

**METABOLIC AND HORMONAL RISK FACTORS FOR ENDOMETRIAL
HYPERPLASIA IN REPRODUCTIVE-AGE WOMEN: A CRITICAL
INTEGRATIVE REVIEW OF PATHOBIOLOGY, REAL-WORLD
EVIDENCE AND PRECISION RISK ASSESSMENT****Xolboyeva Mashhura Yoqutjon qizi****Termez University of Economics and Service****Faculty of Medicine****ABSTRACT**

Background: Endometrial hyperplasia in reproductive-age women encompasses biologically heterogeneous lesions whose risk is shaped by chronic anovulation, unopposed estrogen exposure, adiposity, insulin resistance, diabetes and associated inflammation. The clinical distinction between hyperplasia without atypia and endometrial intraepithelial neoplasia/atypical endometrial hyperplasia (EIN/AEH) is essential because the latter can coexist with or precede carcinoma. **Objective:** To critically integrate mechanistic, clinical and regulatory evidence linking metabolic and hormonal exposures to endometrial proliferative disorders and to propose a transparent real-world research framework. **Methods:** An evidence-informed integrative review prioritised professional guidelines, systematic reviews and comparative studies, together with official methodological documents from ACOG, RCOG, WHO, FDA and EMA. Source reliability, pathological definitions, confounding, outcome ascertainment and transportability were assessed qualitatively. No systematic-review screening counts, effect estimates or pooled analyses were invented. **Results:** Chronic ovulatory dysfunction and excess estrogenic stimulation provide the dominant endocrine framework, whereas obesity adds increased peripheral aromatization, reduced sex hormone-binding globulin, altered adipokines and inflammatory signaling. Hyperinsulinemia may act through insulin and IGF-related pathways, but insulin resistance is not a validated stand-alone diagnostic test for endometrial hyperplasia. PCOS creates a clinically meaningful high-risk context, although BMI, age, bleeding, medications and tissue diagnosis must be evaluated individually. Histology remains decisive for classifying hyperplasia and EIN/AEH. Real-world datasets can estimate comparative treatment effectiveness and recurrence only when selection, confounding, exposure changes and missing biopsies are explicitly addressed. Artificial intelligence may support structured reporting and risk prediction but requires external validation, calibration and governance. **Conclusions:** Integrated metabolic and reproductive assessment is justified in persistent abnormal uterine bleeding, but no biomarker or machine-learning score replaces appropriate endometrial sampling. Fertility-sensitive management should follow histological risk category, informed preferences and adequate surveillance. Prospective registries and target-trial emulations would improve locally applicable evidence, including in Uzbekistan.

Keywords: endometrial hyperplasia; endometrial intraepithelial neoplasia; atypical hyperplasia; obesity; insulin resistance; chronic anovulation; polycystic ovary syndrome; estrogen; progesterone; metabolic syndrome; real-world evidence; causal inference; artificial intelligence; fertility preservation.

1. INTRODUCTION

1.1 Clinical framing

Endometrial hyperplasia (EH) is not synonymous with endometrial cancer, nor is every thickened endometrium an instance of hyperplasia. The diagnostic entity is defined histologically, and terminology matters. Hyperplasia without atypia and endometrial intraepithelial neoplasia/atypical endometrial hyperplasia (EIN/AEH) differ in underlying biology, concurrent carcinoma risk and recommended management. ACOG emphasizes careful exclusion of concurrent malignancy when EIN/AEH is suspected, while the RCOG guideline describes the central role of abnormal uterine bleeding in presentation. The etiologic question is therefore not merely which women exhibit ultrasound thickness but which endocrine-metabolic profiles predict biopsy-confirmed lesion categories and meaningful subsequent outcomes [1–3].

1.2 The reproductive-age problem

In women who have not reached menopause, physiological endometrial thickness changes throughout the cycle, making uncontextualised ultrasonographic thresholds misleading. Prolonged irregular bleeding, oligomenorrhea, chronic anovulation and infertility may prompt assessment, but clinical interpretation is inseparable from the last menstrual period, pregnancy exclusion, recent progestogen exposure and pregnancy intentions. The metabolic profile is neither incidental nor sufficient: adiposity, dysglycemia, polycystic ovary syndrome and chronic inflammation interact with ovulatory function. A tissue-based reference standard and an explicit clinical indication for sampling are critical to avoid both missed precancer and unnecessary invasive procedures [1–4].

