

Diagnostic Significance of the FIGO PALM–COEIN Classification in the Evaluation of Abnormal Uterine Bleeding: An Integrative Critical Review

Xolboyeva Mashhura Yoqutjon qizi

Email – yoqutjonovnam@gmail.com

Termez University of Economics and Service

Faculty of Medicine

Abstract

Background: Abnormal uterine bleeding (AUB) is a symptom rather than a single disease, and diagnostic failure may follow when the bleeding phenotype is mistaken for an etiological diagnosis. The International Federation of Gynecology and Obstetrics (FIGO) PALM–COEIN system provides a common vocabulary for structural and nonstructural causes, but its practical diagnostic value depends on examination strategy, phenotype documentation, access to investigations, and competing risks. **Objective:** To critically evaluate the role of PALM–COEIN in triage, etiological work-up, longitudinal care, and real-world evidence generation, with particular attention to digital decision support and feasibility in Uzbekistan. **Methods:** An integrative narrative synthesis was organized around FIGO revisions, NICE and ACOG guidance, comparative diagnostic principles, and authoritative methodological and regulatory resources. No primary patient data were collected, no original meta-analysis was performed, and advanced causal or artificial-intelligence methods are described as potential study designs rather than completed analyses. **Results:** Classification supplies an operational differential diagnosis but does not independently establish etiology or quantify the accuracy of ultrasound, hysteroscopy, or histopathology. Pregnancy exclusion and assessment of hemodynamic stability precede routine classification. Clinical features help choose investigations: hysteroscopy is particularly relevant to suspected cavity pathology, whereas ultrasonography assesses myometrial disease. Ovulatory, coagulation, endometrial and iatrogenic processes frequently require history and laboratory interpretation rather than imaging alone. Multiple PALM–COEIN categories may coexist, and apparently incidental lesions should not automatically be considered causative. Real-world implementation requires structured bleeding data, standardized coding, quality assurance and careful control of confounding. **Conclusion:** FIGO PALM–COEIN is most useful as a diagnostic reasoning and communication framework embedded within risk-sensitive assessment; it is not a stand-alone diagnostic test. Implementation should be audited against clinically

meaningful outcomes such as diagnostic delay, anemia, unnecessary procedures, patient-reported symptoms and safe referral.

Keywords: abnormal uterine bleeding; FIGO; PALM–COEIN; heavy menstrual bleeding; endometrial polyps; adenomyosis; leiomyoma; endometrial hyperplasia; ovulatory dysfunction; hysteroscopy; transvaginal ultrasonography; real-world evidence; clinical decision support; diagnostic safety.

1. Introduction

Clinical framing

AUB comprises disturbances of bleeding frequency, regularity, duration or volume and requires a structured symptom description before causal attribution. FIGO System 1 formalizes the phenotype; System 2, PALM–COEIN, organizes possible etiologies. The distinction matters because a patient may have both a sonographically visible lesion and an independent endocrine driver of irregular bleeding. Calling every visible fibroid the cause of bleeding creates attribution bias and risks inappropriate intervention [1,2].

The dual FIGO architecture

The structural group contains Polyp (P), Adenomyosis (A), Leiomyoma (L), and Malignancy/hyperplasia (M). Nonstructural categories are Coagulopathy (C), Ovulatory dysfunction (O), Endometrial disorders (E), Iatrogenic causes (I) and Not otherwise classified (N). AUB-L is further characterized by anatomical relationship to the cavity. The classification supports coexistence rather than insisting on a single diagnosis, reflecting the heterogeneity of reproductive-age bleeding [1].

Why the framework matters

The central diagnostic challenge is not only distinguishing a structural lesion from a functional process; it is deciding which findings are actionable given the person's age, pregnancy intentions, stability, anemia burden, medication history and malignancy risk. Standardization can improve referral documentation and longitudinal audit, but cannot repair poor examination technique or absent histological confirmation [2,3].

Real-world data and evidence

Real-world data (RWD) include routine electronic medical records, prescriptions, laboratory records, registries and patient-generated reports; real-world evidence (RWE) describes clinical inferences obtained through appropriate analysis of such data. In AUB, RWD may estimate diagnostic delay and treatment persistence but often incompletely records actual bleeding volume, adverse experiences and the reasons for treatment choices. Administrative diagnostic codes are not equivalent to validated FIGO phenotypes [7,8].

Efficacy, effectiveness and trials

Randomized controlled trials (RCTs) estimate treatment efficacy under defined eligibility, protocol and follow-up conditions. Effectiveness in routine care may differ when patients have multiple coexisting etiologies, irregular adherence, unavailable investigations or incomplete follow-up. RCTs reduce confounding in specified comparisons; nonrandomized RWE contributes transportability and longer-term trajectories but remains vulnerable to confounding by indication, immortal-time bias and measurement error. Neither source should be treated as universally superior [7,9].

