

Mid-Luteal Serum Progesterone in Infertile Women With Hypothyroidism Compared With Euthyroid Women With Proven Fertility: An Exploratory Case-Control Study

Saman Saeed Ansari^{1, 2}, Tashfeen Ikram¹, Aisha Bashir³, Qanita Mahmud⁴, Ambreen Anjum⁵, Tahir Maqbool², Muhammad Tahir^{2,6}

¹ Department of Physiology, Rashid Latif Medical College, Lahore, Pakistan EMAIL: Saman.saeed@rlmc.edu.pk (S.S.A.); EMAIL: tashfeen.ikram@rlmc.edu.pk (T.I) ORCID ID:0009-0001-7231-3634; ORCID ID: 0009-0003-5397-0721

² Institute of Molecular Biology and Biotechnology, The University of Lahore, Lahore, Pakistan EMAIL: tahir.maqbool@imbb.uol.edu.pk (T.M) ORCID ID: 0000-0003-3706-2582

³ Department of Diagnostic Radiology, Foothills Medical Centre, Calgary, Alberta, Canada EMAIL: aisha_bashir@hotmail.com (A.B) ORCID ID: 0009-0004-8422-5756

⁴ Department of Physiology, Fatima Jinnah Medical University, Queen's Road, Lahore, Pakistan EMAIL: qanitamahmud@gmail.com (A.M) ORCID ID: 0009-0006-1370-0251

⁵ Department of Physiology, Al-Aleem Medical College, Lahore, Pakistan EMAIL: dranjum28@yahoo.com (A.A) ORCID ID: 0000-0002-6053-9281

⁶ Department of Cell Biology and Anatomy, New York Medical College, Valhalla, New York, USA EMAIL:mtahir2@nymc.edu (M.T) ORCID ID: 0000-0001-8652-2042

CORRESPONDING AUTHOR: Saman Saeed Ansari, Email: saman.saeed@rlmc.edu.pk

ABSTRACT

Importance Hypothyroidism is linked to female infertility, but its association with luteal-phase progesterone remains uncertain.

Objective To compare mid-luteal serum progesterone in infertile women with untreated hypothyroidism and euthyroid women with proven fertility.

Design, Setting, And Participants This exploratory case-control study enrolled 107 women aged 25 to 44 years with untreated hypothyroidism and primary (n = 59) or secondary (n = 48) infertility, along with 95 euthyroid controls with proven fertility, at Arif Memorial Hospital in Lahore, Pakistan. Twelve enrolled controls were excluded for ineligibility. Women using hormonal medications, including ovulation induction agents or progesterone, were not eligible. Blood was collected 7 days before the next menses.

Exposures Serum progesterone, assessed continuously and in study-defined categories (<10, 10-20, and >20 ng/mL).

Main Outcomes And Measures The outcome was case status. Groups were compared using Mann-Whitney U tests and Hodges-Lehmann (HL) location shifts. Firth penalized logistic regression estimated adjusted odds ratios (aORs) controlling for age and reproductive hormones.

Results Median (IQR) progesterone was 18.5 (15.3-24.8) ng/mL in cases and 10.0 (8.0-13.5) ng/mL in controls (HL shift, 8.0 ng/mL; 95% CI, 6.6-10.0 ng/mL; P < .001). Progesterone exceeded 20 ng/mL in 39 cases (36.4%) and 1 control (1.1%) (aOR vs 10-20 ng/mL, 25.4; 95% CI, 4.9-264.5). Estimates were similar after restricting age and modeling it flexibly. Among 83 age-matched pairs, the paired HL shift was 7.1 ng/mL (95% CI, 4.8-9.1 ng/mL). The association was nonlinear: values below 10 ng/mL were more common in controls than cases (41.1% vs 14.0%). Progesterone did not differ between overt and subclinical hypothyroidism.

Conclusions And Relevance Infertile women with untreated hypothyroidism had higher single mid-luteal progesterone concentrations than euthyroid women with proven fertility. Because the groups differed in both fertility and thyroid status, the observed difference cannot be attributed specifically to hypothyroidism. The findings are hypothesis-generating and do not establish pathological progesterone excess or diagnostic utility.

Keywords: Mid-Luteal, Serum Progesterone, Infertile Hypothyroidism, Euthyroid Women Proven Fertility.

Introduction

BACKGROUND

Infertility affects an estimated 17.5% of adults during their lifetime (1). Hypothyroidism is a common, treatable contributor: it affects up to 12% of the population, is more prevalent in women (2), and has a pooled prevalence of 28% among infertile women in India (3). Subclinical hypothyroidism is also common among subfertile women in Pakistan (4). Thyroid hormones act on the ovary and endometrium (5), local triiodothyronine (T₃) signaling supports endometrial receptivity (6), and experimental hypothyroidism disrupts uterine progesterone signaling (7). In South Asia, hypothyroidism is often first detected during infertility evaluation, when thyroid and reproductive hormones are measured together, offering an opportunity to examine their relationship within the same women.

Despite this biological rationale, the reproductive benefit of thyroid treatment remains uncertain, and recent guidance rates much of the evidence as low to moderate quality (8, 9). Levothyroxine did not increase live births in euthyroid women with thyroid peroxidase antibodies (10), and women who developed and received treatment for hypothyroidism conceived less often than euthyroid women (11). In a meta-analysis of randomized trials involving pregnant women with subclinical hypothyroidism, levothyroxine was associated with

lower pregnancy-loss risk but no statistically significant improvement in live birth; these results do not directly establish effects on preconception fertility (12). Thyroid autoimmunity may impair fertility independently of thyroid function (13). Downstream hormonal pathways including prolactin (14, 15), ovarian reserve (16, 17), and steroid-binding proteins (18) may therefore contribute to infertility.

Luteal progesterone is a less well-characterized component of the relationship between thyroid dysfunction and fertility. Measured approximately 1 week before expected menstruation, serum progesterone can provide presumptive evidence of recent ovulation. However, pulsatile secretion limits interpretation of a single concentration, which cannot reliably establish luteal adequacy or diagnose luteal phase deficiency. Luteal phase deficiency has not been established as an independent cause of infertility (19). Experimental findings are inconsistent: T_3 stimulated progesterone release from human luteal cells (20), high thyrotropin (TSH) impaired luteal cell survival without altering progesterone output (21), and T_3 had no direct effect on rat luteal steroidogenesis (22). In stimulated assisted reproduction cycles, elevated progesterone at ovulation trigger has been associated with lower live-birth rates, although that setting differs from natural-cycle luteal sampling (23). In a randomized substudy of euthyroid women undergoing modified natural-cycle frozen embryo transfer, short-term vaginal progesterone supplementation did not significantly alter TSH or free thyroxine trajectories (24). This finding does not establish the direction of the association between endogenous progesterone and hypothyroidism.