1.3 RWD and RWE

Real-world data (RWD) are routinely collected information about patient status and healthcare delivery, including records, claims, registries and digital sources. Real-world evidence (RWE) is clinical evidence about use, benefits or risks generated by analysing RWD. RWD are not automatically causal evidence. Their utility depends on clear eligibility, exposure, outcome and comparator definitions, data provenance, longitudinal capture and adjustment for measured confounders. In EH research, links between prescriptions, pathology results, weight trajectories, glycemic indices, pregnancy attempts and hysterectomy provide a potential life-course view that small trials cannot easily achieve [5–8].

1.4 RCTs, efficacy and effectiveness

Randomized controlled trials (RCTs) can estimate efficacy under defined inclusion criteria and monitored treatment schedules, limiting confounding by indication through randomization. Effectiveness asks how the intervention performs across routine variation in adherence, access, comorbidity, availability of histological surveillance and fertility priorities. A hormonal regimen demonstrating regression in a trial may show different real-world persistence if women remove a device, discontinue tablets or do not obtain follow-up biopsies. Conversely, observational analyses can mislead when high-risk women are preferentially assigned particular interventions. Neither design can be substituted for the other without understanding its estimand and assumptions [5–9].

1.5 Regulatory context

The FDA guidance on routinely collected electronic records and claims stresses relevance and reliability before regulatory use. EMA promotes RWE through its DARWIN EU infrastructure as complementary evidence in medicines evaluation; these regulatory frameworks do not endorse observational studies indiscriminately. WHO health-AI guidance emphasizes ethics, human rights, transparency, accountability and oversight rather than performance metrics alone. For EH, these principles translate into independently adjudicated histopathology, clinically transparent models, reproducible exposure definitions and safeguards against missing or selectively documented disease [5–10].

1.6 Research objective

The objective of this integrative review is to specify plausible causal mechanisms through which metabolic and endocrine factors contribute to endometrial hyperplasia in reproductive-age women; distinguish established clinical risk contexts from exploratory molecular biomarkers; integrate high-quality evidence on histology, management and follow-up; and propose a research design for longitudinal clinical data. Drug-response modeling, pharmacovigilance and AI are considered methodological extensions of this precise gynecological question, rather than presented as unrelated evidence for other diseases. We explicitly separate findings supported by authoritative guidance from hypotheses awaiting validation.

2. MATERIALS AND METHODS

Review design. This is a critical integrative narrative review, not a systematic review or original clinical study. The manuscript does not claim executed exhaustive database queries, independent duplicate screening, PRISMA counts or a prospectively registered protocol. The source set was assembled from accessible official clinical and regulatory guidance and identifiable peer-reviewed publications. Sources with unverified numerical claims were not used to derive statistics. Searches recommended for full

systematic update include MEDLINE/PubMed, Embase, Scopus, Web of Science and Cochrane Library; journal-specific collections (JAMA, NEJM, The Lancet, Nature Medicine) can complement but do not replace systematic indexing. These suggested sources must not be interpreted as audited searches actually executed here.

Eligibility framework. Priority is assigned to reproductive-age EH/EIN/AEH pathology, abnormal uterine bleeding, metabolic factors (BMI, insulin resistance, type 2 diabetes, dyslipidemia), endocrine variables (chronic anovulation, PCOS, estrogen and progestogen exposure), histological outcomes, therapeutic regression, persistence, recurrence and cancer detection. Professional guidelines and peer-reviewed systematic reviews are privileged. Exclude studies defining EH solely by ultrasound thickness; mixed malignant/pre-malignant series without separable outcomes; case reports used as causal confirmation; and machine-learning demonstrations without independent validation. Studies predating 2020 are retained only where they establish biological, epidemiologic or methodological foundations.

Data synthesis. Claims were grouped by mechanistic plausibility, clinical consistency, measurement validity, study design and transportability. For every reported numeric estimate, the original underlying source and relevant clinical denominator should be retrieved before publication. We do not pool heterogeneous cohorts or assign precision intervals without individual study extraction. Distinguish relative risks of EH onset from odds of biopsy diagnosis, progression among diagnosed EIN/AEH, and relapse following treatment; these cannot be combined as if they described one endpoint.

2.1 Propensity score matching

Define a well-timed treatment decision and estimate each participant's conditional probability of receiving a treatment given pretreatment covariates. Matching can improve observed covariate balance between a levonorgestrel intrauterine system and continuous oral progestin cohort, but cannot eliminate unmeasured confounding, immortal-time bias or differential biopsy ascertainment. Specify the overlap region, caliper, balance standardized mean differences, matched-population estimand and excluded patients.