International governance

WHO emphasizes human oversight and fair, explainable deployment of AI; FDA's RWE framework and evolving AI-enabled device guidance address evidence fitness, verification and lifecycle risk; EMA uses routinely collected healthcare data through mechanisms including DARWIN EU. These initiatives offer transferable principles for research quality, not an endorsement of any AUB-specific prediction model [7,8,10].

Clinical AI opportunities

Image analysis could assist characterization of uterine lesions, while natural-language processing may identify missed AUB descriptions in unstructured gynecology notes. Clinical decision support should first prioritize high-harm omissions such as hemodynamic compromise, possible pregnancy and endometrial pathology. AI predictions cannot replace a complete menstrual and medication history or shared decisions about examination and treatment [10].

2. Materials and Methods

Study design: critical integrative narrative review with a diagnostic-pathway focus. Sources were located through targeted searches of PubMed-indexed FIGO material and clinical guidance, with reference to international methodological documentation. PubMed, Scopus, Web of Science, Embase and Cochrane Library are recommended search repositories for a subsequent reproducible systematic review; no claim is made that each was independently searched or exhaustively screened for this manuscript. Journal sites and institutional regulatory pages were consulted as applicable. Searches prioritize 2020–2026 materials while retaining foundational FIGO terminology and relevant earlier clinical guidance. No protocol was registered and no PRISMA flow diagram, pooled effect estimate or invented screening count is presented.

Inclusion logic prioritized consensus classification documents, professional guidelines, diagnostic comparative evidence, eligible systematic reviews,

methodology papers and regulatory material whose population, indication and version were identifiable. Exclusions concerned pregnancy-related bleeding misclassified as nongravid AUB, studies without relevant etiological resolution, unclear reference standards and claims of algorithmic accuracy unsupported by external validation. The hierarchy of certainty distinguishes confirmed published evidence, reasoned clinical interpretation and prospectively proposed methods.

Extraction framework: for each clinically applicable source we recorded patient context, type of bleeding, reference standard, index test, diagnostic threshold, verification approach, blinding where reported, comorbid causes, and clinical consequences. Comparative observations were synthesized narratively because heterogeneity in case selection, reference standards and endpoints makes naïve pooled accuracy estimates misleading. Risk of bias is discussed using QUADAS-2 concepts for diagnostic work and ROBINS-I concepts for nonrandomized interventions; these tools are not presented as formally scored within this narrative manuscript.

[Propensity-score matching \(PSM\)](#)

For a future comparative AUB cohort, estimate treatment or investigation propensity from pre-index bleeding severity, age, parity, comorbidity, medication exposure and imaging availability; match patients in the area of overlap and assess standardized mean differences. Matching alone does not resolve unmeasured confounding.

[Inverse probability weighting \(IPW\)](#)

Construct stabilized weights from the probability of observed exposure and possibly censoring, inspect positivity violations and extreme weights, and present balance diagnostics. A strategy comparison such as immediate hysteroscopy versus initial sonography requires a clear time zero and a prespecified eligibility rule.

[Cox regression and survival analysis](#)

Time-to-event analysis can model time to definitive diagnosis, recurrence or surgery when censoring is properly represented. Proportional hazards should be assessed; competing events (such as hysterectomy versus medical control) may require cause-specific or subdistribution approaches.

[Kaplan–Meier estimation](#)

Kaplan–Meier curves can estimate time to recurrence or further intervention, but they require unambiguous event definitions, aligned starting dates and defensible noninformative censoring assumptions. They are not appropriate substitutes for cross-sectional diagnostic accuracy studies.

Target trial emulation

Specify eligibility, strategies, assignment time, follow-up, endpoints, causal contrast and analysis plan before examining routine records. Avoid immortal-time bias by aligning treatment assignment with follow-up start; use sensitivity analyses for treatment crossover.

Causal inference

Draw a directed acyclic graph distinguishing confounders (age, pregnancy intentions, anemia severity), mediators (treatment response), colliders (specialist referral) and outcomes. Selection into tertiary services can create collider bias when comparing diagnostic modalities.

Machine learning

Train interpretable baselines (regularized regression) before random forests, gradient boosting or nonlinear classifiers. Evaluate calibration, discrimination, decision-curve net benefit and generalizability rather than AUROC alone; keep patient-level temporal test splits.

Deep learning

Deep neural networks can investigate sonographic images or pathology whole-slide scans, but performance is vulnerable to equipment distribution shift, annotation noise, referral enrichment and label leakage. No diagnostic deep-learning model was developed here.

