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.

Accuracy of Fetal Gender Identification Using Genital Tubercle Angle and Anogenital Distance in First-Trimester Ultrasound

Pimchanoke Puengsema
King Chulalongkorn Memorial Hospital
Wasinee Tianthong
King Chulalongkorn Memorial Hospital
Dhirapatara Charoenvidhya
King Chulalongkorn Memorial Hospital
Patcharaporn Siricharoenthai
Patcharaporn.si@chula.ac.th
King Chulalongkorn Memorial Hospital

Research Article

Keywords: Anogenital distance, Genital tubercle angle, First-trimester ultrasound, Fetal sex, Crown–rump length

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11666136/v1

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

Additional Declarations: No competing interests reported.

Accuracy of Fetal Gender Identification Using Genital Tubercle Angle and Anogenital Distance in First-Trimester Ultrasound

Authors and Affiliations

Pimchanoke Puengsema, M.D.;

Department of Obstetrics and Gynecology, King Chulalongkorn Memorial Hospital, Bangkok, Thailand

Wasinee Tianthong, M.D.;

Department of Obstetrics and Gynecology, King Chulalongkorn Memorial Hospital, Bangkok, Thailand

Assoc. Prof. Dhirapatara Charoenvidhya, M.D.

Department of Obstetrics and Gynecology, King Chulalongkorn Memorial Hospital, Bangkok, Thailand

Corresponding author

Patcharaporn Siricharoenthai, M.D.

Department of Obstetrics and Gynecology, King Chulalongkorn Memorial Hospital, Bangkok, Thailand

Patcharaporn.si@chula.ac.th

+66-868430497

Synopsis

This prospective cross-sectional study evaluated the diagnostic accuracy of genital tubercle angle (GTA), anogenital distance (AGD), and their combined use for first-trimester fetal sex determination in 317 singleton pregnancies. Both markers demonstrated comparable diagnostic performance, while their combination significantly improved discrimination, achieving an AUC of 0.919 and an overall accuracy of 86.8%. These findings suggest that combining GTA and AGD may improve early sonographic fetal sex determination, particularly at later gestational ages.

Keywords: Anogenital distance; Genital tubercle angle; First-trimester ultrasound; Fetal sex; Crown–rump length

Abstract

Background: Early fetal sex determination in the first trimester has important clinical implications, particularly in pregnancies at risk of sex-linked genetic disorders. The genital tubercle angle (GTA) has been widely used for early fetal sex determination, while anogenital distance (AGD) has emerged as a complementary marker; however, data on their comparative and combined diagnostic performance remain limited.

Objective: To evaluate and compare the diagnostic accuracy of GTA, AGD, and their combined use for fetal sex determination in singleton pregnancies at 11–13^{+6} weeks of gestation in a Thai population.

Methods: This prospective cross-sectional study included 317 singleton pregnancies at 11–13^{+6} weeks' gestation. GTA and AGD were measured in a standardized mid-sagittal plane, with three measurements averaged for analysis. Fetal sex was confirmed by postnatal phenotype. Diagnostic performance was assessed using receiver operating characteristic (ROC) analysis.

Results: Both mean GTA and AGD were significantly higher in male than in female fetuses ($p < 0.001$). For GTA, an optimal cut-off of 16.37° yielded a sensitivity of 75.0%, a specificity of 94.5%, and an AUC of 0.864. For AGD, a cut-off of 0.34 cm yielded an AUC of 0.861, with no significant difference in diagnostic performance compared with GTA ($p = 0.881$). The combined model integrating both markers significantly outperformed the individual parameters, achieving the highest AUC of 0.919 (95% CI: 0.884–0.948) and an overall accuracy of 86.8%. Diagnostic accuracy increased with gestational age, reaching 93.9% in the 13–13^{+6} weeks subgroup.

Conclusions: Both GTA and AGD are reliable markers for first-trimester fetal sex determination, with comparable individual performance. Their combined use significantly improves diagnostic

accuracy, particularly at later gestational ages, supporting their role as a noninvasive approach for early fetal sex assessment.

Keywords: Anogenital distance; Genital tubercle angle; First-trimester ultrasound; Fetal sex; Crown–rump length

Introduction:

Accurate determination of fetal sex is a topic with increasing clinical implication, especially in the context of prenatal screening and diagnosis. Traditionally, fetal gender has been confirmed during the mid-trimester ultrasound scan; however, technological advancements and improved imaging resolution have enabled the assessment of external genitalia as early as the late first trimester [1,2]. As noninvasive prenatal testing (NIPT) and first-trimester anatomical scanning become more widely adopted, the potential for early sonographic prediction of fetal sex has become both feasible and relevant [3]

One of the critical factors influencing fetal sex determination by ultrasound is the developmental morphology of the genital tubercle. The genital tubercle angle (GTA) in the mid-sagittal view has been explored as a marker for differentiating male and female fetuses, with promising accuracy from 11 to 13^{+6} weeks' gestation [1, 4, 5]. More recently, anogenital distance (AGD), the distance between the anus and genital tubercle, has been studied as another anatomical parameter that may assist in improving diagnostic precision [6,7,8]. However, despite these efforts, variability exists due to gestational age, fetal position, operator experience, and inter-individual anatomical differences [2, 9].

In clinical practice, the ability to accurately determine fetal sex at an earlier gestation may aid in guiding decisions for further invasive testing in cases of suspected sex-linked genetic conditions [3,7]. Additionally, such early information may enhance parental counseling and psychosocial preparation. Nonetheless, there remains a need for standardized measurement techniques and validation of combined anatomical markers, such as GTA and AGD, across diverse populations [6].

This study aims to evaluate the diagnostic accuracy of fetal sex determination using GTA, AGD, and a combination of both markers in fetuses between 11 and 13^{+6} weeks' gestation.

Furthermore, the research seeks to establish reference ranges for these parameters across this gestational window in a Thai population, thereby contributing local data to the growing body of international evidence.