2.2 Inverse probability weighting

Estimate the propensity to receive each treatment, then construct weights for a defined marginal contrast. Stabilized weights and truncation are important when probabilities approach zero; time-varying exposure and confounding require longitudinal marginal structural models. Weighting does not compensate for inaccurate pathology definitions or limited positivity. Report effective sample size and weight distributions.

2.3 Cox regression and survival analysis

Specify time zero and an event such as first histologically documented recurrence or hysterectomy, accounting for different censoring processes. Cox regression yields

conditional hazard ratios under proportional hazards; compare Schoenfeld residuals or alternative time-varying effects. Noncancer hysterectomy may be a competing event for recurrence and should not automatically be treated as independent censoring.

2.4 Kaplan–Meier estimation

Use product-limit survival functions for time-to-event outcomes only when censoring is plausibly noninformative. When hysterectomy precludes detection of later recurrent lesions, cumulative-incidence estimation under competing risks may be more appropriate. Visits and sampling are irregular, so histological resolution is often interval-censored rather than observed at an exact biological time.

2.5 Target trial emulation

State eligibility, treatment strategies, allocation procedure to be emulated, baseline, grace period, outcomes, follow-up, intercurrent events and causal estimand before looking at treatment outcomes. Clone–censor–weight designs can address dynamic treatment assignment when properly defined. Emulations of conservative management must account for pregnancy intention, histological baseline risk, contraindications and treatment switches.

2.6 Causal inference

A directed acyclic graph can clarify BMI as confounder, mediator or collider depending on the exposure question. A claim that PCOS causes EH through insulin resistance is distinct from the prediction that women with PCOS have elevated biopsy yield. Measurement of estrogen, glucose and BMI at diagnosis can be consequences of prior care and thus inappropriate as baseline confounders. Negative controls and quantitative sensitivity analysis are recommended.

2.7 Machine learning

Penalized regression, gradient-boosted trees and random forests can model nonlinear interactions among bleeding, age, ovulatory status, adiposity and pathology history. Internal cross-validation must respect patient and site clustering. Report discrimination, calibration, decision curves, missingness strategy and external validation; avoid threshold hunting after seeing outcomes.

2.8 Deep learning

Digital histopathology and ultrasound representation learning may augment human pattern recognition; however, slide-scanner shift, staining variability and inconsistent labels can erode performance. Biological labels must distinguish benign hyperplasia from EIN/AEH. Deep-learning performance claims without external datasets and clinically defined comparators are not evidence of improved patient outcomes.

2.9 Natural language processing

NLP extracts structured symptoms, menstrual intervals, histology phrases and treatment intent from clinician narratives, but negation, historical diagnoses and copied notes can introduce errors. A reviewed concept dictionary and manual reference annotation are necessary. Large language models may assist information extraction, not autonomously adjudicate malignancy.

2.10 Missing data and causal validity

Missing biopsies may reflect improvement, disengagement, cost, treatment effects or clinician suspicion and are therefore unlikely to be missing completely at random. Multiple imputation is not automatically valid; characterize missingness, conduct pattern-mixture or tipping-point analyses, and use informative-visit models where indicated.

3. RESULTS: SYNTHESIS OF AVAILABLE EVIDENCE

The following results represent an interpretive synthesis of previously published evidence, not findings obtained from a newly recruited sample. Directional mechanisms and clinical recommendations are separated from hypotheses. No new sample size, relative risk, P value or confidence interval is calculated.

3.1 Classification and histological heterogeneity

The strongest immediate determinant of clinical action is lesion category. Hyperplasia without atypia commonly reflects hormone-dependent proliferation, whereas EIN/AEH carries precancerous significance and warrants careful assessment for concurrent carcinoma. Molecular instability is not distributed uniformly across benign proliferative disorders. Misclassifying EIN as nonatypical EH can distort apparent associations with obesity and age, because the compared groups differ at baseline in disease stage and in probability of hysterectomy [1–3].

The quality of pathology review is a structural part of epidemiology. Expert reassessment, consensus criteria and linkage to subsequent specimens reduce label noise. When an investigator treats a historical diagnostic code as ground truth, the apparent risk-factor gradient may partly represent changing terminology or pathologist practice. Reporting both the original label and central-review label, when available, is preferable to hiding discordance.