Natural-language processing (NLP)

Extract menstrual duration, flooding, intermenstrual bleeding, medication exposure and treatment preference from clinical notes with audited dictionaries or language models; adjudicate negation, temporal context and duplicate episodes.

Reinforcement learning and LLMs

Reinforcement learning remains investigational for sequential care strategies because counterfactual validation is difficult; large language models can assist record abstraction or patient education only with provenance, hallucination safeguards and clinician oversight.

3. Results: Evidence Synthesis

The findings below are literature-derived clinical and methodological findings, not outputs of a new patient cohort. This section is organized as a diagnostic decision architecture rather than as a series of unverified quantitative comparisons.

Category	Clinical target	Primary evidence route
AUB-P	Cavity polyp	Hysteroscopy / appropriately

		selected intracavitary imaging
AUB-A	Adenomyosis	Transvaginal ultrasound, MRI if unresolved
AUB-L	Fibroids	Pelvic ultrasound and localization relative to cavity
AUB-M	Hyperplasia/malignancy	Risk-sensitive histological investigation
AUB-C	Bleeding disorders	Personal/family history and targeted hemostasis work-up
AUB-O	Ovulatory dysfunction	Pattern, endocrine assessment when appropriate
AUB-E	Primary endometrial disorder	Diagnosis after competing structural/systemic causes considered
AUB-I	Medication/device-associated	Chronology of treatment and exposure
AUB-N	Other poorly classified processes	Individualized specialist assessment

3.1 Diagnostic safety before etiological labeling

In acute bleeding, classification must not delay vital-sign assessment or stabilization. A urine or serum pregnancy test is essential when pregnancy is possible, even when the eventual focus is nonpregnant AUB. Symptoms suggesting hypovolemia, syncopal events, severe anemia or active hemorrhage require urgent management. FIGO terminology becomes most useful after immediately dangerous causes have been addressed [4].

3.2 System 1: standardized description

Frequency, regularity, duration, perceived blood loss and intermenstrual timing should be recorded explicitly. Clinicians should distinguish regular heavy bleeding from unpredictable anovulatory episodes and document how symptoms affect functioning. Routine datasets tend to capture a generic “menorrhagia” code without

the nuance needed to determine etiology. A symptom dictionary can be embedded in clinical forms without forcing patients to memorize medical terminology [1,2].

3.3 Polyp and cavity pathology

Endometrial polyps may present with intermenstrual spotting or heavy bleeding, but sonographic detection does not prove causality. Hysteroscopy provides direct visualization and potentially targeted therapy when cavity pathology is suspected. NICE recommends outpatient hysteroscopy under specified clinical patterns, highlighting that procedure choice should depend on prior probability, patient preference and access rather than a uniform test applied to all [3].

3.4 Adenomyosis

Dysmenorrhea, a tender enlarged uterus and characteristic myometrial findings may support adenomyosis. Transvaginal ultrasonography is usually practical for first-line assessment; MRI can be a problem-solving modality when the ultrasound window is limited or concurrent complex uterine disease obscures interpretation. Overlapping endometriosis, fibroids and age-related features complicate causal attribution, and a negative scan does not eliminate all clinically important disease [3].

3.5 Leiomyoma

The probability that a leiomyoma explains bleeding depends on lesion size, number, and especially cavity distortion; classification should therefore include more than a simple present/absent flag. Submucosal lesions are often more clinically relevant to bleeding than incidental serosal lesions, although exceptions occur. AUB-L should be cross-coded with other causes when ovulatory dysfunction or coagulation disorders coexist [1].

3.6 Hyperplasia and malignancy

AUB-M is a diagnostic risk category with disproportionately important consequences. Persistent intermenstrual bleeding, obesity-associated infrequent heavy bleeding, prolonged unopposed estrogen exposure or treatment failure may prompt further evaluation. Selection of biopsy strategy depends on the clinical situation and jurisdictional guidelines; the NICE HMB pathway favors biopsy at hysteroscopy for specifically high-risk women and does not recommend blind biopsy as a routine HMB test [3].

3.7 Coagulopathy

Heavy bleeding starting near menarche, recurrent epistaxis, bruising and a family history of abnormal bleeding may suggest a hemostatic disorder. Screening instruments and specialist investigations must account for prior hormone use, acute illness and sample timing. A normal ultrasound cannot exclude an inherited bleeding

disorder, so diagnostic work must remain broader than structural imaging alone [1,4].

3.8 Ovulatory dysfunction

Irregular cycles with unpredictable heavy bleeding commonly arise from ovulatory dysfunction, but the label should not substitute for determining likely causes such as PCOS, thyroid abnormalities, hypothalamic disorders or perimenopausal transition. FIGO's separate HyPO-P framework adds etiological granularity after ovulatory dysfunction has been identified and helps prevent an overly ovarian-centered interpretation [5,6].