Materials and Methods

This prospective cross-sectional study was conducted at King Chulalongkorn Memorial Hospital, Bangkok, Thailand, from March 2025 to March 2026. The study was approved by the Institutional Review Board of the Faculty of Medicine, Chulalongkorn University (IRB No. 0927/67). All participants provided written informed consent prior to enrollment.

Study Population and Participant Flow

The study included pregnant women ($\geq$18 years) with singleton pregnancies between 11 and 13^{+6} weeks of gestation. Exclusion criteria consisted of prenatally diagnosed chromosomal or structural abnormalities, ambiguous genitalia, intrauterine fetal demise with indeterminate sex, or miscarriage before 20 weeks.

A total of 331 participants were enrolled, with 14 cases (4.2%) subsequently excluded: 12 due to suboptimal visualization (fetal positioning or high maternal BMI), one due to chromosomal abnormality, and one lost to follow-up despite proactive telephone contact. The final analysis included 317 participants, stratified into three groups: 11–11^{+6} weeks (n=96), 12–12^{+6} weeks (n=123), and 13–13^{+6} weeks (n=98). Fetal sex was confirmed by postnatal phenotype documented by neonatologist.

Ultrasonographic Protocol

All examinations were performed using a Voluson E10 system (GE Healthcare, USA) equipped with a C1-6-D convex volume transducer for transabdominal scanning and an RIC5-9-D

microconvex volume transducer for transvaginal scanning. Measurements were acquired by a single experienced investigator (Pimchanoke P.). A transabdominal approach was initially used, followed by transvaginal ultrasound if visualization was suboptimal. In this study, the transvaginal approach was required for 5 participants to achieve optimal visualization of the fetal anatomy. To prioritize fetal safety, examinations adhered to the ALARA principle and total scan time limited to 20 minutes.

Measurements

Measurements were performed in a strict mid-sagittal view of the fetus. The fetus had to be in a neutral position with neither flexion nor hyperextension of the head. Additionally, the fetal crown-rump length (CRL) was magnified to occupy the maximum area of the display screen.

GTA Measurement:

- The first reference line (Horizontal Baseline): A straight line drawn tangent along the flat lower segment of the lumbosacral skin surface, just prior to the onset of the caudal curvature toward the rump.
- The second reference line (Tubercle Axis): A straight line drawn along the longitudinal axis of the genital tubercle, spanning from its base to the distal tip.
- Angle Calculation: The GTA was defined as the angle formed by the intersection of these two reference lines[3], as illustrated in Figure 1.

AGD Measurement:

The distance from the caudal end of the fetus to the base of the genital tubercle[7] (Figure 2).

Three independent measurements were obtained for each parameter from separate acquisitions, and the mean values were used for final analysis. To ensure data quality, the single operator

underwent standardization training under the close supervision and guidance of maternal-fetal medicine specialists. The consistency and reliability of these repeated measurements were evaluated using the Intraclass Correlation Coefficient (ICC), which revealed an ICC of 0.994 (95% CI: 0.981, 0.998) for the GTA and 0.972 (95% CI: 0.929, 0.989) for the AGD.

Statistical Analysis

Statistical analyses were performed using Stata version 19.5 (StataCorp, College Station, TX, USA). Continuous variables were expressed as mean ± standard deviation (SD), and categorical variables were presented as frequencies and percentages. Differences in GTA and AGD between male and female fetuses were compared using t-test.

The diagnostic performance of GTA and AGD was evaluated using Receiver Operating Characteristic (ROC) curve analysis. The optimal cut-off values for fetal sex determination were identified using the Youden Index (Sensitivity + Specificity - 1). A multivariable logistic regression model was employed to assess the synergistic diagnostic value of combining both markers. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy with corresponding 95% confidence intervals (CI) were calculated. A P-value of < 0.05 was considered statistically significant for all analyses.

Results

A total of 331 singleton pregnancies were enrolled, of which 14 were excluded. The final analysis included 317 singleton pregnancies (Figure 3. Participant flow chart). The cohort consisted of 152 (47.9%) male and 165 (52.1%) female fetuses. Fetal sex was confirmed by postnatal outcomes. As shown in Table 1, there were no statistically significant differences between the male

and female groups regarding maternal age, BMI or parity. The mean gestational age at the time of the scan was 12.55 ± 0.72 weeks for males and 12.38 ± 0.76 weeks for females. Baseline maternal characteristics were also comparable across the three gestational age groups (Table 1.1)

In the overall cohort, both the mean GTA and mean AGD were significantly higher in male fetuses compared to female fetuses (GTA: 35.06° ± 21.33° vs. 8.36° ± 5.86°, $p < 0.001$; AGD: 0.41 ± 0.09 cm vs. 0.30 ± 0.06 cm, $p < 0.001$). When stratified by gestational age (Table 2), both mean GTA and AGD values showed a significant progressive increase from 11 to 13^{+6} weeks ($p < 0.001$). Subgroup analyses within each gestational week confirmed that males consistently exhibited significantly larger GTA and AGD measurements compared to females at all time points ($p < 0.001$). Specifically:

- **At 11 to 11^{+6} weeks of gestation** (N = 96; 57 females, 39 males), the mean GTA was 25.89 ± 18.04 in males compared to 8.97 ± 6.48 in females ($p < 0.001$), while the mean AGD was 0.34 ± 0.04 in males and 0.26 ± 0.05 cm in females ($p < 0.001$).
- **At 12 to 12^{+6} weeks of gestation** (N = 123; 63 females, 60 males), the mean GTA was 33.98 ± 20.18 in males versus 7.64 ± 4.36 in females ($p < 0.001$), and the mean AGD was 0.40 ± 0.07 cm in males versus 0.30 ± 0.03 cm in females ($p < 0.001$).
- **At 13 to 13^{+6} weeks of gestation** (N = 98; 45 females, 53 males), the mean GTA reached 43.03 ± 22.19 in males and 8.60 ± 6.81in females ($p < 0.001$), with a corresponding mean AGD of 0.48 ± 0.08 cm in males and 0.36 ± 0.05 cm in females ($p < 0.001$).