3.2 Unopposed estrogen and persistent anovulation

A prolonged proliferative state arises when estrogenic stimulation is insufficiently opposed by cyclic progesterone exposure. Reproductive-age anovulation may reduce regular luteal progesterone production even when circulating estrogen measurements fall within broad laboratory reference intervals. A single estrogen assay rarely reflects cumulative tissue exposure or the balance of local signaling. The relevant exposure is

temporal: repeated cycles without ovulation, irregular endometrial shedding and persistent proliferative cues [1–4].

Chronic anovulation should not be treated as an interchangeable marker for a PCOS diagnosis. Thyroid dysfunction, hyperprolactinemia, physiologic postmenarchal irregularity, medications and other endocrine causes must be considered. Histologic response to a progestogen provides treatment information, but cannot retrospectively prove a single causal hormonal pathway.

3.3 Adiposity and adipose endocrine activity

Adipose tissue can contribute to systemic and local estrogenic milieu through aromatase activity, while obesity is associated with alterations in sex hormone-binding globulin, chronic low-grade inflammation and insulin signaling. Central adiposity may convey biological information not fully captured by BMI. Yet adiposity is a composite measure: diet, medications, socioeconomic context, sleep and reproductive life history may confound observed associations [2–4].

Obesity also changes the clinical pathway: irregular bleeding may be assessed earlier in some settings, while procedure access and anesthesia risk may delay histology in others. Analyses should separate disease occurrence from likelihood of investigation. Treating BMI adjustment as universally correct may obscure adiposity-mediated pathways linking metabolic disease to endometrial proliferation.

3.4 Insulin resistance and hyperinsulinemia

Hyperinsulinemia can accompany reduced insulin sensitivity and may influence growth signaling through insulin/IGF-related receptors, PI3K/AKT and downstream proliferative mediators. These mechanisms plausibly interact with estrogenic effects, obesity-associated inflammation and altered ovarian steroidogenesis. However, a mechanistic hypothesis should not be transcribed as a validated clinical threshold for diagnosing EH. The relationship between fasting insulin and direct uterine tissue stimulation is neither uniform nor reliably isolated from confounding [2,4].

Studies should consider glycemic status, insulin resistance indices, longitudinal diabetes treatment, HbA1c trajectory and waist circumference as distinct constructs. An intervention-induced reduction in insulin concentration does not by itself establish prevention of histological progression. Tissue outcomes and clinically important follow-up remain indispensable.

3.5 Type 2 diabetes and metabolic clustering

Type 2 diabetes frequently co-occurs with obesity, dyslipidemia and hypertension, making attribution to hyperglycemia alone difficult. Medication changes, metabolic disease duration and inflammatory comorbidity can modify both endometrial biology and the chance of clinical review. Consequently, predictive associations between

diabetes codes and EH require sensitivity analyses with more granular biomarkers and treatment histories. Observational effect estimates without BMI and ovulatory data are especially vulnerable to residual confounding.

A multivariable risk score can still be useful for triage even without perfect etiologic decomposition; prediction and causation are different questions. The performance of such a score should be tested against transparent clinical judgment, with the costs of false-negative precancer considered explicitly.

3.6 PCOS as an intersecting endocrine phenotype

Polycystic ovary syndrome combines ovulatory dysfunction, hyperandrogenism and variably expressed metabolic dysfunction. Increased risk of endometrial proliferative pathology is plausibly mediated by prolonged anovulation, adiposity and insulin-related pathways, although phenotype heterogeneity means not every person with PCOS has the same underlying risk. The clinical task is to identify persistent abnormal bleeding and relevant additional risk factors, not to biopsy every woman with a PCOS label.

Temporal misclassification can emerge when a PCOS diagnosis is recorded after referral for bleeding. Research should distinguish incident PCOS recognition from longstanding ovarian dysfunction. Medication data are equally important: combined contraceptives or progestogen exposure may alter risks and observed thickness before formal pathology assessment.

3.7 Progesterone protection and treatment response

Progesterone counterbalances estrogen-driven proliferation through cell-cycle control, differentiation and endometrial secretory transformation. Sustained local progestogen delivery through an LNG-IUS is supported by Cochrane evidence for regression of EH compared with non-intrauterine progestogens, while ACOG distinguishes the higher-stakes management of EIN/AEH and the need for histological surveillance [1,4].

A treatment-comparison dataset must avoid mixing baseline histologies and follow-up durations. The device may remain in place longer than oral adherence persists; effectiveness estimates depend on whether switching and early removal are treated as failure, treatment-policy events or censoring. Adverse-event ascertainment is also treatment-specific.

