Stout in Medical Science, Biology and Chemistry (SMSBC)
ISSN: 2979-2819
DOI: https://doi.org/10.61424/xn3vae15
Journal Homepage: https://stoutjournals.org/index.php/SMSBC

Diagnostic Quality of the Laboratory Work-up for Menstrual Disorders: Hormonal Timing, Appropriateness of Prescriptions, Reclassification and Clinico-therapeutic Impact — A Real-world, Prospective, Observational, Multicentre Study in 2,000 Women Aged 10 to 45 years

Leila Chouali
Faculty of Medicine, Ferhat Abbas Sétif-1 University, Sétif, Algeria; Laboratory of Applied Pediatric Surgery, Ferhat Abbas Sétif-1 University, Sétif, Algeria; Department of Obstetrics and Gynecology, Sétif University Hospital, Sétif, Algeria
✉ **Corresponding Author**: Leila Chouali, **E-mail**: leila_medgyn@yahoo.com

ARTICLE INFORMATION

Submitted: July 10, 2026
Accepted: August 29, 2026
Published: September 07, 2026
Volume: 2
Issue: 1

KEYWORDS

Menstrual disorders; diagnostic stewardship; hormonal testing; hormonal timing; diagnostic quality; adolescents; diagnostic reclassification; PCOS

ABSTRACT

Menstrual disorders are common in adolescents and women of reproductive age. Laboratory diagnostic value depends on clinical context, age, hormonal status, sampling timing and interpretation. This study evaluated laboratory work-up quality and its diagnostic and therapeutic impact. A prospective multicentre observational study was conducted from January 2024 to June 2026 in five healthcare facilities in eastern Algeria. Women aged 10–45 years presenting with menstrual disorders were included. Clinical, ultrasound and laboratory data were collected prospectively. Test appropriateness was assessed according to indication, timing, interpretation and clinical contribution. Multivariable and multilevel analyses were performed. Among 2,000 participants, 1,894 (94.7%) underwent laboratory testing. Biological abnormalities were identified in 56.3%, but only 37.4% were clinically contributive. Inappropriate prescription or interpretation occurred in 21.9%, while clinico-biological discordance occurred in 20.3%. Hormonal timing was appropriate in 58.4% of evaluable cases. An unfavourable DLQC was observed in 36.9%. Diagnostic reclassification occurred in 10.5% and therapeutic changes in 26.1%. Adolescent age, extensive testing and inadequate cycle-timing documentation were the main factors associated with poorer diagnostic quality. Laboratory quality in menstrual disorders depends primarily on clinical relevance, appropriate timing and contextual interpretation, rather than test volume. An age-adapted diagnostic-stewardship strategy may reduce low-value testing, inappropriate diagnoses and diagnostic delays while improving clinical decision-making.

1. Introduction

Menstrual disorders are a frequent reason for consultation in gynaecology, endocrinology and primary care. Their diverse aetiologies require an individualised diagnostic approach based on age, menstrual phenotype, associated clinical features and hormonal context. Laboratory testing is an important component of this evaluation; however, its diagnostic value depends on appropriate indications, timing of sampling, analytical methods and contextual interpretation. Inappropriate or excessive testing may generate uninterpretable results, repeated investigations, diagnostic cascades and inappropriate therapeutic decisions. Hormonal interpretation is particularly sensitive to physiological context. FSH, LH and estradiol should be interpreted according to reproductive status and, when relevant, cycle phase, while progesterone should be assessed according to the presumed timing of ovulation. Unexpected prolactin elevations may require confirmation and contextual reassessment, and androgen interpretation depends partly on assay quality. Adolescents represent a specific challenge, as menstrual irregularity may be physiological during the first years after menarche. Recent recommendations therefore emphasise age- and maturation-specific criteria, particularly for polycystic ovary syndrome (PCOS). In adolescents, PCOS diagnosis requires persistent ovulatory dysfunction combined with hyperandrogenism after exclusion of other causes; ovarian morphology and AMH should not be used alone for diagnosis. Similarly, low-value investigations such as isolated fasting insulin measurement should not be considered central diagnostic tests for menstrual disorders.

Thus, a distinction should be made between a biological abnormality, a clinically meaningful abnormality, and a result that actually modifies diagnosis or management. This perspective is consistent with diagnostic stewardship, which promotes the appropriate test, in the appropriate patient, at the appropriate time, with appropriate interpretation. However, real-world data assessing prescription appropriateness, hormonal timing, interpretation, clinico-biological concordance and the impact of laboratory testing on diagnosis and management remain limited, particularly in intermediate-resource healthcare settings.

The objective of this study was to determine the prevalence and determinants of inappropriate laboratory investigations in women aged 10–45 years presenting with menstrual disorders, and to evaluate their appropriateness, timing, diagnostic yield, clinico-biological concordance, low-value testing, and impact on diagnostic and therapeutic decisions.

Main hypothesis: Inappropriate laboratory testing and interpretation are common and may compromise diagnostic efficiency, leading to unnecessary investigations, diagnostic cascades and delayed diagnosis.