Direct evidence remains limited regarding comparisons of mid-luteal serum progesterone between infertile women with hypothyroidism and euthyroid fertile controls sampled at the same cycle phase. We therefore conducted an exploratory case-control study comparing mid-luteal progesterone in untreated infertile women with hypothyroidism and euthyroid women with proven fertility, and assessed whether differences persisted after accounting for age and other reproductive hormones.

METHODS

Study Design and Participants

The Institutional Review Board of the Institute of Molecular Biology and Biotechnology, The University of Lahore, Lahore, Pakistan, approved this case-control study (approval No. IMBB/BBBC/22/089; March 12, 2022). All participants provided written informed consent. The study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline for case-control studies (25).

Women were recruited from the infertility and endocrine services at Arif Memorial Hospital, Lahore, Pakistan, between April 3, 2023, and May 10, 2024; all samples were analyzed in one laboratory. Cases were women aged 25 to 44 years with infertility defined as failure to achieve a clinical pregnancy after at least 12 months of regular unprotected intercourse and biochemically confirmed hypothyroidism on baseline thyroid testing (TSH ≥ 4.9 mIU/L with the corresponding free hormone profile). Infertility was classified as primary (no previous conception) or secondary (at least 1 previous conception). Hypothyroidism was classified post hoc as overt (FT4 < 0.9 ng/dL, the laboratory lower limit) or subclinical (FT4 ≥ 0.9 ng/dL). Exclusion criteria were thyroid hormone replacement; current or recent use of other exogenous hormonal medications, including oral contraceptives, ovulation induction agents, or progesterone supplements; previous radioactive iodine treatment or thyroid surgery; diagnosed systemic autoimmune disease, active inflammatory disease, or malignant neoplasm; and known male-factor infertility or a tubal or uterine structural abnormality in the couple. Thus, all cases were untreated for hypothyroidism when sampled.

Controls were women aged 25 to 44 years with proven natural fertility (at least 1 previous term pregnancy), a normal reproductive history, and a euthyroid laboratory profile. Controls were recruited independently through the staff coordinator of the Department of Physiology, Rashid Latif Medical College, Lahore, Pakistan, with study sampling at Arif Memorial Hospital, Lahore, Pakistan. Exclusion criteria were a history of infertility, recurrent pregnancy loss, or obstetric complications; current or recent use of exogenous hormonal medications, including oral contraceptives or hormone replacement therapy; thyroid dysfunction or thyroid surgery; and active systemic or autoimmune disease. Controls were enrolled within the same age range as cases to improve age comparability, but neither individual nor frequency matching was used.

Of the 107 enrolled controls, 12 were excluded from all analyses after review showed that they did not meet the eligibility criteria. Two were older than 44 years (48 and 49 years), and 10 had thyroid test results outside the laboratory reference ranges (TSH, 0.5-5.0 mIU/L; FT4, 0.9-2.3 ng/dL), including 5 with TSH levels above 5.0 mIU/L. The final analysis included 107 cases and 95 controls.

Sampling and Measurements

A single blood sample was collected during the mid-luteal phase, 7 days before the next menstrual period, to approximate the expected luteal progesterone peak. Sampling was scheduled according to each woman's usual cycle length and confirmed using the recorded onset date of the subsequent menses. Because luteal-phase length is relatively stable, collection was expected to occur approximately 7 days after ovulation; however, ovulation was not verified by ultrasonography or urinary luteinizing hormone testing. Measurements included erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), TSH, free T₃ (FT3), free thyroxine (FT4), prolactin, follicle-stimulating hormone (FSH), luteinizing hormone (LH), anti-Müllerian hormone (AMH), estradiol, and progesterone. Thyroid antibodies were recorded only for controls. Seventeen text entries were converted to numeric values. Fourteen case FT3 values (range, 4.0-20.5) and 4 case FT4 values (range, 10-15) were inconsistent with the reported units, possibly reflecting different reporting units. Because the original units could not be verified against laboratory reports, these values were excluded rather than converted; the 4 cases without a valid FT4 value were therefore not classified as having overt or subclinical hypothyroidism.

Exposures and Outcome

Case status was the outcome, and serum progesterone was the exposure. Because the association with case status was nonlinear (see Statistical Analysis), adjusted analyses used 3 study-defined categories: low (<10 ng/mL), intermediate (10-20 ng/mL; reference), and high (>20 ng/mL). These categories describe the observed distribution and are not diagnostic thresholds for luteal deficiency or pathological progesterone excess (19). A 25-ng/mL cutoff was evaluated in a sensitivity analysis. Covariates included age, prolactin, LH, FSH, AMH, and estradiol to explore the association after accounting for age and measured reproductive hormones (14, 15, 16, 17, 18), with CRP and ESR included as additional covariates. Thyroid measures defined case status and were therefore not included as covariates. Because the thresholds, covariates, and analyses described below were selected during analysis rather than preregistered, all findings are exploratory.

Sample Size

The planned sample size was derived from pilot data on 20 infertile women sampled on the same schedule and excluded from the main study. Assuming a natural-log progesterone SD of 0.37, 99 women per group provided 80% power (2-sided $\alpha = .05$) to detect a standardized difference of 0.40, equivalent to a geometric mean ratio (GMR) of 1.16. Allowing 8% for incomplete sampling (99×1.08) increased the recruitment target to 107 per group. All 107 cases and 95 of the 107 enrolled controls were eligible and included in the analysis.

Observed natural-log SDs exceeded the assumed value (0.42 in controls and 0.60 in cases). We therefore recalculated the detectable effect using the observed variability rather than conducting a post hoc power analysis based on the observed effect. With 107 cases and 95 controls, the estimated minimum detectable GMR at 80% power was 1.23 (standardized difference, 0.40), and the corresponding GMR at 90% power was 1.27 (Welch t test; 2-sided $\alpha = .05$). Detecting the originally targeted GMR of 1.16 at the observed variability would have required approximately 192 women per group. Accordingly, study precision was evaluated from the CIs around the observed estimates.