The sex-specific distributions of these markers are visualized in Figure 4 and Figure 5.

The diagnostic performance is summarized in Table 3 and Figure 6. For GTA, an optimal cut-off of 16.37° achieved an AUC of 0.864 (95% CI: 0.818–0.906), with a sensitivity of 75.0%

and specificity of 94.5%. For AGD, a cut-off of 0.34 cm yielded an AUC of 0.861 (95% CI: 0.816–0.900). The combined model integrating both GTA and AGD significantly enhanced diagnostic performance, yielding the highest AUC of 0.919 (95% CI: 0.884–0.948) and an overall accuracy of 86.8%.

Diagnostic performance by gestational age group is presented in Table 4 and Figure 7. The accuracy of the combined model increased with advancing gestational age, reaching 93.9% in the 13 to 13^{+6} weeks group. Bootstrap comparison of AUCs (Table 5) demonstrated that the combined model had significantly higher AUC than both GTA alone and AGD alone in the overall cohort (both $p < 0.001$).

In subgroup analyses, the combined model demonstrated a significantly higher AUC than GTA alone across all gestational age groups (all $p < 0.01$), and was superior to AGD alone specifically within GA 12 - 12^{+6} weeks ($p = 0.025$) and GA 13 - 13^{+6} weeks ($p = 0.008$) cohorts. The diagnostic performance of the combined model improved progressively with advancing gestational age. Additionally, no significant difference in diagnostic performance was observed between GTA and AGD when used independently in the overall cohort ($p = 0.881$).

The combined model demonstrated excellent discriminative ability, with an overall AUC of 0.919 and an accuracy of 86.8%. Using an optimal probability cutoff of 0.46, the model achieved a sensitivity of 81.6% and a specificity of 91.5% (Table 3).

Discussion

In this prospective cross-sectional study of singleton pregnancies at 11–13^{+6} weeks of gestation, both GTA and AGD demonstrated good diagnostic performance for early fetal sex determination. The two parameters showed comparable accuracy when used individually, while their combined use significantly improved overall diagnostic performance. In addition, diagnostic

accuracy increased with advancing gestational age, with the highest performance observed in the late first trimester.

The diagnostic performance of GTA in this study is consistent with previous literature. Early work by Efrat et al. demonstrated that the orientation of the genital tubercle can reliably distinguish fetal sex, with accuracy improving as gestational age advances [1]. Subsequent studies, including those by Lubusky et al. and Schaefer et al., have reported high accuracy of first-trimester ultrasound for fetal sex determination, particularly after 12 weeks of gestation [2,3].

AGD also demonstrated good diagnostic performance in our study, consistent with emerging evidence supporting its utility. Previous studies have shown that AGD is significantly longer in male fetuses and can be used to predict fetal sex with reasonable accuracy in the first trimester [7–9]. More recent studies, such as those by Alfuraih et al. and Edris et al., further support the role of AGD as a reliable marker with good discriminative ability[6,8]. Our findings are in agreement with these reports and confirm that AGD performs comparably to GTA when used independently.

A distinctive strength of our study, which addresses a gap in existing literature, is the comprehensive head-to-head comparison of GTA, AGD, and their combined model within the same cohort. While previous research often evaluated these markers in isolation, our integrated approach demonstrates that their synergy significantly optimizes diagnostic yield. Furthermore, although we utilized an overall optimal cut-off for the entire cohort, our stratified analysis and the distribution trends visualized in Figures 4 and 5 highlight the profound influence of anatomical growth. Unlike earlier studies that may rely on a static threshold, our data suggests that the progressive increase in GTA and AGD values—from the early phenotypic transition at 11 weeks to the more defined

structures at 13 weeks—should be carefully considered to minimize misclassification in clinical practice.

The superior performance of the combined model incorporating both GTA and AGD likely reflects the complementary nature of these measurements. GTA represents the directional development of the genital tubercle, whereas AGD reflects androgen-dependent growth of the perineal region. Both parameters are influenced by fetal androgen exposure during early genital differentiation, and their combined use may capture distinct but related aspects of this biological process, thereby improving overall diagnostic accuracy. Similar improvements with combined markers have been reported in recent studies [6].

Our findings also demonstrate that diagnostic accuracy improves with advancing gestational age, which is consistent with previous reports [1–3]. This trend can be explained by progressive sexual differentiation and increasing anatomical clarity as gestation advances. In earlier weeks, subtle anatomical differences may be more difficult to distinguish, whereas later in the first trimester, the genital structures become more clearly defined, facilitating more accurate assessment.

From a clinical perspective, first-trimester ultrasound assessment using a combined model of GTA and AGD provides a non-invasive, accessible, and cost-effective approach for early fetal sex determination, which is particularly valuable in settings where access to NIPT is limited. Enhancing accuracy through this combined approach substantially optimizes its potential and confidence for clinical translation in managing high-risk pregnancies. Most notably, for pregnant women who are carriers of severe sex-linked genetic disorders in resource-limited settings, incorporating these measurements into routine first-trimester anomaly scans serves as an effective clinical triage tool from the initial visit. In cases where a male fetus is indicated, it enables the medical team to initiate early genetic counseling and proactively discuss comprehensive

management strategies with the family in a timely manner. Furthermore, rapid sex prediction significantly mitigates maternal anxiety associated with waiting for standard mid-trimester ultrasound assessments, thereby improving the overall psychological well-being of pregnant women and fostering patient-centered care throughout prenatal management.