2. Materials and Methods

2.1. Study design and period

This was a prospective, multicentre analytical study conducted from January 2024 to June 2026 in five public healthcare facilities in eastern Algeria. Participants were consecutively recruited during consultation for menstrual disorders and followed from initial assessment to final diagnosis, re-evaluation or initial treatment. The study was conducted in accordance with STROBE recommendations

2.2. Study population

The population comprised 2,000 participants aged 10 to 45 years consulting for a menstrual disorder requiring clinical evaluation, with or without biological investigation.

Three age groups were defined a priori:

- 10–19 years: adolescents;
- 20–35 years: women of reproductive age;
- 36–45 years: advanced reproductive age.

Distribution by age and facility

Facility	10–19 years	20–35 years	36–45 years	Total
CHU Sétif	120	300	180	600
EHS El Eulma	80	200	120	400
EPSP Sétif	70	175	105	350
EPSP El Eulma	60	150	90	300
EPH BBA	70	175	105	350
Total	**400**	**1,000**	**600**	**2,000**

This distribution made it possible to study simultaneously the effect of age and that of the level of care.

2.3. Participants and data collection

Women aged 10–45 years with menstrual disorders and sufficient data for the primary endpoint were included. Pregnancy-related disorders, established menopause, insufficient records, and hormonal treatment precluding reliable interpretation were excluded.

Demographic, anthropometric, gynaecological, clinical, laboratory and therapeutic data were collected prospectively, including age, time since menarche, BMI, menstrual phenotype, hormonal contraception and relevant endocrine features.

2.4. Laboratory assessment and hormonal timing

Laboratory tests were prescribed according to clinical indication. Hormonal, thyroid, androgenic and metabolic parameters were assessed when appropriate. Fasting insulin was not considered a routine diagnostic test for PCOS. Ultrasound was performed when clinically indicated; AMH and ovarian morphology were not used alone to diagnose PCOS in adolescents.

For cycle-dependent assays, menstrual timing, hormonal treatment and cycle characteristics were recorded and classified as appropriate, potentially inappropriate or non-evaluable.

2.5. Assessment of diagnostic quality

Test appropriateness was assessed using a standardised grid covering indication, test selection, timing, physiological context, analytical method, interpretation and clinical contribution.

The primary endpoint was the Diagnostic Laboratory Quality Composite (DLQC), based on six domains: indication, test selection, timing, contextual interpretation, clinico-biological concordance and clinical consequence. Each domain was scored 0–1 (range 0–6); 5–6 indicated favourable quality and 0–4 unfavourable quality.

Secondary outcomes included diagnostic yield, clinically contributive abnormalities, clinico-biological discordance, low-value testing, diagnostic cascades, diagnostic delay, Diagnostic Reclassification Rate (DRR), and diagnostic and therapeutic impact.

2.6. Quality control and statistical analysis

Consecutive recruitment, standardised data collection and predefined criteria were used to limit bias. A subsample was independently reassessed by two investigators, with agreement assessed using Cohen's kappa. Continuous and categorical variables were analysed using appropriate statistical tests. Multivariable logistic regression identified factors associated with unfavourable DLQC, with adjusted ORs, 95% CIs and p-values. Sensitivity analyses were performed according to hormonal timing, treatment and age group.

2.7. Ethical considerations, funding and conflicts of interest

The study was conducted after approval by the relevant ethics committee, in accordance with applicable institutional regulations. Data were pseudonymised and kept confidential. Consent was obtained from adult participants and, in minors, according to procedures combining consent from the legal representative and the adolescent's assent. No specific funding was received for this study. The authors declare no conflicts of interest.

3. Results

3.1. Study population and baseline characteristics

Of the 2,143 patients assessed during the study period, 2,000 met the predefined eligibility criteria and were included in the final analysis. The cohort comprised 400 adolescents aged 10–19 years (20.0%), 1,000 women aged 20–35 years (50.0%), and 600 women aged 36–45 years (30.0%). The mean age was 28.9 ± 9.2 years. Participants were recruited from five healthcare facilities in eastern Algeria. A laboratory work-up was performed in 1,894 participants (94.7%).

The main clinical characteristics are presented in Table 1. Overall, 493 participants (24.7%) had a BMI ≥30 kg/m^2, 402 (20.1%) had hirsutism, 361 (18.1%) had acne, 121 (6.1%) reported galactorrhoea, and 214 (10.7%) had acanthosis nigricans. Hormonal contraception was reported in 387 participants (19.4%).