Statistical Analysis

Analyses were performed in Python version 3.13 (Python Software Foundation) using SciPy, statsmodels, and scikit-learn. All markers departed from normality in at least 1 group (Shapiro-Wilk $P < .05$), and 10 of 11 showed unequal variances. Continuous variables were therefore summarized as median (IQR) and compared using the Mann-Whitney U test. The Hodges-Lehmann (HL) estimate—the median of all pairwise case-control differences—was reported with bootstrap 95% CIs (2000 resamples); when distribution shapes differ, it represents a location shift rather than the difference between group medians. Additional effect sizes included rank-biserial correlations (r) and standardized mean differences (SMDs). Log-transformed progesterone was also compared using the Welch t test and expressed as a geometric mean ratio (GMR). P values for the 7 nonthyroid secondary markers were adjusted using the Benjamini-Hochberg method. The primary group comparisons used all eligible participants without weighting or resampling to equalize group sizes; bootstrap resampling was used for uncertainty estimation.

Linearity of continuous progesterone in the log odds was evaluated by comparing a linear term with a restricted cubic spline using a likelihood ratio test. Because linearity was rejected, adjusted analyses used progesterone categories. Firth penalized logistic regression, which reduces small-sample bias when category counts are sparse, estimated adjusted odds ratios (aORs) with profile-likelihood 95% CIs. Model 1 adjusted for age, prolactin, LH, FSH, AMH, and log₂ estradiol; model 2 additionally adjusted for CRP and ESR. Collinearity was evaluated using

variance inflation factors, model fit using the Hosmer-Lemeshow test, and group separation using the area under the receiver operating characteristic curve (AUC).

Within cases, dose-response was assessed using Spearman ρ and age-adjusted linear regression of log progesterone on $\log_2$ TSH; overt and subclinical hypothyroidism were also compared. Sensitivity analyses excluded 1 extreme control value (67 ng/mL), applied a 25-ng/mL threshold, restricted the sample to the common age range (25-40 years), and modeled age with a natural cubic spline. Each case was then matched to the closest unused control by age (caliper, 2 years), and matched pairs were analyzed using the Wilcoxon signed rank test. In a further post hoc analysis, each of the 95 controls was paired with a distinct case by optimal assignment minimizing total age difference without considering progesterone. Because many cases had the same age, this assignment was repeated with 1000 random case orders to assess sensitivity to tie-breaking. Statistical significance was defined as a 2-sided $P < .05$.

RESULTS

Participant Characteristics

A total of 356 women were assessed for eligibility (Figure 1). Of these, 142 were excluded before enrollment, most often for hormone replacement therapy ($n = 44$) or oral contraceptive use ($n = 36$), and 214 were enrolled. The remaining 62 pre-enrollment exclusions comprised other eligibility reasons.

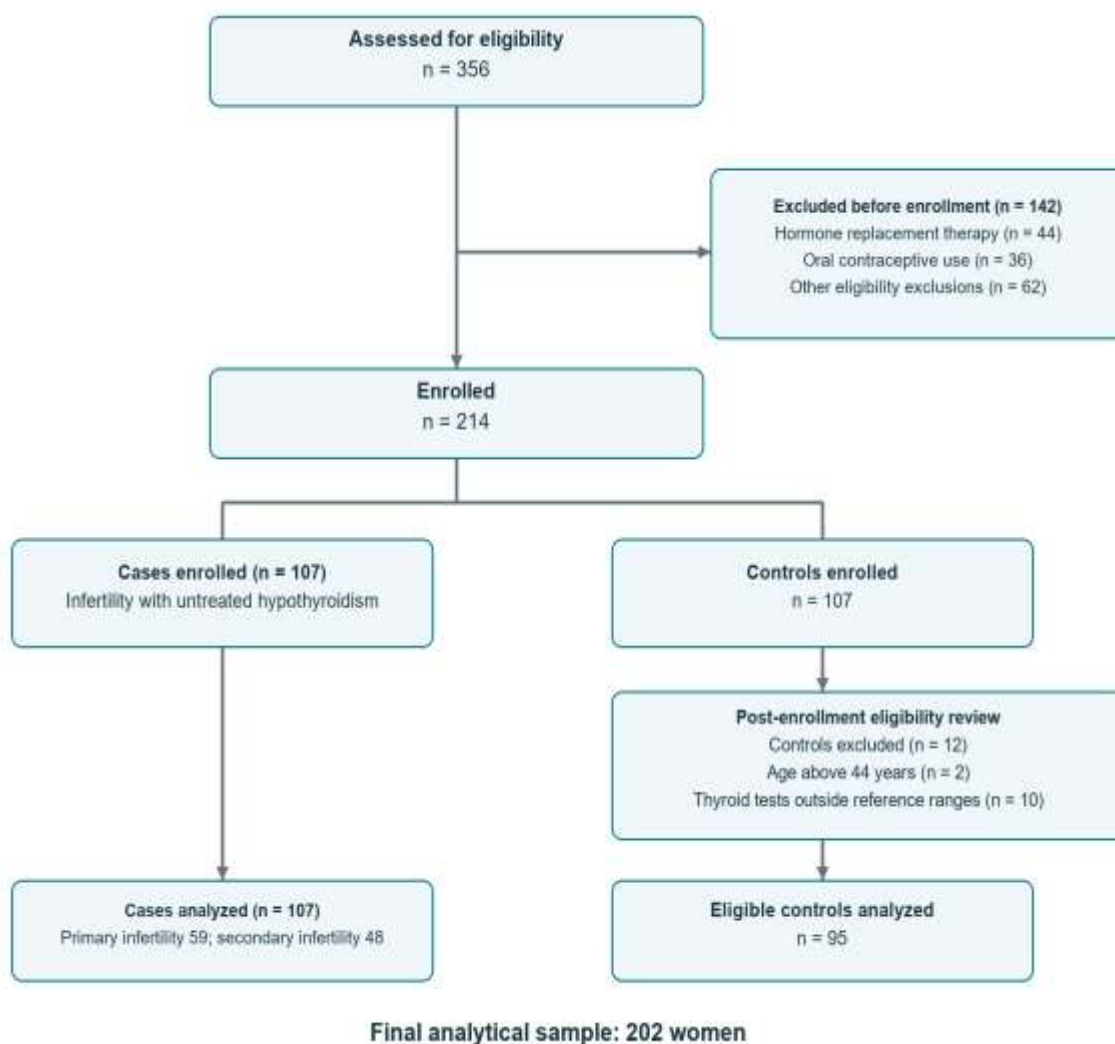


Figure 1. Participant Flow Through Screening, Enrollment, and Analysis. Of 356 women assessed, 142 were excluded before enrollment and 214 were enrolled. Twelve controls were subsequently excluded for ineligibility, leaving 107 cases and 95 controls for analysis. Other pre-enrollment exclusions total 62 women.