This study has several strengths, including its prospective design and the use of a highly standardized measurement protocol. To our knowledge, this is the first study in Thailand to comprehensively evaluate the diagnostic accuracy of combined GTA and AGD for early fetal sex determination. Establishing these baseline sonographic parameters within a Thai cohort provides valuable local reference data that may be more representative of our specific population. Furthermore, all assessments were performed by a single experienced operator, which ensured consistent application of the measurement techniques and minimized inter-observer variability. This uniformity is crucial for establishing reliable diagnostic thresholds for GTA and AGD in a controlled setting.

However, several limitations should be acknowledged. First, since the measurements were conducted by a single operator at a single center, the generalizability of our findings to different clinical settings or less experienced practitioners may be limited. Ultrasound assessment is inherently operator-dependent, and performance may vary based on individual expertise and image quality.

Future research should focus on multicenter validation, standardization of measurement techniques, and exploration of adjunctive tools, including advanced imaging and artificial intelligence, to further improve diagnostic performance.

In conclusion, both GTA and AGD are reliable markers for fetal sex determination in the first trimester, with comparable diagnostic performance when used individually. The combination

of these parameters significantly improves diagnostic accuracy, particularly in later gestational weeks, supporting their role in early, noninvasive fetal sex assessment.

List of abbreviations

AGD: Anogenital distance

AUC: Area under the curve

BMI: Body mass index

CI: confidence interval

GTA: Genital tubercle angle

NIPT: Noninvasive prenatal testing

NPV: Negative predictive value

PPV: Positive predictive value

ROC: Receiver operating characteristic

Declarations

Ethics approval and consent to participate

Ethics approval by Institutional Review Board of the Faculty of Medicine, Chulalongkorn University (IRB No. 0927/67).

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Conflict of interest

The author declare that they have no conflict of interest

Funding

Not applicable

Authors' contributions

All authors contributed substantially to this work. PS and PP designed methodology. PP collected and interpreted data and major contributor in writing the manuscript. WT and DC analyzed and interpreted the patient data. PS was major contributor in interpreting and proved final version of manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank all the staff, fellows and nurses of the Division of Maternal- Fetal Medicine, Department of Obstetrics and Gynecology, King Chulalongkorn Memorial Hospital, Bangkok, Thailand

References

1. Efrat Z, Akinfenwa OO, Nicolaides KH. First-trimester determination of fetal gender by ultrasound. *Ultrasound Obstet Gynecol.* 1999;13(5):305-307.
2. Lubusky M, Studnickova M, Skrivanek A, et al. Ultrasound evaluation of fetal gender at 12–14 weeks. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub.* 2012;156(4):324-329.

3. Schaefer EC, McKenna DS, Sonek JD. First trimester identification of fetal sex by ultrasound. *Arch Gynecol Obstet.* 2024;309:1453-1458.
4. Gharekhanloo F. The ultrasound identification of fetal gender at the gestational age of 11-12 weeks. *J Hum Reprod Sci.* 2014;7(3):209-212.
5. Savirón-Cornudella R, González-Ballano I, Cisneros-Gimeno A, et al. Ultrasound measurement learning of fetal sex during the first trimester: does the experience matter? *Res Rep Focus Ultrasound.* 2015;3:1-9.
6. Alfuraih AM, Almajem BM, Alsolai AA. The Predictive Accuracy of Anogenital Distance and Genital Tubercle Angle for First-Trimester Fetal Sex Determination. *Diagnostics.* 2024;14(16):1811.
7. Sipahi M, Tokgöz VY, Alanya Tosun Ş. An appropriate way to predict fetal gender at first trimester: anogenital distance. *J Matern Fetal Neonatal Med.* 2019;32(12):2012-2016.
8. Edris TI, Mohamed MA, Mohamady LA, et al. Anogenital Distance for Detection of Fetal Sex in First Trimester. *Egypt J Hosp Med.* 2024;94:885-888.
9. Najdi N, Safi F, Hashemi-Dizaji S, et al. First trimester determination of fetal gender by ultrasonographic measurement of anogenital distance: A cross-sectional study. *Int J Reprod Biomed.* 2019;17(1):51-56.

Table 1. Baseline characteristics overall and by confirmed fetal sex

Variable	Overall (317)	Female (165)	Male (152)	p-value
Maternal age (years)	33.08 ± 4.48	32.75 ± 4.61	33.45 ± 4.32	0.163
Gravidarum	1 (1-2)	1 (1-2)	1 (1-2)	0.423
Parity	0 (0-1)	0 (0-1)	0 (0-1)	0.807
Height (cm)	159.12 ± 5.69	158.90 ± 5.40	159.36 ± 6.00	0.481
Weight (kg)	60.46 ± 12.09	60.24 ± 10.15	60.69 ± 13.92	0.749
BMI (kg/m²)	23.86 ± 4.42	23.83 ± 3.92	23.90 ± 4.91	0.891
Gestational age (weeks)	12.46 ± 0.74	12.38 ± 0.76	12.55 ± 0.72	0.037

Continuous variables are shown as mean ± SD or median (IQR).

Categorical variables are shown as n (%). P values compare female versus male.

Abbreviations: BMI, body mass index; cm, centimeter; Kg, kilogram; kg/m2, kilogram per square meter

Table 1.1 Baseline characteristics by gestational age group

Variable	Overall (317)	GA1 (96)	GA2 (123)	GA3 (98)	p-value
Maternal age (years)	33.08 ± 4.48	33.24 ± 4.29	33.13 ± 4.41	32.87 ± 4.77	0.837
Gravidarum	1 (1-2)	1 (1-2)	1 (1-2)	1 (1-2)	0.897
Parity	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-1)	0.633
Height (cm)	159.12 ± 5.69	158.41 ± 4.96	159.40 ± 5.71	159.47 ± 6.29	0.338
Weight (kg)	60.46 ± 12.09	59.89 ± 10.15	60.24 ± 12.10	61.27 ± 13.77	0.708
BMI (kg/m²)	23.86 ± 4.42	23.81 ± 3.88	23.70 ± 4.51	24.12 ± 4.82	0.783
Gestational age (weeks)	12.46 ± 0.74	11.59 ± 0.30	12.44 ± 0.27	13.35 ± 0.27	<0.001
Male sex	152 (47.9%)	39 (40.6%)	60 (48.8%)	53 (54.1%)	0.167

GA1 = 11 to 11^{+6} weeks; GA2 = 12 to 12^{+6} weeks; GA3 = 13 to 13^{+6} weeks.