3.8 Metabolic inflammation and local tissue ecology

Adipokines, macrophage signaling, oxidative stress and altered cytokine networks may modulate epithelial-stromal interactions. Endometrial tissue is an immunologically active environment, and hormonal responses are mediated by local cells as well as systemic concentrations. Such pathways offer candidate biomarkers for future stratification but remain insufficiently specific for an independent EH screening test.

Case-control molecular studies often sample tissue only after surgical selection, creating substantial selection bias. Molecular observations from hysterectomy specimens should not be transported uncritically to reproductive-age women undergoing an outpatient biopsy for irregular bleeding.

3.9 Imaging, histology and diagnostic discordance

Transvaginal ultrasound provides structural context and can identify focal lesions or coexisting pathology, but premenopausal endometrial thickness is cycle-dependent and cannot independently classify atypia. Histopathology is needed when clinical suspicion warrants assessment. Hysteroscopically directed sampling can be especially valuable when focal pathology or concurrent carcinoma is a concern, as reflected in ACOG recommendations for EIN/AEH [1].

Unselected use of a numerical endometrial thickness threshold may inflate false-positive results and delay identification of lesions not meeting the cut-off. Clinical decisions require bleeding severity, age, pregnancy status, risk profile and prior therapy. Screening models should evaluate whether they improve timely, appropriate sampling rather than maximize diagnostic labels.

3.10 Fertility preservation and competing goals

Reproductive-age management balances cancer prevention, bleeding control, histologic response, conception timing and tolerance of therapy. For EIN/AEH, fertility-sparing management is not interchangeable with treatment of benign hyperplasia and requires careful exclusion of cancer and close follow-up. For patients who do not pursue surgery, ongoing progestin therapy and surveillance are central [1].

A model predicting “successful treatment” is clinically incomplete unless it specifies whether success means regression at a defined biopsy, sustained remission, subsequent pregnancy, live birth, or avoidance of hysterectomy. These outcomes are not substitutable and may compete with one another.

3.11 Electronic health records

EHRs contain diagnoses, gynecologic notes, pregnancy status, BMI, glucose measurements, medication orders, biopsy requests and pathology narratives. Their strength is longitudinal clinical context; their weakness is nonstandard documentation, unlinked private laboratories and missing visits outside the health system. Diagnosis codes may reflect suspicion rather than histological proof.

For a credible EH registry, pathology accession numbers should be linked to coded diagnoses and specimen dates. Structured menstrual history and fertility intentions should be captured at baseline, with provenance metadata for imported results.

3.12 Electronic prescriptions and dispensing

E-prescription systems can identify oral progestogen orders, contraceptive prescriptions, metformin exposure and changes in treatment. Orders are not fills, and fills are not ingestion. Linkage to dispensing and device insertion/removal records is necessary to compare realistic treatment strategies.

Prescription histories can support safety surveillance but do not establish indications. A patient receiving metformin may have PCOS, diabetes or another indication; conditioning on drug receipt may create confounding by indication.

3.13 Laboratories and pathology information systems

Centralized laboratory feeds can provide HbA1c, lipid fractions, thyroid markers and pregnancy assays while histopathology systems contain the outcome definition. Quality assurance must address unit conversion, assay drift, repeat specimens and the distinction between diagnostic and surveillance biopsies.

Metabolic values measured shortly after treatment initiation may already be mediators. Rigorous analyses set measurement windows before time zero and distinguish missing testing from normal test results.

3.14 National registries and claims

Cancer registries, procedural datasets and insurance claims can extend follow-up, detect hysterectomy and estimate service utilization. Registries are limited by coverage, reporting delay, pathology granularity and access inequity. Claims are well suited to utilization and cost questions but often lack menstrual-cycle history, obesity severity or histologic subclassification.

A federated research architecture can preserve patient privacy while harmonizing a minimal dataset. Coverage differences across geography and providers must be tested before national generalization.

3.15 Wearables and mobile health

Wearable devices may capture physical activity, sleep, heart rate and some cycle-related variables; apps can register bleeding patterns, medication reminders and patient-reported symptoms. These signals may facilitate earlier recognition of persistent irregular bleeding but are not validated surrogate outcomes for premalignant tissue change.

App participation and device ownership are socially patterned, so models developed only among digitally active users risk inequitable performance. Explicit consent, de-identification and opt-out options should be fundamental, not optional afterthoughts.

3.16 Machine learning for risk stratification

Regularized logistic regression provides a clinically interpretable baseline against which complex models should be compared. Gradient-boosted trees may capture nonlinear interactions among age, BMI, diabetes, ovulatory history and bleeding; however, apparent gains are often overstated by random splits that distribute visits from one patient between training and test sets.