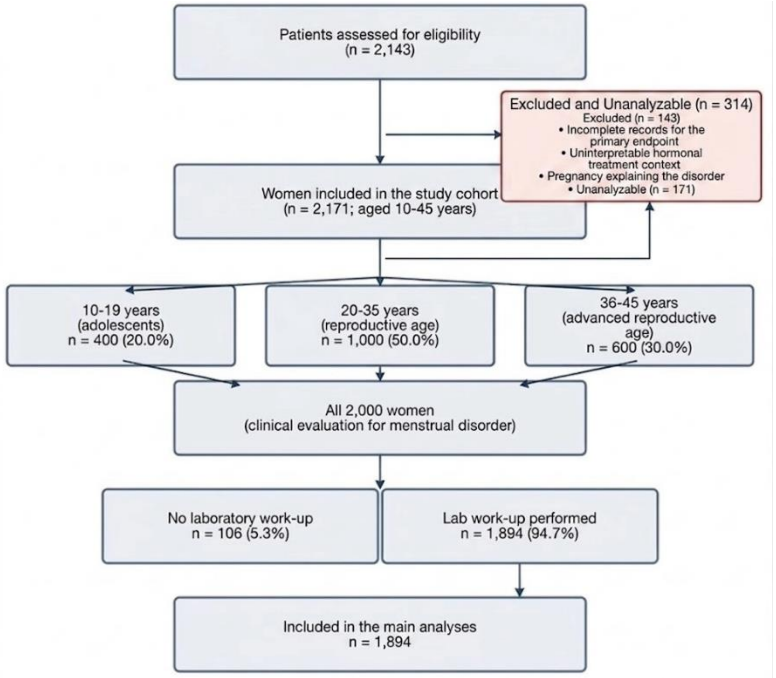

Figure 1. Flow chart of the patients included and analysed.

Table 1. Baseline demographic and clinical characteristics of the study population

Characteristic	n (%)
Total population	2,000 (100)
Age, mean ± SD	28.9 ± 9.2 years
10–19 years	400 (20.0)
20–35 years	1,000 (50.0)
36–45 years	600 (30.0)
BMI ≥30 kg/m²	493 (24.7)
Menarche ≤12 years	1,118 (55.9)
Hirsutism	402 (20.1)
Acne	361 (18.1)
Galactorrhoea	121 (6.1)
Acanthosis nigricans	214 (10.7)
Hormonal contraception	387 (19.4)

3.2. Menstrual phenotype and laboratory testing burden

The most frequent menstrual presentations were oligo/amenorrhoea, observed in 704 participants (35.2%), and irregular cycles, observed in 658 (32.9%). Hypermenorrhoea or menorrhagia was reported in 438 participants (21.9%), while dysmenorrhoea was associated with the menstrual disorder in 124 (6.2%) (Table 2).

Table 2. Menstrual phenotypes in the study population

Menstrual phenotype	n (%)
Oligo/amenorrhoea	704 (35.2)
Irregular cycles	658 (32.9)
Hypermenorrhoea/menorrhagia	438 (21.9)
Associated dysmenorrhoea	124 (6.2)
Other disorders	76 (3.8)

Among the 1,894 participants who underwent laboratory testing, the median number of tests prescribed per patient was 5 [IQR, 3–8]. A total of 811 patients (42.8%) underwent at least six tests. The distribution of testing burden is shown in Table 3.

Table 3. Laboratory testing burden

Number of tests prescribed	n (%)
1–3	512 (27.0)
4–5	571 (30.1)
6–8	489 (25.8)
≥9	322 (17.0)
Total	**1,894 (100)**

3.3. Laboratory testing patterns and appropriateness

TSH (85.7%) and prolactin (81.5%) were the most frequently prescribed tests, followed by FSH (70.8%), LH (67.9%), estradiol (61.9%), and testosterone (53.7%). Insulinaemia and AMH were requested in 30.7% and 27.0%, respectively (Table 4).

Table 4. Laboratory tests prescribed

Laboratory test	n (%)
TSH	1,622 (85.7)
Prolactin	1,544 (81.5)
FSH	1,341 (70.8)
LH	1,286 (67.9)
Estradiol	1,172 (61.9)
Testosterone	1,018 (53.7)
SHBG	624 (33.0)
DHEAS	493 (26.0)
17-OHP	321 (16.9)
AMH	512 (27.0)
Progesterone	447 (23.6)
Insulinaemia	581 (30.7)
HbA1c	744 (39.3)

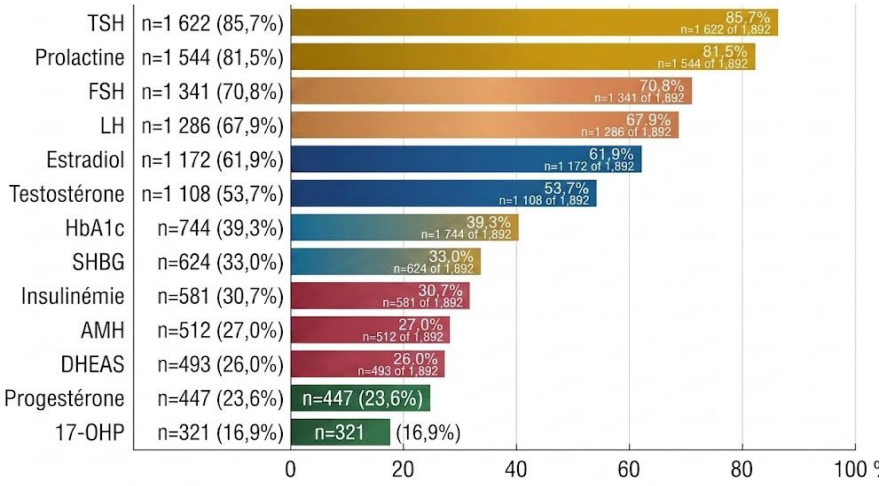

Figure 2. Profile of performed hormonal and metabolic assays

A clear clinical indication was documented in 76.9% of work-ups, while 23.1% included at least one potentially inappropriate prescription. The main issues were excessive testing, unclear indications, extensive hormonal panels, unjustified insulinaemia testing, and repeated tests (Table 5).

