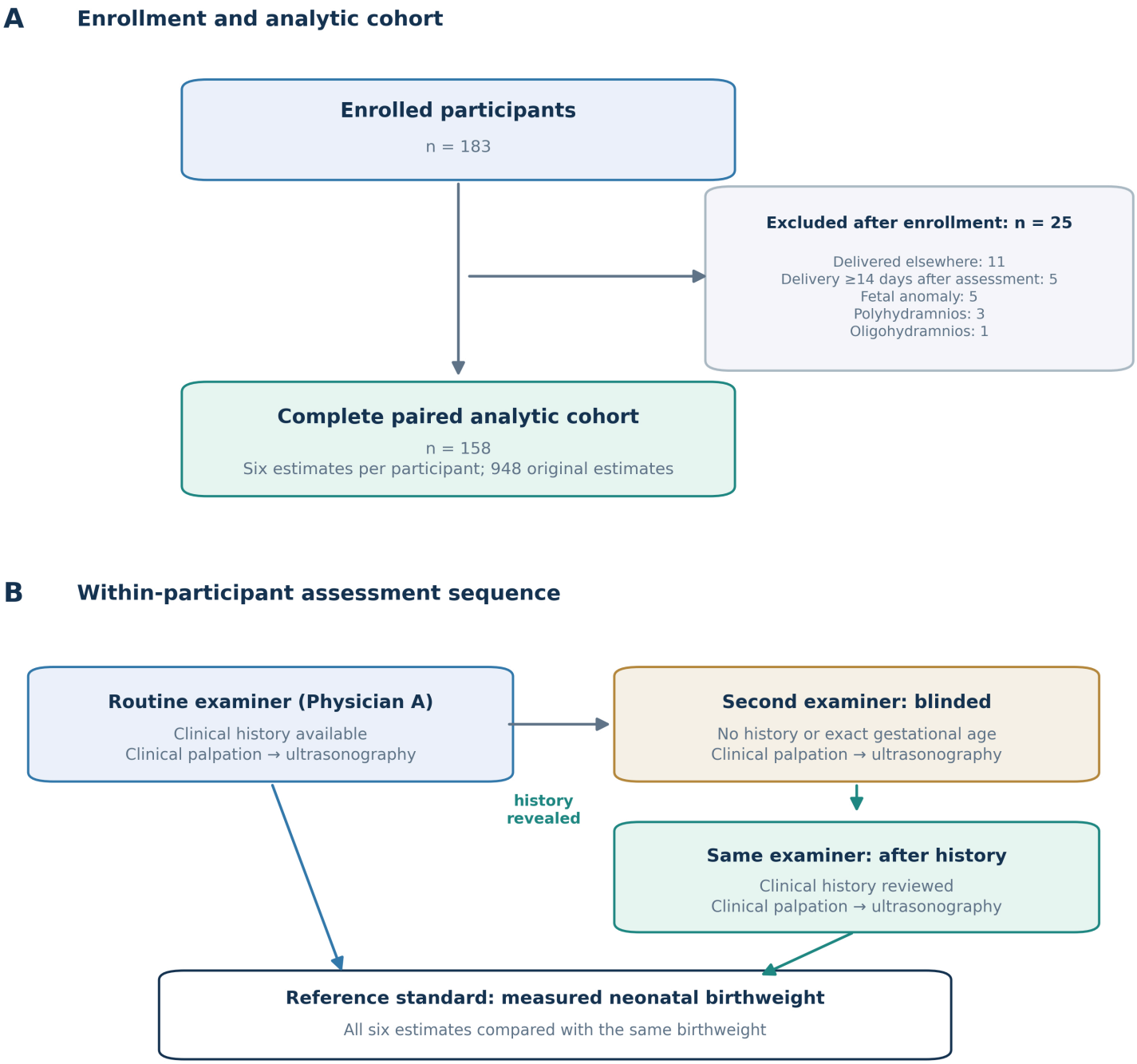


Figure 1

Participant flow and assessment protocol

The primary comparison is routine ultrasound by Physician A versus blinded ultrasound by Physician B. All within-examiner before–after contrasts are exploratory because of recall, anchoring, and fixed-order bias.

Neonatal birthweight is the reference standard.