3.9 Primary endometrial processes

AUB-E denotes local endometrial disturbance in the absence of another sufficiently explanatory process. It is not a default label for every patient with normal ultrasound findings. Endometrial hemostasis involves prostaglandin pathways, vasoconstriction, fibrinolysis and tissue repair; interpretation depends on compatible phenotype and exclusion of competing categories rather than on one routine molecular assay [1].

3.10 Iatrogenic and otherwise classified causes

Anticoagulants, hormonal contraception, intrauterine devices, endocrine medicines and treatment interruptions can influence bleeding patterns. A time-stamped medication history often has greater diagnostic value than indiscriminate hormonal testing. Surgical scars and unusual anatomical processes may enter the N category when they do not fit existing definitions; coding should describe uncertainty instead of falsely creating a definitive diagnosis [1].

3.11 Coexisting etiologies

AUB is seldom an exercise in choosing a mutually exclusive checkbox. A patient with both endometrial polyp and PCOS can have structural and ovulatory contributions, and a person with fibroids and anticoagulation may experience compounded bleeding. A diagnostic record should specify which findings are confirmed, suspected or possibly incidental, and revisit causal attribution after treatment response or additional pathology results.

3.12 Ultrasound selection

Transvaginal ultrasound offers accessible evaluation of uterine anatomy but is reader-dependent and less definitive for certain intracavitary lesions. Transabdominal scanning may help when transvaginal scanning is declined or anatomically unsuitable. Diagnostic sensitivity should never be reported without the patient spectrum, comparator and reference standard; enriched surgical cohorts exaggerate apparent certainty when generalized to primary care [3].

3.13 Hysteroscopy and histology

Hysteroscopy offers direct cavity examination and can facilitate targeted histological sampling. It does not evaluate the external myometrium as well as imaging, and normal appearances do not remove every reason to pursue tissue diagnosis. Biopsy limitations differ between diffuse pathology and focal lesions; a coherent investigation strategy should specify how discordant findings are reconciled [3].

3.14 Hematologic and endocrine tests

A complete blood count helps quantify anemia, with iron status assessed where appropriate; targeted thyroid or endocrine testing is justified when symptoms suggest an underlying dysfunction. Broad hormone panels without a clinical question may create false positives, inconsistent interpretation and avoidable costs. Test ordering should be matched to phenotype and differential diagnosis, not driven by the presence of an administrative AUB code.

3.15 Diagnostic hierarchy and red flags

A risk-stratified pathway separates immediate safety assessment, phenotyping, examination, targeted imaging, histological evaluation when indicated, and later reassessment. It includes explicit review of pregnancy status, bleeding severity, anemia burden, malignancy risk and reproductive intentions. The framework is intended to reduce unsafe delay, not compel invasive investigations in patients with uncomplicated low-risk symptoms [3,4].

3.16 Patient-reported outcomes

Bleeding volume alone incompletely represents disease burden. Fatigue, missed work, anxiety, sexual and reproductive concerns, and dissatisfaction with treatment are relevant outcomes that may alter management despite an unchanged imaging label. Patient-reported outcomes can distinguish symptom control from etiological resolution and should be collected using consistent questions at appropriate follow-up intervals.

3.17 Electronic medical records

Electronic health records are useful for tracing bleeding presentations, laboratory trends, completed investigations and referral intervals. Their main limitation is variable coding granularity: AUB-O may be inferred from a clinician note while the coded diagnosis remains “heavy menstrual bleeding.” A registry built for FIGO analysis must preserve narrative provenance, pregnancy exclusion and certainty of classification.

3.18 Electronic prescribing data

Prescription streams document tranexamic acid, hormonal preparations, anti-inflammatory medicines and other relevant exposures, but a prescription does not

prove medication use. The distinction between prescribed, dispensed and taken doses becomes essential in effectiveness analyses. Linking repeat prescription events to bleeding outcomes without controlling severity risks confusing treatment selection with treatment effects.

3.19 Laboratory and imaging data

Laboratory systems contribute hemoglobin, ferritin and selected endocrine or coagulation results; imaging archives provide structural findings and measurements. Device types, operator experience and report templates differ between institutions. Structured report elements describing lesion location and whether cavity distortion exists allow more clinically relevant classification than imaging “positive” or “negative” fields.

3.20 Registries, insurance and pharmacovigilance

National clinical registries can monitor regional inequities and management trajectories if definitions are consistent. Insurance claims broaden follow-up but encode reimbursement events, not complete symptom histories. Pharmacovigilance reports provide safety signals that require careful adjudication and lack reliable exposed denominators; they must not be interpreted as comparative incidence without additional data.