Table 5. Inappropriate laboratory testing and interpretation

Type of inappropriate testing/interpretation	n (%)
Excessive number of tests	247 (13.0)
Test without clearly documented indication	218 (11.5)
Excessive hormonal panel	189 (10.0)
Insulinaemia without specific indication	176 (9.3)
Unjustified repetition	164 (8.7)
Undocumented cycle phase	306 (15.3)
Incorrect timing	247 (12.4)
Isolated biomarker interpretation	189 (9.5)
Excessive androgen panel	97 (4.9)
Inappropriate AMH use in adolescents	63 (3.2)

Among patients undergoing cycle-dependent hormonal testing, timing was appropriately documented in 58.4%, potentially inadequate in 26.3%, and undocumented in 15.3%. Appropriate timing was lower among adolescents (51.8%) than among women aged 20–35 (61.7%) and 36–45 years (60.2%) (Table 6)

Table 6. Appropriateness of hormonal timing according to age group

Age group	Appropriate timing	Inadequate/undocumented timing
10–19 years	51.8%	48.2%
20–35 years	61.7%	38.3%
36–45 years	60.2%	39.8%
Overall	**58.4%**	**41.6%**

Correct interpretation was observed in 75.3% of work-ups, while 24.7% showed potentially inadequate interpretation. The main issues involved cycle timing, reference intervals, isolated biomarkers, hormonal contraception, and borderline values. Clinico-biological discordance occurred in 20.3%, particularly with extensive testing and poorly documented hormonal timing or clinical indications.

3.4. Biological abnormalities and final diagnoses

A biological abnormality was identified in 1,126/2,000 participants (56.3%), but only 748 (37.4%) had a clinically contributive abnormality. PCOS was the most frequent final diagnosis (552, 27.6%), while 459 (23.0%) had no definitive cause identified (Table 7).

Table 7. Final etiological diagnoses

Final diagnosis	n (%)
PCOS	552 (27.6)
Hyperprolactinaemia	168 (8.4)
Dysthyroidism	153 (7.7)
Hypothalamic dysfunction	142 (7.1)
Uterine/ovarian pathology	151 (7.6)
Primary ovarian insufficiency	82 (4.1)
Non-PCOS hyperandrogenism	76 (3.8)
Iatrogenic cause	71 (3.6)
Other causes	146 (7.3)
No cause identified	459 (23.0)

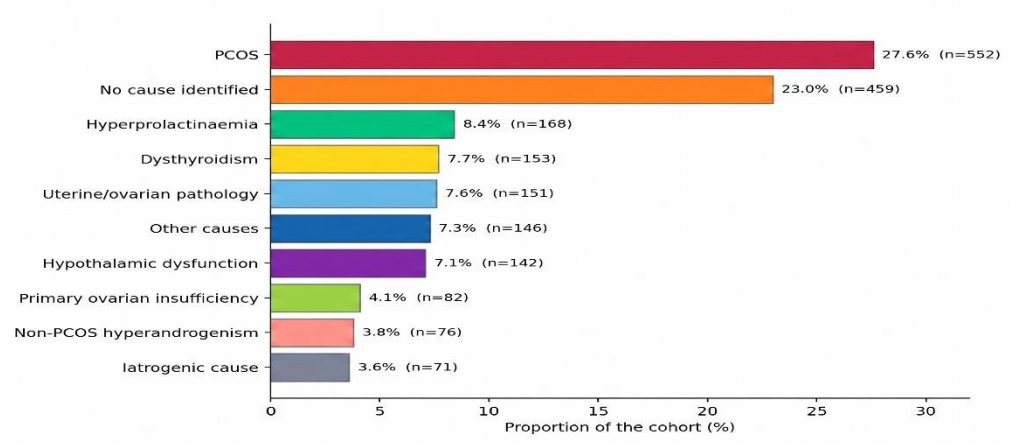

Figure 3: Breakdown of identified etiologies in the study

PCOS was most frequently diagnosed among women aged 20–35 years (326, 32.6%), compared with 96 adolescents (24.0%) and 130 women aged 36–45 years (21.7%). In contrast, primary ovarian insufficiency was more prevalent among women aged 36–45 years (9.8%).

3.5. Diagnostic quality and clinico-biological concordance

After standardised review, 63.1% of work-ups were fully conforming, 22.8% partially conforming, and 14.1% non-conforming, with lower conformity among adolescents (54.8%). Unfavourable DLQC was observed in 36.9%, highest among adolescents (45.2%). The mean DSS was 4.1 ± 1.3/6, with high scores in 50.3% of participants (Table 8).