After enrollment, 12 controls were excluded as ineligible, leaving 202 for analysis (Figure 1): 107 cases with infertility with coexisting hypothyroidism (59 [55.1%] primary; 48 [44.9%] secondary) and 95 controls with proven fertility. Median (IQR) age was 33 (28–36) years in cases and 31 (29–38) years in controls ($P = .41$; SMD, -0.16). Eleven controls (11.6%) were older than 40 years, the oldest case age. No progesterone values were missing. As expected, cases had higher TSH and lower FT3 and FT4 levels than controls. They also had higher prolactin, AMH, ESR, CRP, and FSH levels and lower LH and estradiol levels, and all nonthyroid differences remained significant after false discovery rate correction (Table 1). Among cases with a valid FT4 value, 41 (39.8%) had overt and 62 (60.2%) had subclinical hypothyroidism. Although active inflammatory conditions were an exclusion criterion, 36 cases (33.6%) had CRP above 10 mg/L, compared with 1 control (1.1%).

Table 1. Age and Serum Markers in Infertile Women With Hypothyroidism and Euthyroid Fertile Controls

Characteristic	Controls (n = 95), median (IQR)	Cases (n = 107), median (IQR)	HL difference (95% CI) ^a	r ^b	SMD	P value ^c	q value ^d
Age, y	31 (29–38)	33 (28–36)	–1.0 (–2.0 to 1.0)	–0.07	–0.16	.41	NA
Progesterone, ng/mL	10.0 (8.0–13.5)	18.5 (15.3–24.8)	8.0 (6.6 to 10.0)	0.62	0.99	<.001	NA
Prolactin, ng/mL	14.5 (12.0–19.2)	24.5 (20.0–27.3)	8.0 (6.5 to 10.0)	0.69	1.47	<.001	<.001
AMH, ng/mL	3.8 (2.9–4.4)	5.5 (3.5–7.0)	1.7 (1.0 to 2.0)	0.46	0.83	<.001	<.001
ESR, mm/h	12.0 (10.0–15.5)	17.0 (11.5–35.0)	5.0 (3.0 to 8.0)	0.41	0.95	<.001	<.001
CRP, mg/L	7.0 (6.0–8.0)	8.5 (6.6–12.8)	1.6 (1.0 to 3.0)	0.38	0.81	<.001	<.001
FSH, mIU/mL	9.6 (7.8–10.8)	10.5 (8.0–12.0)	1.0 (0.1 to 2.0)	0.22	0.21	.006	.007
LH, IU/L	9.6 (7.0–12.7)	5.5 (4.0–15.0)	–3.0 (–4.0 to –1.6)	–0.34	–0.21	<.001	<.001
Estradiol, pg/mL	226 (150–311)	200 (51–300)	–30.5 (–70.0 to 0.0)	–0.17	–0.31	.04	.04
TSH, mIU/L	3.3 (2.2–4.0)	6.0 (5.5–6.8)	3.0 (2.5 to 3.2)	0.99	2.66	<.001	NA
FT3, pg/dL ^e	246 (224–321)	122 (115–128)	–129 (–168 to –120)	–1.00	–3.23	<.001	NA
FT4, ng/dL ^e	2.0 (1.8–2.0)	0.9 (0.7–2.0)	–0.8 (–1.1 to –0.5)	–0.52	–1.11	<.001	NA

Abbreviations: AMH, anti-Müllerian hormone; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; FSH, follicle-stimulating hormone; FT3, free triiodothyronine; FT4, free thyroxine; HL, Hodges-Lehmann; LH, luteinizing hormone; NA, not applicable; SMD, standardized mean difference; TSH, thyrotropin. ^a Difference calculated as case minus control; 95% CI obtained by bootstrap. ^b Rank-biserial correlation coefficient. ^c Calculated using the Mann-Whitney U test. ^d Benjamini-Hochberg adjustment across the 7 nonthyroid secondary markers; thyroid markers were not adjusted because they defined case status. ^e After excluding values with inconsistent units, case counts were 93 for FT3 and 103 for FT4.

Primary Outcome Luteal Progesterone

Median (IQR) progesterone was 18.5 (15.3–24.8) ng/mL in cases and 10.0 (8.0–13.5) ng/mL in controls (HL location shift, 8.0 ng/mL; 95% CI, 6.6–10.0 ng/mL; $P < .001$; $r = 0.62$). The GMR was 1.57 (95% CI, 1.36–1.81), exceeding the smallest GMR detectable with 80% power at the observed variability (1.23). Most controls (68 of 95 [71.6%]) had values of 7 to 15 ng/mL, whereas the reported case distribution showed two concentrations of values: 80 cases (74.8%) had values above 15 ng/mL and 15 (14.0%) had values of 7 ng/mL or less (Figure 2). Progesterone was not correlated with age in either group (cases, $p = 0.11$; controls, $p = 0.03$; both $P \geq .27$).

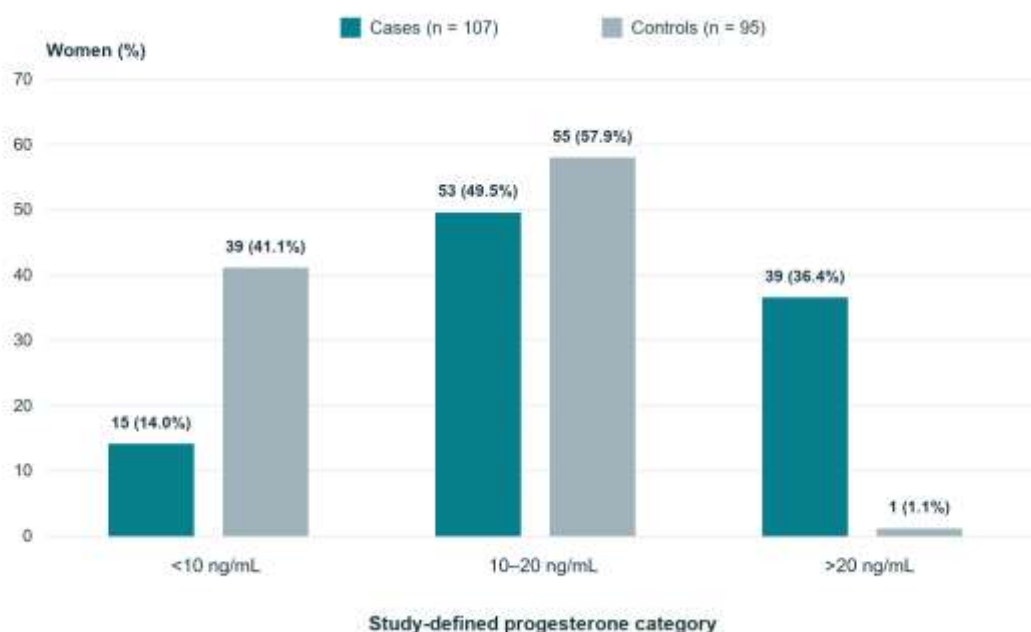


Figure 2. Distribution of Study-Defined Mid-Luteal Progesterone Categories. Bars show the percentage of women in each category among 107 cases and 95 eligible controls. Labels give the numerator and percentage. Categories are descriptive and do not establish luteal deficiency or pathological progesterone excess. Progesterone Above Study-Defined Thresholds

Progesterone above 20 ng/mL occurred in 39 cases (36.4%) and 1 control (1.1%) (Firth OR, 36.3; 95% CI, 9.3–327.6; $P < .001$) (Table 2). The wide CI reflects the single control in this category. Values above 25 ng/mL showed the same pattern.