3.21 Wearables and mobile applications

Cycle-tracking applications could provide longitudinal phenotype information that episodic clinic visits miss, but app adoption is selective and menstrual logs are subject to recall, reporting and algorithmic bias. Wearable physiology cannot currently replace etiological examination. A clinically safe implementation requires informed consent, clear data retention rules, and meaningful access for patients without smartphones.

3.22 RWE evaluation of diagnostic pathways

Routine data can compare pathways such as immediate hysteroscopy versus an ultrasound-first strategy for time to diagnosis and number of unnecessary procedures. However, women referred directly for hysteroscopy may have more concerning symptoms or higher malignancy risk. Propensity adjustment requires recorded indication, severity, referral pathways, fertility intent and clinician availability; residual confounding must be highlighted rather than hidden.

3.23 Treatment efficacy versus effectiveness

Clinical trials of symptom treatments should not be translated automatically to everyday AUB care where adherence, concurrent causes and follow-up differ. An intervention can reduce bleeding without correcting the underlying cause, and an etiologically definitive procedure may be unacceptable to someone prioritizing

fertility. Patient-centered comparative effectiveness therefore requires both symptom and preference-sensitive outcomes.

3.24 Machine learning and image recognition

A supervised model could prioritize women for additional evaluation by combining symptom pattern, age, anemia, obesity, medication exposures and imaging findings. The correct benchmark is usual clinical care, not an artificially weak algorithm. External calibration, error analysis in younger versus older patients and prospective studies of clinician override are prerequisites to any clinical recommendation.

3.25 Deep learning and natural language processing

Deep-learning analysis of ultrasound may assist lesion segmentation, but a reliable ground truth is particularly challenging when reference diagnoses differ by pathology, imaging or treatment response. NLP applied to free-text notes may identify previously undocumented intermenstrual bleeding, although negation, menstrual history copied between visits and historical medication exposures can contaminate labels.

3.26 Reinforcement learning and large language models

Sequential treatment optimization via reinforcement learning is an appealing research idea but observational datasets rarely contain complete counterfactual outcomes. Large language models may support explanatory summaries or clinical documentation only if unsupported recommendations are actively blocked. Neither method should independently assign malignancy risk or bypass tissue-based evaluation.

3.27 Drug-effect prediction across conditions

Prediction models developed in diabetes, cardiology, oncology, infectious disease or neurology provide useful methodological lessons on treatment heterogeneity, but they cannot be imported as calibrated AUB prediction models. Shared exposures (obesity, anticoagulation, chronic disease medication) influence management and safety; disease-specific endpoints and validation remain mandatory.

3.28 Safety analytics and adverse drug reactions

Monitoring adverse reactions to antifibrinolytics, hormonal therapy or devices demands the distinction between suspected reports and confirmed causal events. Signal detection can flag unexpected reporting patterns; formal evaluation requires exposure denominator, indication and competing causes. Safety models should not reward undertreatment of severe bleeding by treating all medication exposure as an avoidable hazard.

3.29 Precision gynecology

A personalized approach integrates reproductive goals, anemia severity, lesion anatomy, endocrine profile and medication contraindications. Biomarker-driven risk prediction is less mature than imaging and histology for many AUB etiologies. Pharmacogenomic or pharmacogenetic testing should not be promoted as routine FIGO triage in the absence of validated clinical utility and guideline endorsement.

3.30 Cost, access and clinical implementation

Triage efficiency depends on local access to ultrasound, hysteroscopy, pathology and follow-up. A simple high-quality FIGO form may outperform a technically sophisticated system that lacks staffing and patient continuity. Economic assessment should consider missed malignancy, repeat visits, time away from work, fertility consequences and avoidable emergency treatment, not only unit procedure price.

4. Discussion

Interpretation of classification utility

PALM–COEIN is a semantic and diagnostic organization system rather than a laboratory assay with an intrinsic sensitivity and specificity. Claims that FIGO classification itself “detects” a percentage of diseases conflate diagnostic pathways with coding. Comparative effectiveness should instead ask whether embedding the classification changes correct referral, missed disease, time to appropriate investigation or unnecessary surgery. A classification is useful when it modifies behavior and measurement, not simply when every clinic form contains nine checkboxes.

Internal and external validity

Tertiary gynecology services are enriched for refractory and complicated presentations. A cohort in which all participants undergo hysteroscopy differs profoundly from unselected primary-care patients, many of whom may have low-risk bleeding requiring no invasive investigation. Spectrum effects influence positive and negative predictive values; verification bias occurs if histology is performed only in individuals with abnormal imaging. Both issues can make a diagnostic pathway appear more reliable than it will be in the intended population.