Abbreviations: BMI, body mass index; cm, centimeter; GA, gestational age; Kg, kilogram; kg/m2, kilogram per square meter

Table 2. GTA and AGD distribution table by gestational age

Variable	Overall (317)	GA1 (96)	GA2 (123)	GA3 (98)	p-value
GTA measurement 1 (degrees)	21.31 ± 22.75	15.07 ± 15.14	19.85 ± 21.17	29.27 ± 28.16	<0.001
GTA measurement 2 (degrees)	20.34 ± 21.28	14.45 ± 16.30	20.23 ± 21.02	26.26 ± 24.30	<0.001
GTA measurement 3 (degrees)	21.84 ± 20.83	18.01 ± 17.39	21.40 ± 20.51	26.14 ± 23.54	0.023
GTA mean (degrees)	21.17 ± 20.34	15.84 ± 14.99	20.49 ± 19.53	27.22 ± 24.14	<0.001
AGD measurement 1 (cm)	0.36 ± 0.10	0.30 ± 0.08	0.35 ± 0.09	0.43 ± 0.10	<0.001
AGD measurement 2 (cm)	0.35 ± 0.09	0.29 ± 0.07	0.35 ± 0.08	0.42 ± 0.09	<0.001
AGD measurement 3 (cm)	0.35 ± 0.10	0.29 ± 0.06	0.34 ± 0.07	0.42 ± 0.11	<0.001
AGD mean (cm)	0.35 ± 0.09	0.29 ± 0.06	0.35 ± 0.07	0.42 ± 0.09	<0.001

GA1 = 11 to 11^{+6} weeks; GA2 = 12 to 12^{+6} weeks; GA3 = 13 to 13^{+6} weeks.

Abbreviations: AGD, anogenital distance; cm, centimeter; GA, gestational age; GTA, genital tubercle angle

Table 3. Overall diagnostic accuracy for prediction of male sex

Test	n	AUC (95% CI)	cutoff	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	accuracy
GTA	317	0.864 (0.818, 0.906)	16.37	75.0% (67.3%, 81.7%)	94.5% (89.9%, 97.5%)	92.7% (86.6%, 96.6%)	80.4% (74.1%, 85.8%)	85.2%
AGD	317	0.861 (0.816, 0.900)	0.34	81.6% (74.5%, 87.4%)	77.6% (70.4%, 83.7%)	77.0% (69.7%, 83.3%)	82.1% (75.1%, 87.7%)	79.5%
GTA + AGD	317	0.919 (0.884, 0.948)	0.46	81.6% (74.5%, 87.4%)	91.5% (86.2%, 95.3%)	89.9% (83.6%, 94.3%)	84.4% (78.2%, 89.3%)	86.8%

AUC confidence intervals were estimated by nonparametric bootstrap with 3,000 resamples. Optimal cutoff was chosen by the Youden index.

Abbreviations: AGD, anogenital distance; AUC, area under the curve; cm, centimeter; GA, gestational age; GTA, genital tubercle angle, NPV, negative predictive value; PPV, positive predictive value

Table 4. Diagnostic accuracy by gestational age group

Group	Test	n	AUC (95% CI)	Accuracy
GA1	GTA	96	0.772 (0.661, 0.872)	80.2%
	AGD	96	0.882 (0.800, 0.946)	83.3%
	GTA + AGD	96	0.904 (0.835, 0.963)	88.5%
GA2	GTA	123	0.873 (0.798, 0.937)	87.0%
	AGD	123	0.905 (0.842, 0.955)	87.0%
	GTA + AGD	123	0.944 (0.895, 0.981)	91.1%
GA3	GTA	98	0.925 (0.865, 0.973)	87.8%
	AGD	98	0.905 (0.841, 0.959)	83.7%
	GTA + AGD	98	0.968 (0.932, 0.994)	93.9%

GA1 = 11 to 11^{+6} weeks; GA2 = 12 to 12^{+6} weeks; GA3 = 13 to 13^{+6} weeks.

Abbreviations: AGD, anogenital distance; AUC, area under the curve; CI, confidence interval; GA, gestational age; GTA, genital tubercle angle

Table 5. Bootstrap comparisons of AUCs

Group	Comparison	Mean AUC difference	95% CI low	95% CI high	Bootstrap p-value
Overall	Combined vs GTA	0.055	0.028	0.083	<0.001
	Combined vs AGD	0.058	0.028	0.091	<0.001
	GTA vs AGD	0.004	-0.049	0.055	0.881
GA1	Combined vs GTA	0.132	0.046	0.224	0.002
	Combined vs AGD	0.022	-0.022	0.073	0.365
	GTA vs AGD	-0.111	-0.226	0.009	0.071
GA2	Combined vs GTA	0.072	0.021	0.134	0.003
	Combined vs AGD	0.038	0.004	0.081	0.025
	GTA vs AGD	-0.032	-0.113	0.046	0.432
GA3	Combined vs GTA	0.043	0.011	0.081	0.003
	Combined vs AGD	0.062	0.013	0.119	0.008

Group	Comparison	Mean AUC difference	95% CI low	95% CI high	Bootstrap p-value
	GTA vs AGD	0.020	-0.055	0.099	0.634

GA1 = 11 to 11^{+6} weeks; GA2 = 12 to 12^{+6} weeks; GA3 = 13 to 13^{+6} weeks.

Positive AUC difference indicates that the first test in the comparison had a larger AUC than the second.