Table 8. Diagnostic quality indicators and DLQC

Indicator	Result
Conforming work-up	1,196/1,894 (63.1%)
Partially conforming	431/1,894 (22.8%)
Non-conforming	267/1,894 (14.1%)
Favourable DLQC	1,196/1,894 (63.1%)
Unfavourable DLQC	698/1,894 (36.9%)
Clinico-biological discordance	384/1,894 (20.3%)
Mean DSS	4.1 ± 1.3/6
DSS 0–2	238 (12.6%)
DSS 3–4	704 (37.2%)
DSS 5–6	952 (50.3%)

3.6. Diagnostic reclassification, overdiagnosis and diagnostic cascades

Potential overdiagnosis occurred in 6.3%, mainly involving PCOS, while underdiagnosis affected 4.2%. Diagnostic reclassification occurred in 10.5%, more frequently among adolescents (15.5%). Diagnostic cascades followed abnormal results in 30.7%, with 22.3% potentially unnecessary. False-positive biological signals occurred in 12.1%, and 30.1% experienced diagnostic delays >3 months. Overall, 2,874 tests were classified as low added value, increasing with the number of tests prescribed.

3.7. Factors associated with inappropriate work-up and unfavourable DLQC

Multivariable analysis identified adolescence and ≥6 prescribed tests as the strongest predictors of poor diagnostic quality. Recent hormonal contraception was also associated with poorer quality, while hirsutism was protective. Multilevel analysis showed modest facility-level variability (ICC 5.2%), with an age × facility interaction (p=0.018).

Table 9. Multivariable analyses of factors associated with inappropriate work-up and unfavourable DLQC

Factor	Inappropriate work-up aOR (95% CI)	p	Unfavourable DLQC aOR (95% CI)	p
Age 10–19 vs 20–35 years	1.71 (1.28–2.28)	<0.001	1.78 (1.35–2.35)	<0.001
Age 36–45 vs 20–35 years	1.18 (0.91–1.54)	0.214	—	—
≥6 tests prescribed	2.86 (2.17–3.77)	<0.001	2.91 (2.24–3.78)	<0.001
Recent hormonal contraception	1.48 (1.08–2.03)	0.015	1.52 (1.12–2.07)	0.008
Hirsutism	0.74 (0.55–0.99)	0.043	0.72 (0.53–0.97)	0.031
Oligo/amenorrhoea	0.81 (0.62–1.07)	0.135	0.84 (0.64–1.10)	0.208

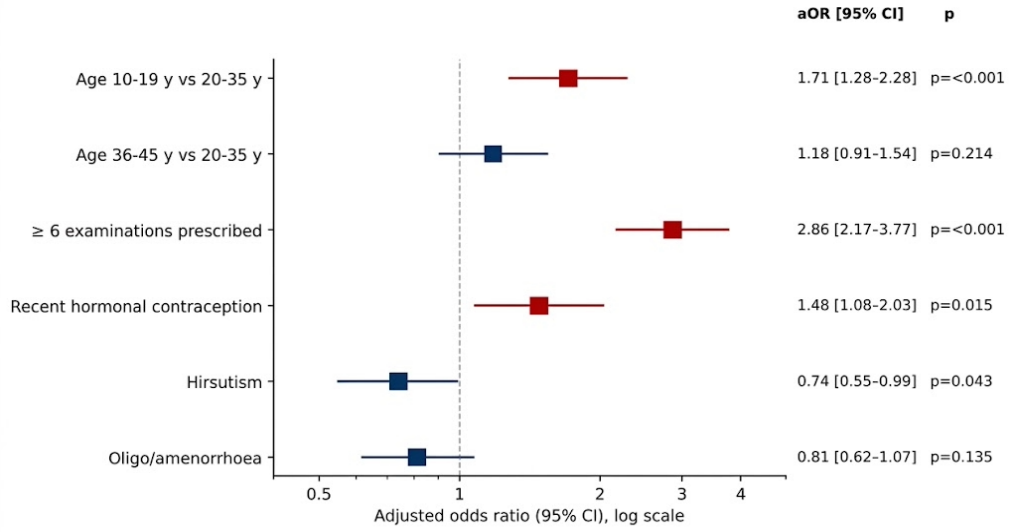

Figure 4. Factors independently associated with an inappropriate biological prescription/interpretation (multivariable model).

3.8. Diagnostic and clinical management impact

Laboratory testing changed the diagnosis in 29.7% of participants and identified clinically contributive abnormalities in 37.4%, while prompting justified additional investigations in 8.7%. Management changed in 26.1%. Among those with biological abnormalities, 76.8% received appropriate management, whereas 15.0% had suboptimal management. Overall, adolescents showed poorer diagnostic quality, and extensive testing was associated with lower diagnostic value (Table 10).