Model deployment should follow prospective evaluation with calibration by clinical site, decision-curve analysis and monitoring of missed EIN/AEH. Interpretable models may be preferable when training samples are small or clinical consequences of error are large.

3.17 Deep learning and computational pathology

Histopathology whole-slide images could support segmentation of glands, stromal context and morphologic irregularity. Unlike administrative diagnosis codes, image models work close to the disease-defining tissue phenotype. Yet training labels may contain interobserver disagreement, and deployment performance can vary by staining, tissue preparation and scanner.

External benchmarking should include expert-pathologist adjudication, borderline lesions, mixed specimens and uncertainty reporting. An algorithm predicting atypia is a second reader, not a substitute for the clinician or pathologist.

3.18 NLP and large language models

Clinical narratives frequently encode irregular bleeding duration, ovulation status, documented fertility preferences and reasons for choosing conservative therapy. NLP could extract such variables while preserving temporal and negation context. Large language models may draft research-ready summaries, but invented clinical details and inappropriate certainty create unacceptable risk.

A safe workflow requires retrieval anchored to the original chart, human review of generated labels, audit trails and strict control of identifiable patient text. Aggregate validation metrics should report errors involving pregnancy, cancer exclusion and treatment discontinuation.

3.19 Reinforcement learning limitations

Sequential decision-making seems attractive for individualized surveillance intervals, but reinforcement learning typically requires dense, reliable longitudinal state and reward information. EH care has infrequent biopsies and ethically asymmetric harms: missing progressive atypia is more consequential than an avoidable low-risk revisit.

Offline reinforcement learning on observational treatment histories may reproduce existing undertreatment and referral biases. Such systems should be regarded as

experimental and not deployed without prospective safeguards and regulatory assessment.

3.20 Diabetes treatment effectiveness as a transportability lesson

Comparative effectiveness methods developed for glucose-lowering medications demonstrate the importance of adherence, treatment changes and competing risk in longitudinal observational work. For EH, metabolic medications should be assessed only against meaningful endometrial outcomes, not assumed beneficial because they improve glycemic markers.

Linking metabolic treatment to biopsy-confirmed lesion regression demands control for baseline BMI, PCOS, initial histology, concurrent progestogen and differing probabilities of follow-up biopsy.

3.21 Cardiovascular, oncology, infection and neurology comparators

Across cardiovascular disease, oncology, infection and neurology, real-world treatment evaluation faces shared issues of confounding by indication, heterogeneous baseline risk, informative censoring and competing events. These therapeutic areas illustrate reusable study-design principles, not evidence that drugs used in those fields directly prevent EH.

Oncology provides particularly useful examples of pathology adjudication and time-varying endpoints; infectious disease highlights rapid treatment changes; neurology highlights long-term symptom capture. Mechanisms and treatment effects must not be transferred across diseases without disease-specific clinical validation.

3.22 Pharmacovigilance and ADRs

Safety monitoring for progestogens and intrauterine therapies should record irregular bleeding, pain, device expulsion, discontinuation, gastrointestinal symptoms and other clinically assessed adverse outcomes. Passive reports can identify signals but are affected by stimulated reporting and lack reliable exposure denominators.

Drug-safety signal detection may use disproportionality scores, but such signals are hypothesis-generating, not incidence estimates or causal proofs. Registry-based active surveillance with known exposure periods is preferable for estimating comparative risks.

3.23 Pharmacogenomics and biomarkers

Molecular signatures involving progesterone response, PTEN-associated alterations and other endometrial pathways are of research interest, but routine germline pharmacogenetic selection of EH progestogen therapy is not established. Tissue markers must be judged for incremental prognostic value above histology and clinical factors.

An analytically accurate biomarker may still have little clinical utility when results do not change treatment or surveillance. Tissue availability, affordability, assay repeatability and equitable access should be part of validation.

3.24 Precision medicine and decision support

Precision medicine in EH begins with clinically credible stratification: histology, symptoms, reproductive plans, metabolic context, contraindications and surveillance feasibility. A decision-support tool should identify the clinical action, provide calibrated probabilities and indicate uncertainty.

The reference care pathway remains human accountable. Automated recommendations must surface exceptions such as pregnancy possibility, severe bleeding and cancer suspicion, and explain why histologic confirmation is required.