Abbreviations: AGD, anogenital distance; AUC, area under the curve; CI, confidence interval; GA, gestational age; GTA, genital tubercle angle

Figure 1: Ultrasound image demonstrating GTA measurement

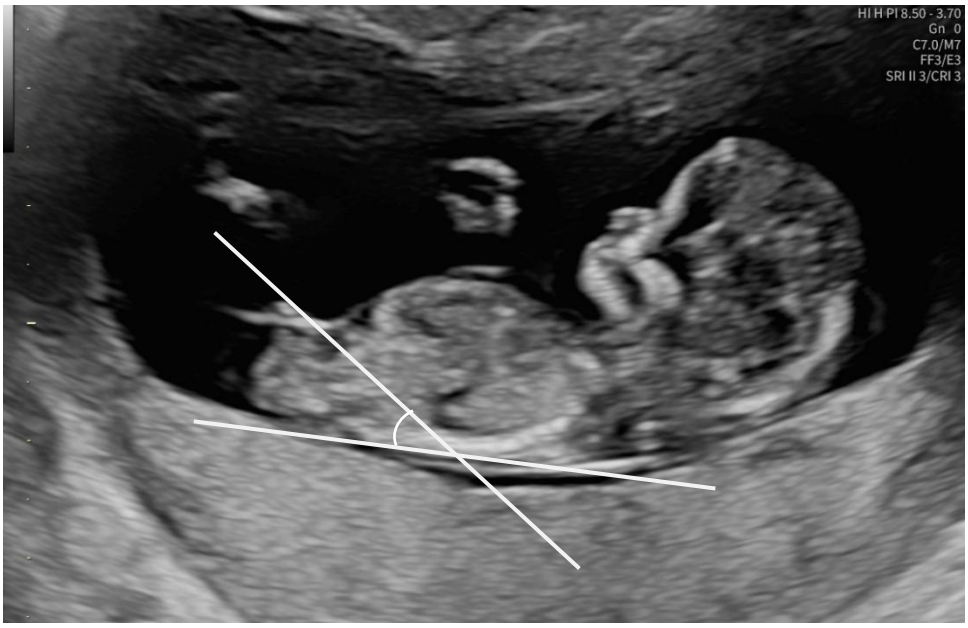


Figure 2: Ultrasound image demonstrating AGD measurement



Figure 3. Study Participant Flowchart

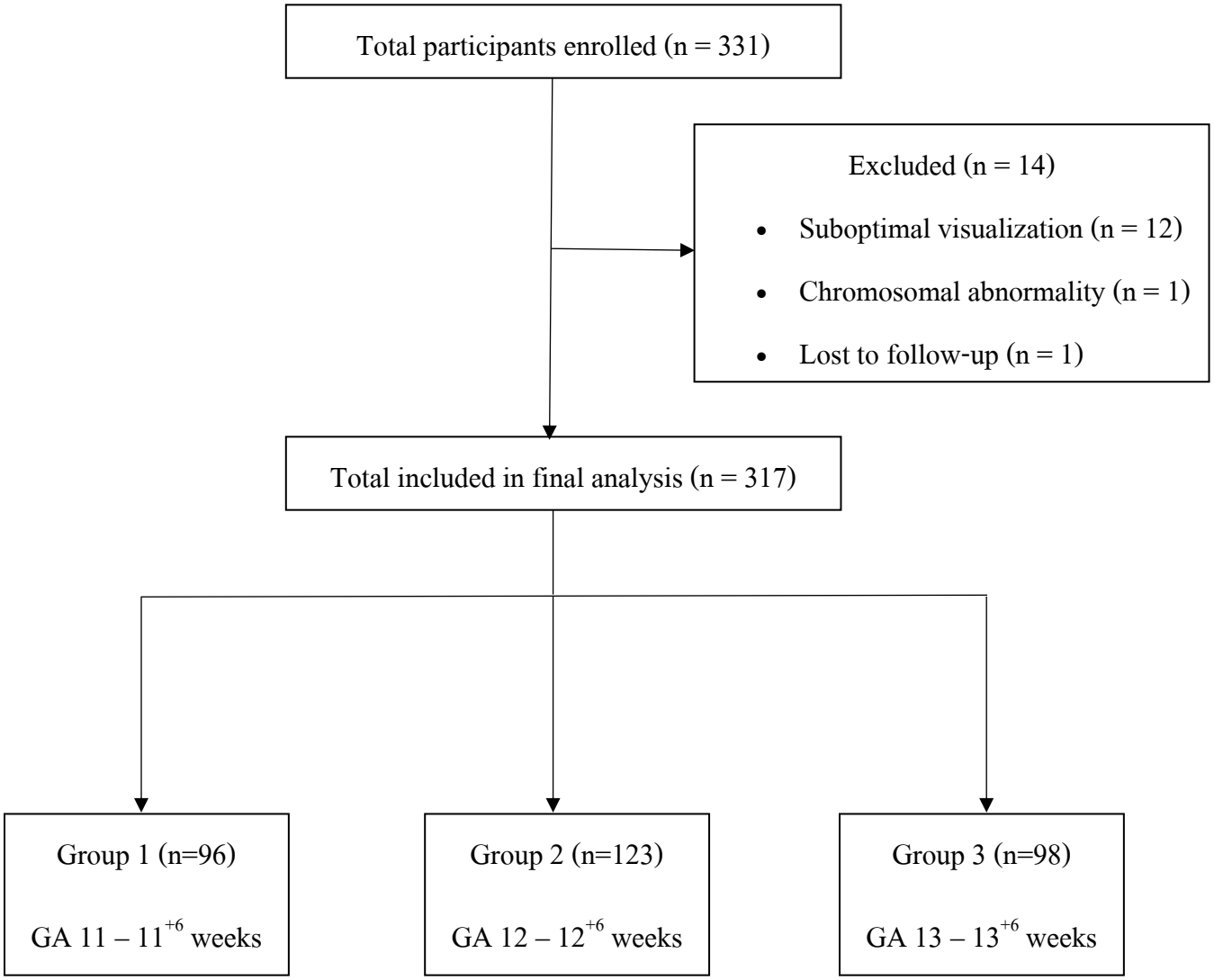
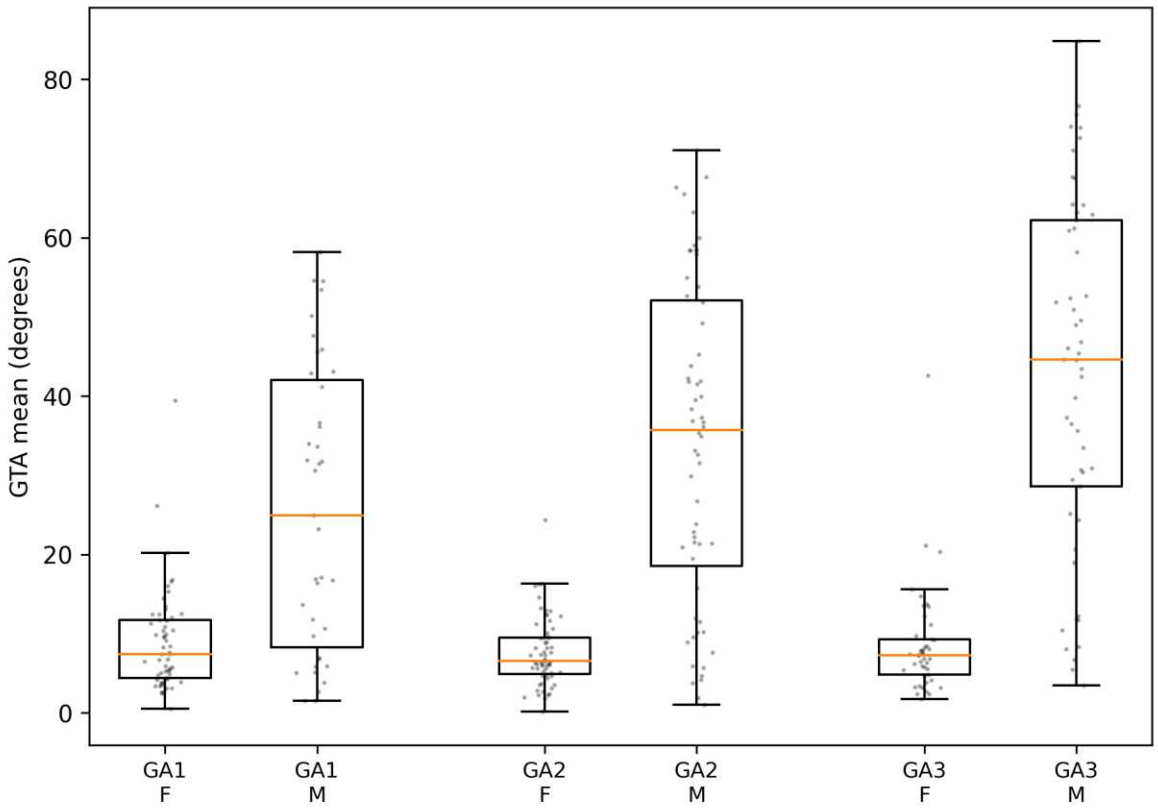