Table 2. Progesterone Above Study-Defined Thresholds

Threshold	Cases, No./total (%)	Controls, No./total (%)	OR (95% CI) ^a	P value ^b
>20 ng/mL	39/107 (36.4)	1/95 (1.1)	36.3 (9.3–327.6)	<.001
>25 ng/mL	22/107 (20.6)	1/95 (1.1)	16.6 (4.1–151.0)	<.001

Abbreviation: OR, odds ratio.

^a Firth penalized likelihood with profile-likelihood 95% CI.

^b Fisher exact test.

Multivariable Analysis

Continuous progesterone was nonlinearly associated with the log odds of case status (likelihood ratio $\chi^2(2) = 53.4$; $P < .001$). Relative to the intermediate category, high progesterone was strongly associated with case status, whereas low progesterone was less common among cases (Table 3). After adjustment for age and reproductive hormones (model 1), the aORs were 25.4 (95% CI, 4.9–264.5) for high progesterone and 0.49 (95% CI, 0.18–1.25) for low progesterone. Estimates for high progesterone were similar when age was restricted to 25 to 40 years or modeled with a spline. Variance inflation factors were 1.21 or lower. The Hosmer-Lemeshow test showed no evidence of poor fit ($P = .62$), although its power to detect miscalibration is limited. Progesterone alone yielded an AUC of 0.81 (95% CI, 0.74–0.88), indicating group separation within this sample rather than diagnostic utility.

Table 3. Firth Logistic Regression for Case Status by Progesterone Category (Reference, 10-20 ng/mL)

Analysis	Cases/controls, No.	<10 ng/mL, OR (95% CI)	>20 ng/mL, OR (95% CI)
Unadjusted	107/95	0.41 (0.20–0.81)	27.3 (6.9–248.2)
Model 1 ^a	107/95	0.49 (0.18–1.25)	25.4 (4.9–264.5)
Model 2 ^b	107/95	0.19 (0.04–0.71)	41.0 (4.3–756.2)

Analysis	Cases/controls, No.	<10 ng/mL, OR (95% CI)	>20 ng/mL, OR (95% CI)
Model 1, age 25–40 y	107/84	0.54 (0.20–1.40)	21.1 (4.2–217.9)
Model 1, age as spline	107/95	0.55 (0.20–1.46)	20.4 (4.0–209.8)
Model 1, age-balanced subset ^c	95/95	0.40 (0.14–1.07)	24.8 (4.9–255.7)

Abbreviation: OR, odds ratio (adjusted in all rows except the first). Estimates are from Firth penalized logistic regression with profile-likelihood 95% CIs. ^a Adjusted for age, prolactin, luteinizing hormone, follicle-stimulating hormone, anti-Müllerian hormone, and log₂ estradiol. ^b Model 1 plus C-reactive protein and erythrocyte sedimentation rate. ^c Post hoc age-balanced subset: each eligible control paired with a distinct case by optimal assignment minimizing the total age difference (ties broken by record order).

Secondary and Sensitivity Analyses

Among cases, progesterone was not correlated with TSH ($\rho = -0.01$; $P = .96$), FT3, FT4, or other hormones (all $P \geq .32$). The age-adjusted progesterone ratio per doubling of TSH was 1.28 (95% CI, 0.74–2.21), although the estimate was imprecise because the TSH range was narrow (IQR, 5.5–6.8 mIU/L). Progesterone did not differ between overt and subclinical hypothyroidism (median, 18.0 vs 19.0 ng/mL; $P = .78$) or between primary and secondary infertility (18.5 vs 18.3 ng/mL; $P = .58$). Table 4 shows the category distributions.

Table 4. Luteal Progesterone Categories (Mid-Luteal Sampling, 7 Days Before the Next Menses)

Category	Cases, No. (%)	Controls, No. (%)
<10 ng/mL	15 (14.0)	39 (41.1)
10–20 ng/mL	53 (49.5)	55 (57.9)
>20 ng/mL	39 (36.4)	1 (1.1)

In an exploratory comparison, the 15 cases with progesterone below 10 ng/mL had higher ESR (median, 38 vs 16.5 mm/h; $P = .02$) and CRP (12.5 vs 8.0 mg/L; $P = .04$), lower AMH (3.0 vs 5.5 ng/mL; $P = .02$), and similar age (31 vs 33 years; $P = .91$) compared with the other cases. Excluding the extreme control value did not alter the results ($r = 0.64$; $P < .001$). Among 83 age-matched pairs 54 matched exactly on age; median age, 31 years in cases and 30 years in controls median progesterone was 18.0 vs 10.0 ng/mL (paired HL shift, 7.1 ng/mL; 95% CI, 4.8–9.1 ng/mL; $P < .001$). Values above 20 ng/mL occurred in 28 cases and 1 control.

In the post hoc analysis of 95 pairs using every eligible control, 58 pairs were identical in age; median age was 32 years in cases and 31 years in controls. Median progesterone was 19.5 vs 10.0 ng/mL (paired HL shift, 8.1 ng/mL; 95% CI, 6.3–9.8 ng/mL; $P < .001$). Values above 20 ng/mL occurred in 38 cases and 1 control (model 1 aOR, 24.8; 95% CI, 4.9–255.7) (Table 3). Because age tie-breaking determined which 12 cases remained unpaired, the assignment was repeated with 1000 random case orders. Across repetitions, the unpaired HL shift ranged from 7.5 to 8.8 ng/mL (median, 8.0 ng/mL), and 30 to 38 of 95 cases had values above 20 ng/mL, compared with 1 control in every repetition.