3.25 Uzbekistan: digital implementation strategy

For Uzbekistan, a practical implementation route is a multicenter gynecology–pathology–endocrinology registry with harmonized diagnostic terms and interoperable identifiers. Ministry-led digitalization and e-prescription initiatives can create useful infrastructure only if technical interoperability, authorization, privacy protection and pathology data quality are verified locally; neither national coverage nor clinical completeness should be assumed.

An initial pilot could capture menstrual history, PCOS features, BMI, HbA1c, imaging context, tissue category, prescriptions, treatment changes, repeat biopsy and fertility outcomes. The pilot should publish coverage and missingness audits before reporting associations, then test externally across regions.

3.26 Uzbekistan: clinical pharmacology and governance

Clinical pharmacology services could contribute to progestogen exposure definitions, medication reconciliation and active adverse-event monitoring. The scientific agenda should connect specialists in gynecology, endocrinology, pathology, statistics, health informatics and bioethics. A national protocol would benefit from prespecified target-trial questions and independent data governance.

Before using AI at the point of care, institutions should establish bias assessment, security review, annotation policy, local validation and human override. Public reporting of limitations is a clinical safety function rather than a publicity weakness.

3.27 Temporal structure of the exposure

An exposure recorded at first biopsy may not represent the biologically relevant window. Longstanding obesity, repeated anovulatory cycles, newly treated diabetes and short-term hormonal contraception differ in duration and accumulation. A longitudinal design should record the earliest reliable occurrence of each exposure and define latency windows. Landmark analyses may help distinguish measured baseline

risk from treatment-related changes; nevertheless, landmark selection itself can bias the target population when survival to the landmark is related to disease severity.

3.28 Outcome definition and verification

An apparently lower incidence of atypia among patients with fewer biopsies may represent under-ascertainment rather than protection. Histology-confirmed regression requires a defined specimen, adequate sampling and an assessment time window. A composite endpoint of “no recorded EH” is generally inadequate, because unobserved disease cannot be classified as resolution. Sensitivity analysis could compare complete follow-up, inverse-probability-of-observation adjustment and conservative assumptions about unsampled patients.

3.29 Selection pathways

Referral to a tertiary fertility clinic depends on infertility duration, insurance or payment capacity, geography and clinician preferences; each may correlate with metabolic status and likelihood of biopsy. A case series from such a clinic does not represent all women of reproductive age. Transportability analysis should compare distributions of age, BMI, parity, bleeding characteristics and access to pathology between source and intended target populations.

3.30 The estrogen measurement paradox

When single-point estradiol levels are not elevated, a patient may still experience prolonged unopposed estrogenic activity because the absence of luteal progesterone changes the temporal hormonal balance. Conversely, one isolated high serum estrogen measurement does not diagnose hyperplasia. Research relying on dichotomized estrogen alone loses cycle timing and fails to capture endometrial responsiveness. Better endocrine phenotyping combines cycle history, ovulation evidence, medication history and appropriately timed assays.

3.31 Metabolic mediation

A causal mediation analysis might ask how much of the association between PCOS and EH operates through obesity or insulin resistance. Yet mediators and confounders can affect one another over time, and the assumptions for natural direct and indirect effects are stringent. The resulting percentages are particularly unreliable when BMI and insulin are measured after the development of bleeding. Mechanistic language should therefore remain proportionate to data quality.

3.32 Clinical risk communication

A clinician may communicate both relative association and absolute expected benefit, but without a validated reproductive-age prediction model it is inappropriate to assign an individualized numeric probability. Communicating uncertainty does not imply indifference: persistent abnormal bleeding with relevant metabolic risk merits timely

clinical evaluation. Risk discussions should avoid stigmatizing women with higher body weight or assuming that lower BMI excludes precancerous lesions.

3.33 Intervention fidelity

Pharmacological treatment can be described at prescription, dispensing, initiation, persistence and biologically effective exposure levels. Device insertion date and documented removal are critical for hormonal intrauterine therapies. Intention-to-treat analogues in observational data answer a different question from per-protocol analyses that require adjustment for adherence-related confounders. Carefully specifying the estimand prevents selective reporting of whichever analysis looks most favorable.

3.34 Biopsy as both outcome and intervention

Biopsy is not merely a neutral observation. Its performance is shaped by bleeding intensity, previous reassuring studies, fertility plans and clinician judgment, and it can trigger treatment changes. If observation schedules depend on latent disease risk, standard regression may misrepresent remission. Joint models of examination intensity and disease transition, although methodologically demanding, may clarify the distortion introduced by irregular surveillance.