Table 10. Main diagnostic and clinical management outcomes

Outcome	n (%)
Clinically contributive biological abnormality	748/2,000 (37.4)
Diagnostic change	593/2,000 (29.7)
Justified additional investigation	174/2,000 (8.7)
Change in clinical management	521/2,000 (26.1)
Treatment initiated	283
Treatment modified	146
Inappropriate treatment discontinued	42
Specialist referral	50
Appropriate management of biological abnormality	865/1,126 (76.8%)
Excessive action	91/1,126 (8.1%)
Insufficient action	103/1,126 (9.1%)
No action documented	67/1,126 (5.9%)
Diagnostic reclassification	210/2,000 (10.5%)
Potential overdiagnosis	126/2,000 (6.3%)
Potential underdiagnosis	84/2,000 (4.2%)
Diagnostic delay >3 months	602/2,000 (30.1%)

4. Discussion

4.1. Main finding: distinguishing biological abnormality from clinical usefulness

The main finding was the dissociation between biological abnormalities and clinical usefulness: 56.3% of patients had an abnormal result, but only 37.4% had a clinically contributive abnormality. This highlights that laboratory quality should be assessed by diagnostic yield rather than the number of abnormal results. Hormonal results require interpretation according to age, cycle timing, physiological context and analytical method. This approach is consistent with current PCOS and menstrual-disorder guidelines, which recommend targeted, clinically integrated testing rather than indiscriminate hormonal panels [1–7].

4.2. Biological burden and multiplication of tests

The median of five tests per patient and the higher risk of unfavourable DLQC with ≥6 tests (aOR 2.91; 95% CI 2.24–3.78) indicate a substantial laboratory burden. However, this may reflect clinical complexity rather than overtesting alone. The rise in low-value testing from 8.7% to 34.6% with increasing test numbers supports a targeted, sequential and clinically driven diagnostic strategy rather than indiscriminate testing.

4.3. TSH, prolactin, FSH and LH: frequent tests requiring contextualisation

TSH and prolactin were the most frequently prescribed tests (85.7% and 81.5%), reflecting their relevance in menstrual disorders [5,10]. FSH and LH were also frequently requested (70.8% and 67.9%), although their value depends on the clinical phenotype and hormonal context [5,6]. Thus, the main issue was not the tests themselves, but their contextual appropriateness and interpretation.

4.4. Timing of sampling: a cross-cutting determinant of biological quality

Hormonal timing was a major quality gap: only 58.4% of cycle-dependent assays had adequately documented timing, while 41.6% were inadequate or undocumented. This is particularly relevant for FSH, LH, estradiol and progesterone, whose interpretation depends on cycle context [5,6]. Systematic documentation of the last menstrual period, cycle length and presumed cycle day could substantially improve interpretability, especially in adolescents.

4.5. Adolescence: a population particularly vulnerable to overdiagnosis

Adolescents showed the poorest diagnostic quality, with 29.5% inappropriate testing and 45.2% unfavourable DLQC. This reflects the difficulty of distinguishing physiological pubertal maturation from endocrine pathology. Current recommendations emphasise time since menarche and require persistent ovulatory dysfunction with hyperandrogenism for PCOS diagnosis, supporting cautious follow-up of at-risk adolescents rather than premature labelling [1–3,7].

4.6. Prolactin: distinguishing a biological abnormality from pathological hyperprolactinaemia

Prolactin was prescribed in 81.5% of patients. Although clinically relevant, moderate isolated elevations should be interpreted cautiously because physiological and pre-analytical factors may affect levels. Confirmation and, when appropriate, assessment for macroprolactin can prevent unnecessary investigations [10,19]. Thus, prolactin yield should be judged by its clinical contribution, not simply by the frequency of abnormal results.

4.7. Androgens: the importance of clinical phenotype and analytical quality

Testosterone was prescribed in 53.7% of patients, with excessive androgen panels in 4.9%. Assessment should be guided by the clinical phenotype and interpreted cautiously when elevations are borderline. Given the low testosterone concentrations in women, LC-MS/MS-based assays offer greater analytical reliability than direct immunoassays [11–16]. A targeted, sequential approach is therefore preferable.

4.8. AMH and ovarian ultrasound: avoiding overdiagnosis of PCOS

PCOS was the most frequent final diagnosis (27.6%). In adults, AMH may support PCOS assessment but should not be used in isolation. In adolescents, current recommendations are stricter: AMH and ovarian morphology should not be used to diagnose PCOS [1–4]. The 63 inappropriate AMH prescriptions in adolescents therefore highlight an important gap between practice and recommendations.

4.9. Fasting insulinaemia: an example of a test that is biologically available but of low decisional value

Fasting insulinaemia was prescribed in 30.7% of patients, including 176 without a specific indication. Although insulin resistance is relevant to PCOS pathophysiology, routine insulin assays have limited clinical utility and are not diagnostic criteria for PCOS [1,4]. This finding illustrates the gap between test availability and clinical usefulness.

4.10. PCOS, isolated biomarker and diagnostic reclassification

PCOS was the most frequent diagnosis and a major potential source of overdiagnosis. 13.9% of diagnoses were based mainly on an isolated biomarker, with 33.5% subsequently reclassified; the overall DRR was 10.5%. These findings reinforce the need for an integrated clinical and biological assessment, particularly in adolescents, in whom PCOS diagnosis requires persistent ovulatory dysfunction and hyperandrogenism after exclusion of other causes [1–3].