DISCUSSION

In this exploratory case-control study of 202 women, untreated women with infertility with coexisting hypothyroidism had higher mid-luteal progesterone than euthyroid women with proven fertility, and values above 20 ng/mL occurred almost exclusively in cases. This difference persisted after adjustment for age and other reproductive hormones, restriction or flexible modeling of age, and age matching. However, the association was nonlinear, and progesterone did not vary with TSH or differ between overt and subclinical hypothyroidism. The direction differs from clinical reports associating thyroid dysfunction with impaired luteal activity. Cases were untreated, hormonal medication was excluded in both groups, and sampling was timed to the recorded onset of the next menses. These precautions reduce some alternative explanations but do not eliminate physiological variability or residual differences in cycle timing. The principal uncertainty is that the groups differed in fertility as well as thyroid status.

Values above 20 ng/mL should not independently be considered pathological. Mid-luteal progesterone varies widely in normally ovulatory women and can exceed this threshold. A single concentration can provide presumptive evidence of recent ovulation but cannot assess luteal adequacy (19). All but 1 of the 202 participants had values of at least 3 ng/mL; this observation does not provide independent confirmation of

ovulation or establish normal luteal function. Similarly, values below 10 ng/mL—observed in 41.1% of controls and 14.0% of cases—are descriptive and do not establish luteal deficiency.

Available clinical evidence is heterogeneous and does not provide a directly equivalent comparison. Among 304 infertile women evaluated with thyrotropin-releasing hormone testing, 37 had subclinical hypothyroidism; of these, 8 had luteal phase deficiency and 3 were anovulatory, and levothyroxine increased mean luteal progesterone (26). In a prospective cohort of 86 euthyroid women, lower total thyroxine was associated with lower maximum luteal urinary pregnanediol glucuronide (8.1 vs 11.7 µg/mg creatinine for low vs high total thyroxine) (27). A major review links hypothyroidism with oligomenorrhea, altered sex hormone-binding globulin and sex steroid levels, and reduced fertility (28). Experimental findings are mixed: T₃ stimulated progesterone release from human luteal cells (20), high TSH impaired luteal cell survival without changing progesterone output (21), and T₃ had no direct effect on rat luteal steroidogenesis (22). A retrospective study found no significant progesterone difference between women with polycystic ovary syndrome with and without subclinical hypothyroidism; however, samples were obtained in the early follicular phase or an equivalent amenorrheic window, limiting comparison with mid-luteal sampling (29). The human studies considered here do not establish increased natural-cycle luteal progesterone as a consistent feature of hypothyroidism. The absent TSH gradient within cases does not support a dose-response relationship in this sample, although the narrow TSH range limits that analysis.

Because this finding conflicts with previous literature, alternative explanations should be considered. Treatment is unlikely: cases were untreated for hypothyroidism, and women using ovulation induction agents, progesterone, or other hormonal medication were excluded. Thus, levothyroxine, which increased luteal progesterone in an earlier clinical study (26), clomiphene-based ovarian stimulation, which increased reproductive hormone concentrations (30), and exogenous progesterone supplementation (24) are less plausible explanations under the reported medication exclusions. Timing collection relative to the subsequent menstrual period reduced cycle-phase uncertainty; however, the absence of independently verified ovulation and serial measurements prevents exclusion of residual timing differences and pulsatile variation. The main alternative explanation is the comparison group. Cases were infertile, whereas controls had proven fertility; therefore, the design cannot distinguish effects of hypothyroidism from those of infertility or its other causes, and reproductive history differed by design. The contrast may also reflect low control values as much as high case values, because 41.1% of controls had progesterone below 10 ng/mL. A similarly sampled euthyroid infertile group is needed to resolve this issue.

Several biological mechanisms could be investigated, although none was measured in this study or establishes an explanation for the observed difference. First, altered progesterone clearance could raise circulating concentrations without increased luteal secretion. The present data cannot distinguish secretion from clearance, and cortisol kinetics cannot be assumed to represent progesterone metabolism. Thyroid status is associated with sex hormone-binding globulin and sex steroid concentrations (18, 28), but these findings do not directly establish altered progesterone binding or clearance. Second, corpus luteum function may be prolonged in some experimental settings. In pregnant hypothyroid rats, luteal 20α-hydroxysteroid dehydrogenase was reduced; luteolytic prostaglandin F_{2α} and LH decreased, prolactin increased, and luteal and circulating progesterone remained elevated (31). However, prolactin is luteotropic in rodents, whereas hyperprolactinemia generally impairs reproductive function in women (14), limiting extrapolation across species and from pregnancy to natural nonpregnant cycles. Third, very high TSH may stimulate the ovary through cross-reactivity with gonadotropin pathways. Case reports describe multicystic ovaries and spontaneous ovarian hyperstimulation in severe primary hypothyroidism, with improvement after levothyroxine (32, 33). Multiple follicles and corpora lutea are a possible hypothesis, but ovarian morphology and follicular development were not assessed here, and case TSH values had a narrow range without a progesterone gradient. Finally, experimental hypothyroidism alters uterine progesterone signaling in rats (7), and experimental T₃ signaling supports endometrial receptivity (6). Whether a comparable disturbance in endometrial response occurred in these women remains untested. Testing these mechanisms will require a euthyroid infertile comparison group, serial luteal sampling, measurement of progesterone metabolites and binding proteins, and appropriately designed endometrial assessment.

Other hormonal findings were mixed. Higher prolactin in cases was physiologically plausible (14), although a Riyadh cohort found no statistically significant association between hyperprolactinemia and hypothyroidism (15). Higher AMH contrasted with a report linking higher TSH to lower AMH (16), while age and thyroid-related factors complicate interpretation of ovarian reserve markers (17). PCOS could not be assessed and remains a potential source of residual confounding; lower LH concentrations do not exclude it. Although one cohort reported similar thyroid-disease prevalence in women with PCOS and controls (34), that population-level finding cannot resolve PCOS status in the present sample. The subgroup with progesterone below 10 ng/mL had higher

inflammatory markers and lower AMH, but these nonspecific findings do not establish autoimmune thyroiditis. Thyroid autoimmunity may affect reproductive outcomes (13), yet antibodies were not measured in cases. Moreover, 33.6% of cases had CRP above 10 mg/L despite exclusion of active inflammatory conditions, suggesting possible unrecognized inflammation. Because prolactin and other reproductive hormones may lie on pathways connecting thyroid status and luteal function, their inclusion in adjusted models may introduce overadjustment rather than isolate a causal thyroid effect.