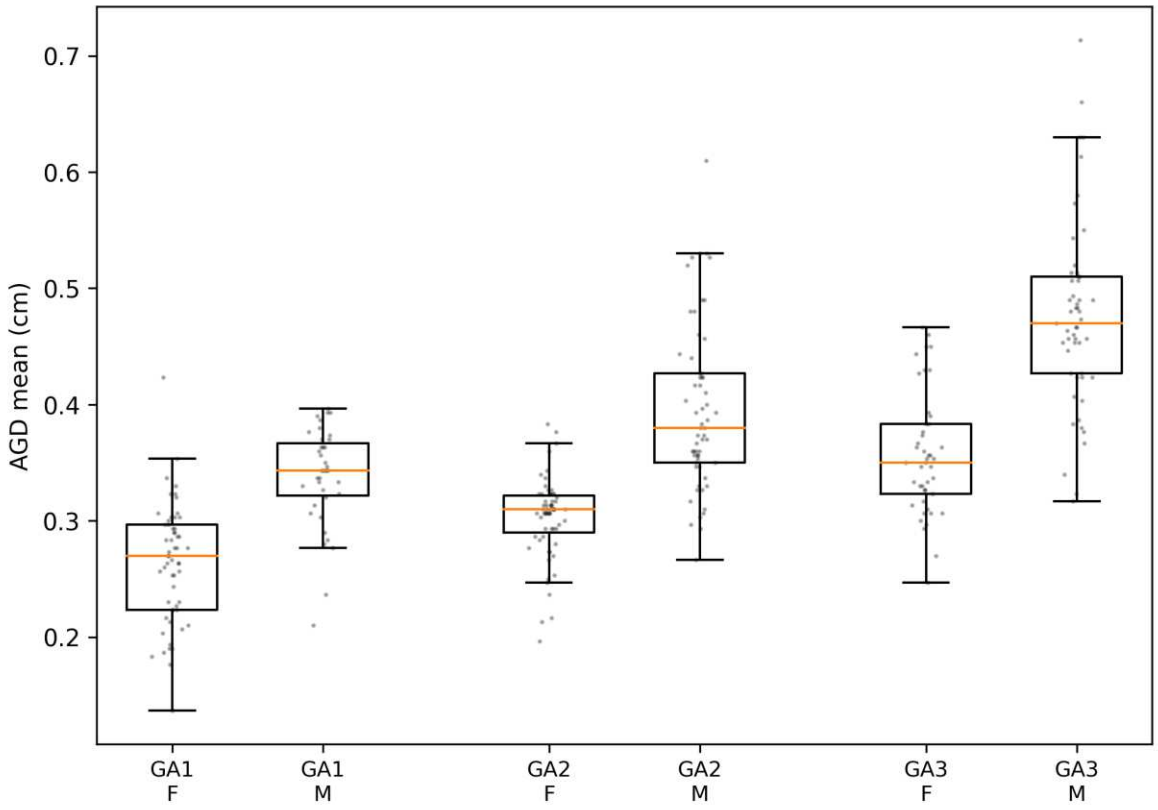
Figure 4. Distribution of GTA mean by GA group and confirmed fetal sex



GA1 = 11 to 11^{+6} weeks; GA2 = 12 to 12^{+6} weeks; GA3 = 13 to 13^{+6} weeks.