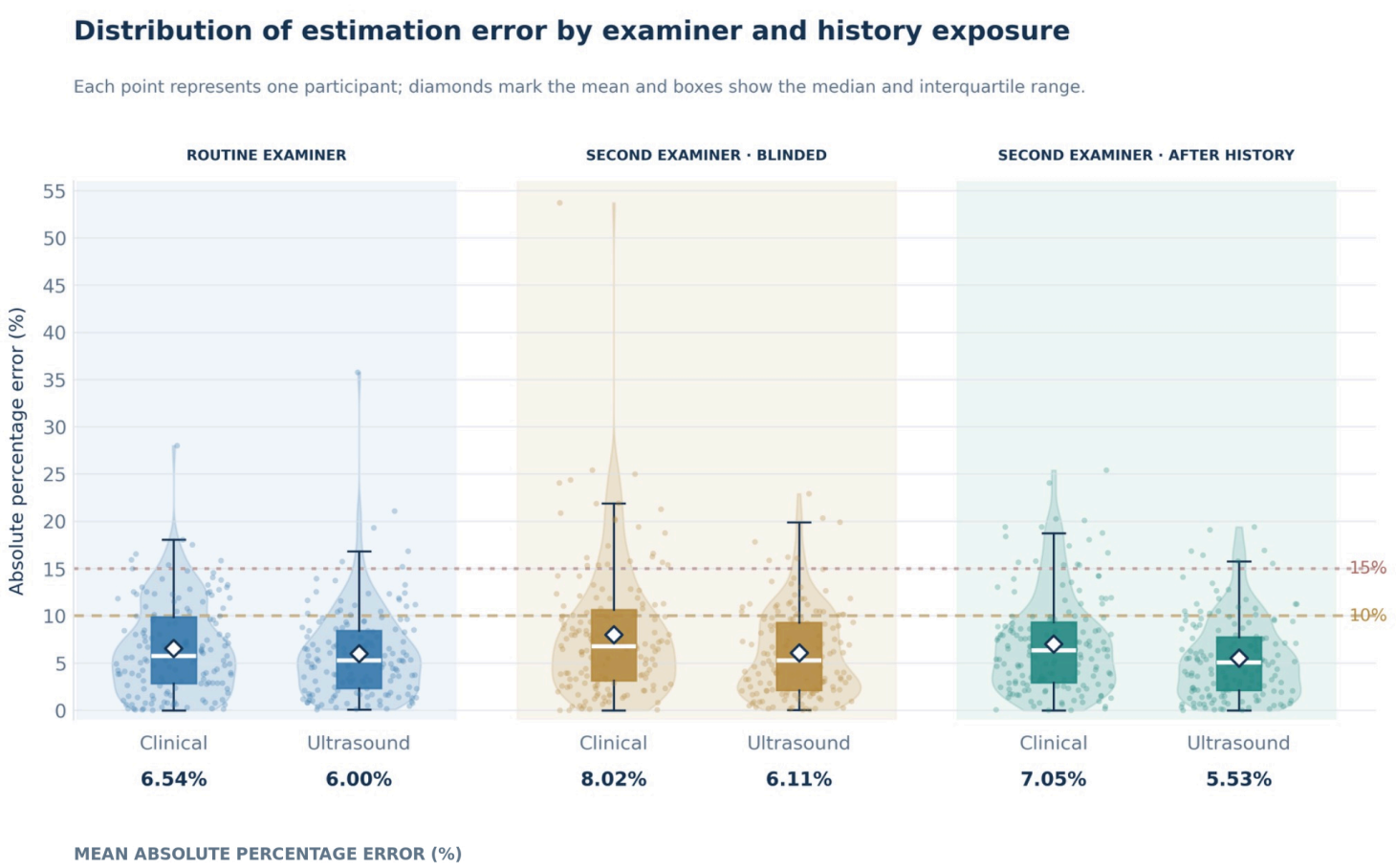


Figure 2

Distribution of absolute percentage error by estimation method

Violin plots, box plots, and participant points describe APE for the six assessment conditions. Diamonds indicate means; box centers and spans indicate medians and interquartile ranges. Dashed lines mark 10% and 15% error.

Supplementary Files

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

- FetalWeightSupplementary.docx