4.11. Biological abnormalities, clinico-biological discordance and overdiagnosis

The gap between 56.3% biological abnormalities and 37.4% contributive abnormalities highlights the limited clinical yield of some abnormal results. Clinico-biological discordance occurred in 20.3% of patients and was associated with extensive testing and isolated biomarker interpretation. Potential overdiagnosis (6.3%) and underdiagnosis (4.2%) indicate that the challenge lies not only in test volume, but in the quality of diagnostic reasoning and contextual interpretation.

4.12. Low-value tests and diagnostic cascades

Overall, 2,874 low-value tests (19.6%) were identified. Their proportion increased from 8.7% with 1–3 tests to 34.6% with ≥9 tests. Diagnostic cascades occurred in 30.7% of patients, with 22.3% considered potentially unnecessary. These findings support diagnostic stewardship through targeted testing focused on results likely to change management.

4.13. Repeated work-ups and diagnostic delay

Laboratory work-up was repeated in 25.7% of patients, although only 8.7% were considered unjustified. The median diagnostic delay was 2.4 months, exceeding three months in 30.1% overall and 38.5% of adolescents. Although delays are multifactorial, their coexistence with repeated testing, diagnostic cascades and a 10.5% DRR supports a more structured, sequential diagnostic strategy.

4.14. Conformity with recommendations and the gap between recommendations and practice

Overall, 63.1% of work-ups conformed to recommendations, while 22.8% were partially conforming and 14.1% non-conforming. Conformity was lower among adolescents (54.8%). These findings highlight the persistent gap between updated international recommendations and clinical practice, supporting continuing education, decision-support tools and standardised local protocols [1–3].

4.15. Differences between care facilities

Unfavourable DLQC varied from 31.8% to 42.1% across facilities. The ICC of 5.2% indicated modest but meaningful organisational variability, while the significant age × facility interaction ($p=0.018$) suggested that care context may affect age groups differently. These findings support the use of multilevel analysis rather than direct ranking of facilities.

4.16. Therapeutic consequences: demonstrating clinical causality

Therapeutic impact should be defined by clinical contribution rather than treatment change alone. A contributive result should demonstrate a clear sequence: biological result → diagnostic re-evaluation → justified therapeutic decision. This distinction strengthens the validity of the "contributive abnormality" concept and avoids attributing clinical impact to a test based solely on temporal association.

5. Impact and clinical implications

This study highlights a substantial gap between laboratory testing and clinical value in the evaluation of menstrual disorders. The findings suggest that diagnostic quality should not be judged by the number of tests performed or abnormal results detected, but by the extent to which investigations provide reliable, clinically meaningful information that changes diagnostic reasoning or management.

From a clinical perspective, the results support a shift from an exhaustive, test-intensive approach toward a targeted, sequential and value-based diagnostic strategy. Laboratory requests should be linked to a clearly defined clinical question, with appropriate documentation of menstrual and hormonal context. Particular attention should be paid to cycle timing, time since menarche in adolescents, and the interpretation of borderline or isolated hormonal abnormalities.

This approach may reduce false-positive biological signals, inappropriate PCOS labelling, unnecessary repeat testing, diagnostic cascades and diagnostic delays, while preserving investigations with genuine clinical yield. In adolescents, an age- and menarche-specific approach is particularly important to avoid interpreting physiological pubertal variations as pathological findings.

At the healthcare-system level, the implementation of standardised diagnostic pathways, local laboratory protocols, clinician education and clinical decision-support tools could improve adherence to recommendations and reduce variability between levels of care. Overall, diagnostic stewardship should aim not to minimise testing, but to ensure that each investigation has a clear indication, appropriate timing, reliable interpretation and demonstrable clinical value.

6. Conclusion

This multicentre study shows that the quality of the laboratory work-up for menstrual disorders cannot be assessed by the number of tests prescribed, nor by the frequency of biologically abnormal results alone. It relies above all on their clinical relevance, their adequacy to the physiological context, their timing, their analytical quality, and their capacity to justifiably change diagnostic reasoning or management.

Our results suggest that a non-negligible proportion of prescriptions or interpretations could be potentially inappropriate, with particular attention warranted for adolescents, in whom interpretation of hormonal parameters must take into account time since menarche, as well as for the different contexts and levels of care, which are likely to influence prescribing practices.

Optimisation of the laboratory work-up should therefore favour a targeted, hierarchical approach, based on the clinical question, phenotype, age, cycle phase, ongoing treatments and pre-test probability. Limiting low-value tests, notably when they are prescribed systematically without a clear clinical indication, could help reduce overdiagnosis, unnecessary secondary investigations and biological costs, while improving diagnostic efficiency.

Thus, the guiding principle should be: the right test, in the right patient, at the right time, with the right interpretation, for the right clinical decision.

In practice, the relevance of the work-up must take precedence over its complexity: fewer systematic tests, but more tests that are justified, correctly performed, and interpreted within their clinical context.