Comparative guideline considerations

FIGO provides an etiological language, NICE recommends differential selection of investigation based on clinical suspicion, and ACOG guidance emphasizes stabilization in acute AUB. These documents answer different questions and should not be treated as competing diagnostic accuracy trials. An appropriate pathway combines their complementary functions: emergency triage, symptom

characterization, diagnosis-specific investigation, informed choice and outcome reassessment. Differences across national biopsy policies also mean that a procedure recommended in one setting cannot simply be described as universal.

Algorithmic bias

Bias can enter at symptom documentation, specialist referral, imaging access, pathology verification and algorithm training. If only insured, urban or tertiary-care patients receive definitive diagnostic work-up, an AI system may learn to predict service access rather than biological pathology. Demographic group audits should report both predictive accuracy and rates of missed high-risk pathology across age, geography, financial access and disability groups, subject to lawful and ethical data collection.

Data quality and missingness

Missing laboratory values may reflect clinical judgment, patient reluctance, transport cost or unavailable tests; consequently, missingness is rarely completely at random. Complete-case analysis can bias estimated diagnostic patterns. Multiple imputation may be defensible for appropriate variables under explicit assumptions, but missing histology from women never offered a biopsy requires a different verification-bias strategy. Every model should publish completeness tables and clarify what each “unknown” field represents.

Explainable AI and calibration

Feature attribution tools can identify influential predictors but cannot establish biological causation. A model assigning high importance to referral source may expose site practice rather than disease mechanism. Predicted risk must be calibrated in settings with different baseline incidence; for safety-critical use, threshold selection should reflect the harm of missed malignancy and the burden of unnecessary procedures. Clinician override, audit logs and model monitoring should remain available.

Ethical AI

WHO principles emphasize autonomy, safety, transparency, accountability, inclusiveness and responsiveness to context. For reproductive health records, privacy risks are substantial because menstruation, fertility plans and sexual-health information are highly sensitive. Patients need understandable notice about data reuse, meaningful consent where applicable, and routes to correct inaccuracies. Human clinicians must retain responsibility for the interpretation of uncertain or conflicting information [10].

Regulatory science

An AI-enabled device should be assessed for its specified intended use, not a vague claim to enhance gynecology. Developers need a locked evaluation dataset,

analytic validation, clinical validation, postdeployment surveillance and change-control procedures. FDA's general RWE and device frameworks, and EMA's DARWIN EU infrastructure, provide examples of evidence-generation and governance directions. Their existence does not constitute regulatory approval of a specific AUB model.

Clinical decision support

The most immediately feasible intervention may be a rules-based reminder rather than a neural network. A tool can prompt exclusion of pregnancy, record severity, identify anemia assessment needs, document medications and ensure that postmenopausal bleeding is not incorrectly managed under reproductive-age FIGO pathways. Every alert should show its logic and support contextual dismissal; excessive false alerts can degrade clinical vigilance.

RWD and transportability

Healthcare pathways depend on insurance coverage, transport, cultural preferences and clinical capacity. A treatment effect estimated from a highly resourced service may not transport to rural settings where follow-up and diagnostic histology are delayed. Multisite analysis should explicitly evaluate effect modification by access and center rather than assuming one nationwide estimate applies everywhere. Reporting must distinguish evidence for classification harmonization from evidence for patient benefit.

Treatment safety across specialties

The discussion of drug-effect prediction in diabetes, cardiovascular disease, oncology, infectious disorders and neurology is methodologically relevant because comorbidity and co-prescription affect bleeding risks. Nevertheless, drug-response models validated for those conditions should not be repackaged as gynecological clinical tools without direct validation. Anticoagulant-associated bleeding and possible drug interactions are clinically germane to AUB; generalized AI enthusiasm is not a substitute for indication-specific pharmacovigilance.

Clinical heterogeneity and fertility

A definitive procedure is not equivalent to an optimal outcome for every patient. Fertility intention, age, anemia, pain and preference can change the relative value of medical and procedural approaches even when the FIGO category is the same. Decision analyses should therefore specify both clinical and patient-valued endpoints, avoiding a single success variable that masks unacceptable tradeoffs. The ability to code coexisting PALM–COEIN categories supports this patient-centered reasoning.

Temporal classification and reassessment

Bleeding etiology may change after pregnancy, initiation of contraception, menopause transition, development of a fibroid or treatment of an endocrine disorder. The classification should be time-stamped. A static chart label entered years earlier may lead clinicians to ignore new risk factors or a new pattern of intermenstrual bleeding. Structured reassessment is especially important when empirical treatment fails or iron deficiency recurs, and systems should preserve the evidence supporting each previous attribution.