Boxplots and overlaid individual observations for GTA mean, stratified by gestational age group and confirmed fetal sex.

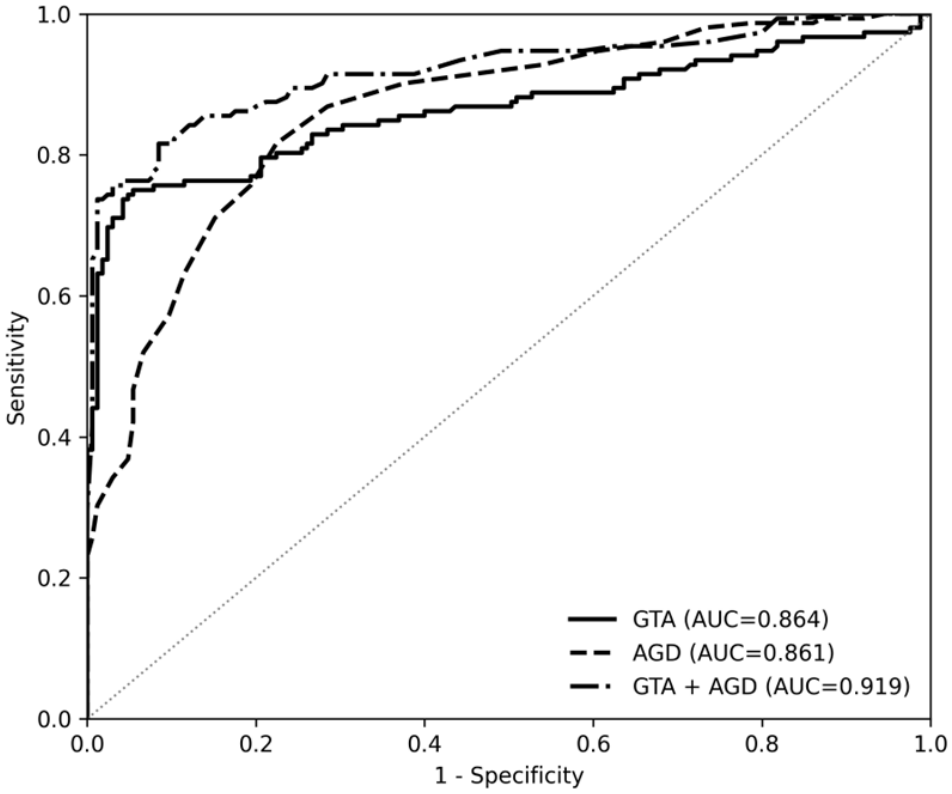
Figure 5. Distribution of AGD mean by GA group and confirmed fetal sex



GA1 = 11 to 11^{+6} weeks; GA2 = 12 to 12^{+6} weeks; GA3 = 13 to 13^{+6} weeks.

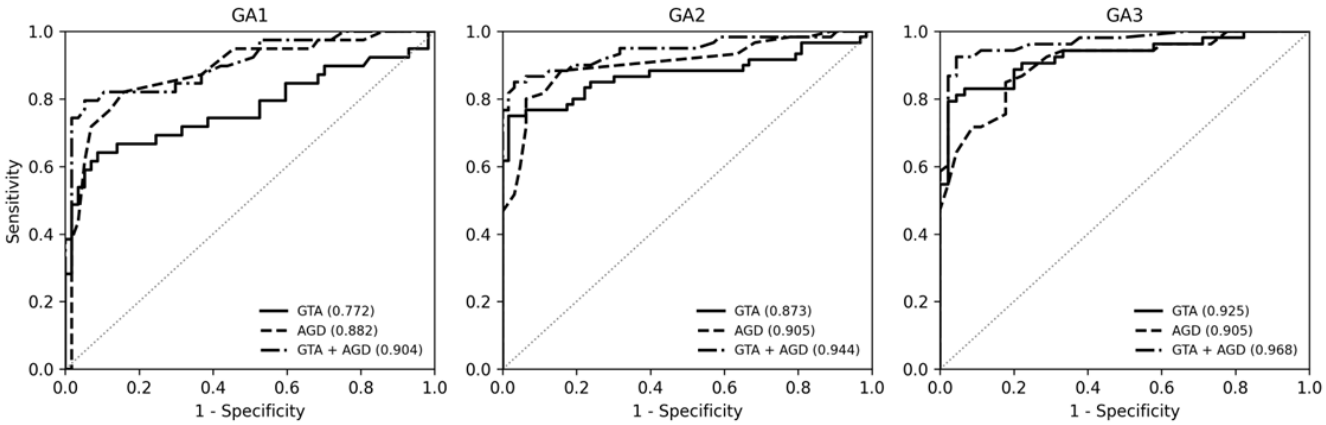
Boxplots and overlaid individual observations for AGD mean, stratified by gestational age group and confirmed fetal sex.

Figure 6. Overall ROC curves



ROC curves for GTA mean, AGD mean, and the combined logistic model using both markers for prediction of male sex in the overall cohort.

Figure 7. ROC curves stratified by gestational age group



GA1 = 11 to 11^{+6} weeks; GA2 = 12 to 12^{+6} weeks; GA3 = 13 to 13^{+6} weeks.

ROC curves for GTA mean, AGD mean, and the combined model within GA1, GA2, and GA3.