Funding: This research received no external funding

Conflicts of Interest: The authors declare no conflict of interest

ORCID iD: 0009-0001-1320-7874

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers

Artificial Intelligence (AI) Use Disclosure: The authors declare that no artificial intelligence tools were used in the preparation of this manuscript

References

[1] Teede HJ, Tay CT, Laven JJE, Dokras A, Moran LJ, Piltonen TT, et al. Recommendations from the 2023 International Evidence-based Guideline for the Assessment and Management of Polycystic Ovary Syndrome. *Hum Reprod.* 2023;38(9):1655-1679. doi:10.1093/humrep/dead156.

[2] Peña AS, Witchel SF, Boivin J, Burgert TS, Ee C, Hoeger KM, et al. International evidence-based recommendations for polycystic ovary syndrome in adolescents. *BMC Med.* 2025;23:151. doi:10.1186/s12916-025-03901-w.

[3] Peña AS, Witchel SF, Hoeger KM, Oberfield SE, Vogiatzi MG, Misso M, et al. Adolescent polycystic ovary syndrome according to the international evidence-based guideline. *BMC Med.* 2020;18:72.

[4] Christ JP, Cedars MI. Current Guidelines for Diagnosing PCOS. *Diagnostics (Basel).* 2023;13(6):1113.

[5] Practice Committee of the American Society for Reproductive Medicine. Current evaluation of amenorrhea: a committee opinion. *Fertil Steril.* 2024;122(1):52-61. doi:10.1016/j.fertnstert.2024.02.001.

[6] Gordon CM, Ackerman KE, Berga SL, Kaplan JR, Mastorakos G, Misra M, et al. Functional hypothalamic amenorrhea: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab.* 2017;102(5):1413-1439.

[7] American College of Obstetricians and Gynecologists. Menstruation in girls and adolescents: using the menstrual cycle as a vital sign. Committee Opinion No. 651. *Obstet Gynecol.* 2015;126(6).

[8] Munro MG, Critchley HOD, Fraser IS; FIGO Menstrual Disorders Committee. The two FIGO systems for normal and abnormal uterine bleeding symptoms and classification of causes of abnormal uterine bleeding in the reproductive years: 2018 revisions. *Int J Gynaecol Obstet.* 2018;143(3):393-408.

[9] Deligeoroglou E, Karountzos V. Abnormal uterine bleeding including coagulopathies and other menstrual disorders. *Best Pract Res Clin Obstet Gynaecol.* 2018;48:51-61.

[10] Melmed S, Casanueva FF, Hoffman AR, Kleinberg DL, Montori VM, Schlechte JA, et al. Diagnosis and treatment of hyperprolactinemia: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab.* 2011;96(2):273-288.

[11] Handelsman DJ, Wartofsky L. Requirement for mass spectrometry sex steroid assays in the Journal of Clinical Endocrinology and Metabolism. *J Clin Endocrinol Metab.* 2013;98(10):3971-3973.

[12] Handelsman DJ, Ly LP. An accurate mass spectrometry assay for testosterone and dihydrotestosterone in human serum. *J Clin Endocrinol Metab.* 2019;104(6):2281-2289.

[13] Barth JH, Field HP, Yasmin E, Balen AH. Defining hyperandrogenism in polycystic ovary syndrome: measurement of testosterone and androstenedione by liquid chromatography-tandem mass spectrometry and analysis by receiver operator characteristic plots. *Eur J Endocrinol.* 2010;162(3):611-615.

[14] Legro RS, Schlaff WD, Diamond MP, Coutifaris C, Casson P, Brzyski RG, et al. Total testosterone assays in women with polycystic ovary syndrome: precision and correlation with hirsutism. *J Clin Endocrinol Metab.* 2010;95(12):5305-5313.

[15] Morley LC, Tang T, Yasmin E, Norman RJ. Reproductive hormones for diagnosis of polycystic ovary syndrome in adolescents and adults: a systematic review and meta-analysis. *Hum Reprod Update.* 2017;23(3).

[16] Keevil BG. How do we measure androgens in women? *J Clin Pathol.* 2016;69(7):559-563.

[17] Forslund M, Melin J, Stener-Victorin E, Hirschberg AL, Teede H, Vanky E, et al. International evidence-based guideline on assessment and management of PCOS—A Nordic perspective. *Acta Obstet Gynecol Scand.* 2024;103(1):7-12.

[18] Turcu AF, Rege J, Auchus RJ, Rainey WE. 11-Oxygenated androgens in health and disease. *Nat Rev Endocrinol.* 2020;16(5):284-296.

[19] Elenkova A, Genov N, Abadzhieva Z, Vasilev V, Kirilov G, Zacharieva S. Macroprolactinemia in patients with prolactinomas: prevalence and clinical significance. *Exp Clin Endocrinol Diabetes.* 2013;121(4):201-205.

[20] Nanba AT, Rege J, Ren J, Auchus RJ, Rainey WE, Turcu AF. 11-Oxygenated C19 steroids do not decline with age in women. *J Clin Endocrinol Metab.* 2019;104(7):2615-2622.