Validation study design

A future prospective study could consecutively recruit patients presenting with AUB, record FIGO System 1 phenotype before imaging, and assign candidate etiologies using standardized references, blinded where possible. Prespecified primary outcomes should include diagnostic concordance, appropriate test utilization and clinically important delays. Secondary outcomes could track anemia, patient-reported bleeding and treatment burden. Such a design would test the effect of implementing classification, rather than falsely attributing performance metrics to terminology alone.

Causal estimands in pathway evaluation

Suppose an observational comparison evaluates ultrasound-first versus hysteroscopy-first care. The relevant causal question is the outcome that would occur if the same eligible patients could follow either strategy from a common start date. Confounders include clinical suspicion, biopsy indications, referral provider and waiting-list length. Post-baseline treatment decisions can act as mediators; conditioning on them without careful design changes the estimand. Target trial emulation and sensitivity analyses are therefore not optional decorations but guards against a misleading conclusion.

Limitations of the present synthesis

This manuscript is a critical narrative review, not a registered systematic review. Search coverage is not claimed as exhaustive; included sources vary from consensus classifications to clinical guidance and research-method references. No original patient dataset, pooled odds ratio, study-selection count or risk-of-bias score was generated. Suggestions for local implementation and AI validation are hypotheses and operational proposals, not observed findings. This distinction is necessary to preserve scientific integrity.

Research priorities

Priority questions include whether structured FIGO classification shortens time to appropriate testing, whether primary-care ultrasound-first strategies are safe in low-risk women, and whether electronic phenotyping can reliably distinguish

incidental lesions from probable causative pathology. Future studies should prioritize external validation, patient-reported outcomes, referral equity and pragmatic effectiveness. AUC, calibration and diagnostic sensitivity matter only alongside harms, benefits, operational feasibility and patient experience.

5. Uzbekistan: Digital Health and Clinical Research Priorities

Health-system context

Implementation proposals should align with the Uzbekistan Ministry of Health and existing national digitalization programs, with verification of currently operating modules before any claim of nationwide coverage. The proposed FIGO registry is not represented here as a deployed national platform. Regional hospitals differ in pathology access, gynecologic ultrasound capacity and informatics maturity; a phased implementation should begin with reliable clinical documentation.

Electronic health records and e-prescribing

An AUB registry should record menstrual symptom phenotype, pregnancy exclusion, blood count, selected imaging findings, histology when appropriate and all suspected etiological categories. Electronic prescription data could support evaluation of tranexamic acid and hormonal treatment persistence, with explicit distinction between prescribed and dispensed medicine. Secure identifiers and interoperability rules are essential before linking gynecology and pharmacology records.

National registries and clinical pharmacology

A minimal dataset could be adopted through participating clinics and evaluated for interrater reliability before national expansion. Clinical pharmacology teams could assist analysis of anticoagulant-associated bleeding, interactions and adverse reaction reports. Signal detection must not substitute for review of the suspected reaction, and patients should not be deprived of indicated therapy simply because a surveillance rule flagged the prescription.

Pragmatic training and ultrasound quality

Training can focus on FIGO System 1 bleeding nomenclature, PALM–COEIN coexistence coding, lesion localization, recognition of anemia and safe referral. Ultrasonography templates should describe submucosal relationships and features compatible with adenomyosis. A blinded sample of reports could be audited for agreement against specialist review, with remediation before introducing automated image classification.

AI readiness and equitable rollout

Pilot tools should be auditable and support clinicians rather than autonomously determine diagnosis. A rules-based checklist for urgent bleeding and indicated

follow-up may be more practical than deep learning. If machine learning is studied, validation must include rural and urban settings, age strata and varying access to hysteroscopy, because historical inequalities can otherwise be learned by the model.

Proposed multicenter protocol

A prospective observational program could compare completeness of FIGO documentation and diagnostic trajectory before and after implementation in multiple clinics. Collect baseline risk, ultrasound access, histology decisions, anemia outcomes, referral times and patient-reported burden. Use a predefined analysis plan, cluster-aware statistics and qualitative clinician feedback. Institutional ethics review, secure data storage and explicit reporting of missing data are required.

Recommendations for policy

A practical sequence is to define a minimum dataset, train health workers, introduce quality-controlled ultrasound reporting, standardize referral thresholds, evaluate audit metrics, and only then assess advanced AI tools. Indicators should measure missed serious disease, unnecessary procedures and patient experience rather than the percentage of electronic charts with a completed drop-down field.

6. Conclusion

FIGO PALM–COEIN creates a consistent etiological language for nonpregnant reproductive-age abnormal uterine bleeding, but its principal diagnostic contribution is structured reasoning and appropriate selection of further investigation. It cannot independently diagnose polyps, cancer, coagulation disorders or endocrine causes. A high-quality pathway combines symptom characterization, urgent safety assessment, individualized examination, relevant imaging and pathology, recognition of coexisting causes, and longitudinal reassessment. Routine-data analytics and AI could strengthen audit and decision support only after validation, bias assessment and governance. Uzbekistan-specific deployment should prioritize a minimal interoperable dataset, workforce training, appropriate referral and patient safety before algorithmic sophistication.