These findings do not support using progesterone as a clinical marker of thyroid-related infertility. Evidence linking elevated progesterone to lower live-birth rates comes from measurements obtained at ovulation trigger during stimulated assisted reproduction cycles (23) and cannot be extrapolated to luteal-phase sampling in natural cycles. Similarly, although the TABLET trial found that levothyroxine did not improve live-birth rates in euthyroid women with thyroid peroxidase antibodies (10), it does not show that treating established hypothyroidism fails to improve fertility; current guidance recommends achieving euthyroidism before conception (8). Clinically, high mid-luteal progesterone in women with hypothyroidism should therefore be interpreted in the context of cycle timing and current treatment.

Limitations and Strengths

This study has several limitations. Because controls had proven fertility and no euthyroid infertile group was included, the between-group contrast combines differences in fertility, thyroid status, and reproductive history. Recruitment through different clinical and institutional routes may also introduce selection bias. Twelve controls were excluded after enrollment for ineligibility, indicating incomplete initial screening. Controls were not age-matched at recruitment; restriction, flexible age modeling, and post hoc age balancing produced consistent estimates but do not reproduce a prospectively matched design. Ovulation was not independently confirmed by ultrasonography or urinary LH testing. Retrospective confirmation of subsequent menses reduces timing uncertainty but does not establish that every sample captured the same interval after ovulation. A single progesterone measurement cannot separate persistent elevation from pulsatile or cycle-to-cycle variation. Body mass index was unavailable, PCOS could not be assessed, and thyroid antibodies were not measured in cases, leaving important sources of residual confounding. One-third of cases had CRP above 10 mg/L despite exclusion of active inflammatory conditions. Eighteen thyroid measurements with uncertain units were excluded, limiting thyroid-profile interpretation and leaving four cases without subtype classification. The case TSH threshold of 4.9 mIU/L overlapped slightly with the control reference interval extending to 5.0 mIU/L, and thyroid subtype classification was post hoc; borderline biochemical classification may therefore be uncertain. Absence of assay-specific platform and precision information limits assessment of analytical reproducibility and measurement error. Progesterone thresholds, covariates, and sensitivity analyses were selected during analysis. Sparse high-progesterone control counts produced wide odds-ratio confidence intervals, and additional matching analyses remain exploratory. The case-control design cannot establish temporality, and findings from a single South Asian hospital may not generalize to other populations.

Strengths include untreated cases, exclusion of hormonal medication in both groups, sampling timed relative to the subsequent menses, analysis in a single laboratory, reporting of effect estimates with confidence intervals, nonparametric and penalized estimation, and multiple sensitivity analyses. These features support further investigation of the observed group difference without establishing its cause or clinical applicability.

CONCLUSIONS

Infertile women with untreated hypothyroidism had higher single mid-luteal serum progesterone concentrations than euthyroid women with proven fertility. Because the groups differed in both thyroid status and fertility, the observed difference cannot be attributed specifically to hypothyroidism. These findings do not establish pathological progesterone excess or support diagnostic use. Further studies should include euthyroid infertile comparators, independently characterized ovulation timing, repeated luteal measurements, and assessment of endometrial response.

Article Information

Author Contributions

Saman Saeed Ansari: concept and design; acquisition, analysis, or interpretation of data; drafting of the manuscript; critical review of the manuscript for important intellectual content; statistical analysis.

Tashfeen Ikram: concept and design; acquisition, analysis, or interpretation of data; drafting of the manuscript; critical review of the manuscript for important intellectual content.

Aisha Bashir: concept and design; statistical analysis, interpretation of data; drafting of the manuscript; critical review of the manuscript for important intellectual content.

Qanita Mahmud: concept and design; acquisition, analysis, or interpretation of data; drafting of the manuscript; critical review of the manuscript for important intellectual content; statistical analysis.

Ambreen Anjum: concept and design; acquisition, analysis, or interpretation of data; drafting of the manuscript; critical review of the manuscript for important intellectual content; statistical analysis.

Tahir Maqbool: concept and design; drafting of the manuscript; critical review of the manuscript for important intellectual content.

Muhammad Tahir: concept and design; drafting of the manuscript; critical review of the manuscript for important intellectual content.

Dr Saman Saeed Ansari had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of Interest Disclosures: The authors have no competing or conflicting interests to declare.

Funding/Support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Sharing Statement: Deidentified participant data and the analysis code are available from the corresponding author on reasonable request.

AI Declaration: Generative AI tools were used solely for language and grammatical refinement. The authors verified the final manuscript and remain fully responsible for its accuracy, integrity, and scientific content.

REFERENCES

1. Cox CM, Thoma ME, Tchangalova N, et al. Infertility prevalence and the methods of estimation from 1990 to 2021: a systematic review and meta-analysis. *Hum Reprod Open*. 2022;2022(4):hoac051. doi:10.1093/hropen/hoac051
2. Chaker L, Papaleontiou M. Hypothyroidism: a review. *JAMA*. 2025;334(19):1750-1760. doi:10.1001/jama.2025.13559
3. Ojha PR, Kumar A, Kumar R. Pooled prevalence of hypothyroidism among Indian females with infertility: a systematic review & meta-analysis. *Indian J Med Res*. 2024;159(6):627-636. doi:10.25259/IJMR_987_23
4. Arif S, Nasrullah FD, Jaleel R, Hasnain FSU, Abbasi AU. Frequency of subclinical hypothyroidism in subfertile females and the effect of thyroxine replacement. *J Coll Physicians Surg Pak*. 2025;35(11):1413-1417. doi:10.29271/jcpsp.2025.11.1413
5. Concepción-Zavaleta MJ, Coronado-Arroyo JC, Quiroz-Aldave JE, Concepción-Urteaga LA, Paz-Ibarra J. Thyroid dysfunction and female infertility: a comprehensive review. *Diabetes Metab Syndr*. 2023;17(11):102876. doi:10.1016/j.dsx.2023.102876
6. Zhang R, Peng M, Liu T, et al. T3-THR signaling governed by DIO2 contributes to endometrial receptivity by regulating epithelial cell membrane fluidity. *Endocrinology*. 2025;166(6):bqaf070. doi:10.1210/endocr/bqaf070
7. da Silva TQM, Barbosa EM, Santos LC, et al. Hypothyroidism alters uterine kisspeptin system and activity modulators in cyclic rats. *Int J Mol Sci*. 2025;26(2):543. doi:10.3390/ijms26020543
8. Korevaar TIM, Leung AM, Alexander EK, et al. American Thyroid Association 2026 guidelines for thyroid disease in preconception, pregnancy, and postpartum. *Thyroid*. 2026;36(5):481-544. doi:10.1177/10507256261445624
9. Unuane D, Velkeniers B, Poppe KG. Thyroid disorders and female infertility. *J Clin Endocrinol Metab*. 2026;111(5):e1239-e1251. doi:10.1210/clinem/dgag039
10. Dhillon-Smith RK, Middleton LJ, Sunner KK, et al. Levothyroxine in women with thyroid peroxidase antibodies before conception. *N Engl J Med*. 2019;380(14):1316-1325. doi:10.1056/NEJMoa1812537
11. Gill S, Cheed V, Morton VAH, et al. Evaluating the progression to hypothyroidism in preconception euthyroid thyroid peroxidase antibody-positive women. *J Clin Endocrinol Metab*. 2023;108(1):124-134. doi:10.1210/clinem/dgac525
12. Provinciatto H, Moreira MVB, Neves GR, et al. Levothyroxine for subclinical hypothyroidism during pregnancy: an updated systematic review and meta-analysis of randomized controlled trials. *Arch Gynecol Obstet*. 2024;309(6):2387-2393. doi:10.1007/s00404-024-07512-3
13. Bucci I, Giuliani C, Di Dalmazi G, Formoso G, Napolitano G. Thyroid autoimmunity in female infertility and assisted reproductive technology outcome. *Front Endocrinol (Lausanne)*. 2022;13:768363. doi:10.3389/fendo.2022.768363

14. Haidenberg-David F, Sidauy-Adissi J, Moscona-Nissan A, et al. Overview of hyperprolactinemia: general approach and reproductive health implications. *Arch Med Res.* 2024;55(8):103102. doi:10.1016/j.arcmed.2024.103102
15. Alwahhabi AY, Labban RH, Alshaikh MY, et al. Prevalence of hyperprolactinemia and its association with menstrual abnormalities and hypothyroidism among infertile women: a cross-sectional study from Saudi Arabia. *Saudi Med J.* 2026;47(7):1240-1245. doi:10.15537/1658-3175.8813
16. Amidi B, Naderi N, Ghafarzadeh M, Abasinezhad A, Yari F, Kaviani Charati M. Impact of hypothyroidism on ovarian reserve in infertile reproductive-aged women: a cross-sectional study. *J Obstet Gynaecol India.* 2026;76(3):206-212. doi:10.1007/s13224-025-02147-z
17. Giusti M, Mittica M. Evaluation of anti-Müllerian hormone in pre-menopausal women stratified according to thyroid function, autoimmunity and age. *Thyroid Res.* 2022;15(1):15. doi:10.1186/s13044-022-00133-5
18. Kjaergaard AD, Marouli E, Papadopoulou A, et al. Thyroid function, sex hormones and sexual function: a Mendelian randomization study. *Eur J Epidemiol.* 2021;36(3):335-344. doi:10.1007/s10654-021-00721-z
19. Practice Committee of the American Society for Reproductive Medicine; Practice Committee of the Society for Reproductive Endocrinology and Infertility. Diagnosis and treatment of luteal phase deficiency: a committee opinion. *Fertil Steril.* 2026;126(4):633-639. doi:10.1016/j.fertnstert.2026.06.014
20. Datta M, Roy P, Banerjee J, Bhattacharya S. Thyroid hormone stimulates progesterone release from human luteal cells by generating a proteinaceous factor. *J Endocrinol.* 1998;158(3):319-325. doi:10.1677/joe.0.1580319
21. Taggi M, Capponi C, Bertani N, et al. Role of thyroid stimulating hormone in the maintenance and functioning of the human corpus luteum. *J Endocrinol Invest.* 2024;47(7):1719-1732. doi:10.1007/s40618-023-02269-z
22. Navas PB, Redondo AL, Cuellar-Carrión FD, et al. Luteal expression of thyroid hormone receptors during gestation and postpartum in the rat. *Thyroid.* 2014;24(6):1040-1050. doi:10.1089/thy.2013.0332
23. Lim YC, Hamdan M, Maheshwari A, Cheong Y. Progesterone level in assisted reproductive technology: a systematic review and meta-analysis. *Sci Rep.* 2024;14(1):30826. doi:10.1038/s41598-024-81539-z
24. Grove-Laugesen D, Saupstad M, Gabrielsen AV, et al. Effect of luteal phase support with progesterone on thyroid function: a sub-study of the FET optimizing trial. *Eur Thyroid J.* 2026;15(4):ETJ260099. doi:10.1530/ETJ-26-0099
25. 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. *PLoS Med.* 2007;4(10):e296. doi:10.1371/journal.pmed.0040296
26. Bispink L, Brändle W, Lindner C, Bettendorf G. Preclinical hypothyroidism and disorders of ovarian function [in German]. *Geburtshilfe Frauenheilkd.* 1989;49(10):881-888. doi:10.1055/s-2008-1036104
27. Jacobson MH, Howards PP, Darrow LA, et al. Thyroid hormones and menstrual cycle function in a longitudinal cohort of premenopausal women. *Paediatr Perinat Epidemiol.* 2018;32(3):225-234. doi:10.1111/ppe.12462
28. Krassas GE, Poppe K, Glinioer D. Thyroid function and human reproductive health. *Endocr Rev.* 2010;31(5):702-755. doi:10.1210/er.2009-0041
29. Hua R, Li JJ, Yang W, Hua Y, Ma YL. Impact of subclinical hypothyroidism on endocrine features in patients with polycystic ovary syndrome. *Eur J Med Res.* 2025;30(1):725. doi:10.1186/s40001-025-02910-y
30. Montenegro IS, Kuhl CP, Schneider RA, et al. Use of clomiphene citrate protocol for controlled ovarian stimulation impairs endometrial maturity. *JBRA Assist Reprod.* 2021;25(1):90-96. doi:10.5935/1518-0557.20200056
31. Hapon MB, Motta AB, Ezquer M, Bonafede M, Jahn GA. Hypothyroidism prolongs corpus luteum function in the pregnant rat. *Reproduction.* 2007;133(1):197-205. doi:10.1530/REP-06-0035
32. Guerra M, Marado D, Silva F, Almeida MC. Severe primary hypothyroidism and ovarian hyperstimulation syndrome in a spontaneous pregnancy: a case report. *AME Case Rep.* 2024;8:26. doi:10.21037/acr-23-13
33. Zhuang W, Wu L, Shao R, Zhang K, Ye L. Natural conception complicated by spontaneous ovarian hyperstimulation syndrome in the setting of severe primary hypothyroidism: a case series. *Medicine (Baltimore).* 2025;104(41):e44825. doi:10.1097/MD.00000000000044825
34. van der Ham K, Stekelenburg KJ, Louwers YV, et al. The prevalence of thyroid dysfunction and hyperprolactinemia in women with PCOS. *Front Endocrinol (Lausanne).* 2023;14:1245106. doi:10.3389/fendo.2023.1245106