Funding: None declared for this manuscript. Conflicts of interest: To be declared by the actual authors. Ethics approval: Not applicable to a literature-only review. Data availability: No original individual-level patient dataset was generated.

References

1. 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. doi:10.1002/ijgo.12666.

2. Munro MG, Critchley HOD, Fraser IS. Research and clinical management for women with abnormal uterine bleeding in the reproductive years: more than PALM-COEIN. *BJOG.* 2017;124:185-189. doi:10.1111/1471-0528.14431.

3. National Institute for Health and Care Excellence. Heavy menstrual bleeding: assessment and management. NICE guideline NG88. Updated 2021. <https://www.nice.org.uk/guidance/ng88>.

4. American College of Obstetricians and Gynecologists. Management of acute abnormal uterine bleeding in nonpregnant reproductive-aged women. Committee Opinion No. 557. 2013. <https://www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2013/04/management-of-acute-abnormal-uterine-bleeding-in-nonpregnant-reproductive-aged-women>.

5. International Federation of Gynecology and Obstetrics. FIGO ovulatory disorders classification system. *Hum Reprod.* 2022;37(10):2446. <https://academic.oup.com/humrep/article/37/10/2446/6670602>.

6. The FIGO ovulatory disorders classification system. *Int J Gynaecol Obstet.* 2022. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10086853/>.

7. World Health Organization. Ethics and governance of artificial intelligence for health. Geneva: WHO; 2021. <https://www.who.int/publications/i/item/9789240029200>.

8. US Food and Drug Administration. Framework for FDA's Real-World Evidence Program. Silver Spring: FDA; 2018. <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>.

9. European Medicines Agency. Data Analysis and Real World Interrogation Network (DARWIN EU). <https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu>.

10. World Health Organization. WHO issues first global report on artificial intelligence in health and six guiding principles. 2021. <https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use>.

11. National Institute for Health and Care Excellence. Heavy menstrual bleeding: rationale and impact. NG88. <https://www.nice.org.uk/guidance/ng88/chapter/rationale-and-impact>.

12. National Institute for Health and Care Excellence. Heavy menstrual bleeding: recommendations for research. NG88. <https://www.nice.org.uk/guidance/ng88/chapter/Recommendations-for-research>.

13. National Institute for Health and Care Excellence. Heavy menstrual bleeding: recommendations. NG88.
<https://www.nice.org.uk/guidance/ng88/chapter/Recommendations>.
14. World Health Organization. Ethics and governance of artificial intelligence for health: executive summary. 2021.
<https://www.who.int/publications/i/item/9789240037403>.
15. European Medicines Agency. Real-world evidence.
<https://www.ema.europa.eu/en/human-regulatory-overview/research-development/real-world-evidence>.
16. US Food and Drug Administration. Real-world evidence.
<https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>.
17. US Food and Drug Administration. Artificial intelligence and machine learning in software as a medical device. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>.
18. World Health Organization. Global strategy on digital health 2020–2025.
<https://www.who.int/publications/i/item/9789240020924>.
19. International Council for Harmonisation. E6(R3) Good Clinical Practice.
<https://www.ich.org/page/efficacy-guidelines>.
20. Equator Network. STROBE Statement: strengthening reporting of observational studies in epidemiology. <https://www.strobe-statement.org/>.
21. Equator Network. PRISMA 2020 statement. <https://www.prisma-statement.org/>.
22. Equator Network. TRIPOD statement. <https://www.tripod-statement.org/>.
23. Equator Network. STARD 2015: reporting diagnostic accuracy studies.
<https://www.equator-network.org/reporting-guidelines/stard/>.
24. National Institutes of Health. PubMed biomedical literature database.
<https://pubmed.ncbi.nlm.nih.gov/>.
25. Cochrane. Cochrane Library. <https://www.cochranelibrary.com/>.
26. International Federation of Gynecology and Obstetrics. Official FIGO website and clinical resources. <https://www.figo.org/>.
27. World Health Organization. Digital health programme.
<https://www.who.int/health-topics/digital-health>.
28. European Medicines Agency. Big data.
<https://www.ema.europa.eu/en/about-us/how-we-work/big-data>.
29. National Institute for Health and Care Excellence. Clinical knowledge and guidance portal. <https://www.nice.org.uk/guidance>.

30. Uzbekistan Ministry of Health. Official public information portal.
<https://ssv.uz/>.