3.35 Digital equity

A digital study may overrepresent urban residents, insured patients or those with smartphones. If wearable-derived activity improves model discrimination only because it proxies socioeconomic access, the apparent advantage may disappear during wider deployment. Model auditing should present subgroup calibration and data completeness; alternative low-technology pathways must remain available for patients with limited connectivity.

3.36 Clinical meaning of explainability

Variable importance is not equivalent to causal attribution. An explanation indicating that obesity influenced a score cannot establish that weight loss would change the predicted histology by the same amount. Counterfactual outputs can be particularly misleading if the predictor represents age, pre-existing pathology or another nonmanipulable exposure. Clinical explainability requires understandable decision logic and documentation of uncertainty, not just graphical attribution.

3.37 Surveillance as a care pathway

After histologic diagnosis, the length and intensity of follow-up depend on lesion category, therapy, response and patient goals. For EIN/AEH treated conservatively, ACOG recommends repeat histologic assessment within three to six months [1]. This recommendation should not be indiscriminately copied to patients with other diagnoses or to all symptom-free women. Documentation of missed visits is crucial for both patient safety and effectiveness research.

3.38 Regulatory interpretation

International regulatory use of RWE places emphasis on fit-for-purpose data, prespecified protocols, reproducibility and bias control. DARWIN EU is a regulatory evidence network, not itself an EH clinical guideline. FDA medical-product guidance should be cited as methodological context rather than as proof of endocrine causation. WHO AI principles are normative governance guidance, distinct from disease-specific randomized evidence.

3.39 Drug-safety interpretation

Spontaneous adverse-event reporting is designed to discover potential problems, not estimate how many women will experience them. A comparison of raw report counts for LNG-IUS and oral progestogens would be particularly misleading because time at risk, reporting behavior and indications differ. Active safety studies should define initiation, exposure gaps, device status and competing pregnancy, and distinguish expected bleeding patterns from severe harms.

3.40 Value of negative findings

If a complex AI model does not outperform an interpretable model using age, bleeding history, ovulation status and BMI, this is clinically useful. The finding could reduce cost and improve reproducibility. Similarly, a metabolic biomarker that fails to discriminate EIN/AEH after adjustment for histology and reproductive context should not be added to routine workflows merely because its molecular mechanism seems compelling.

3.41 Replication and generalizability

External replication across laboratories and clinics is indispensable because histology prevalence, health-system coding, medication access and ultrasound practice vary. Internal random-split testing cannot establish geographic transportability. A reproducible package should specify data dictionaries, phenotype algorithms, temporal definitions, statistical plans, deidentification procedures and sensitivity analyses without releasing personally identifiable data.

3.42 Research priorities

The most valuable next studies are prospective reproductive-age cohorts with biopsy-confirmed categories, serial ovulation and metabolic measures, and treatment records sufficient to distinguish intervention effects from secular changes. Comparative effectiveness analyses should prespecify differential biopsy surveillance as a central threat. Cost-effectiveness and equity analyses should test whether risk stratification improves timely access to appropriate tissue diagnosis rather than simply reducing the total number of procedures.

3.43 Estrogen receptor signaling across cell types

Estrogen-mediated transcription operates within epithelial and stromal compartments, but the response of endometrial tissue depends on local receptor abundance and co-regulators as well as circulating ligand. Cell-cycle progression, differentiation and repair participate in normal cyclic remodeling. In chronic anovulation, repeated proliferative signaling may persist without a physiologic period of progesterone-dominant maturation. The resulting biological narrative is persuasive, yet observational studies rarely measure intratissue estrogen levels, receptor states and the preceding ovulatory history together. Associations between serum estradiol and tissue classification are therefore limited by timing, tissue heterogeneity and previous hormone exposure. Research should predefine collection relative to the menstrual phase and separately evaluate current use of estrogen-containing therapy.

3.44 Progesterone biology and resistance

A progestogen can suppress proliferation through several pathways, and its effect depends on dose, delivery site, adherence and disease biology. Histological nonresponse may reflect missed doses, device expulsion, inaccurate baseline diagnosis, biologically less responsive tissue or persistent hormonal stimulation. The phrase progesterone resistance should not become a substitute for testing those explanations. Future trials that stratify response by receptor expression should still be anchored in reproducible pathology and clinically meaningful outcomes. They must distinguish tissue markers obtained before treatment from markers altered by therapy and should evaluate whether measurement actually changes management rather than merely correlating with later regression.

3.45 PCOS ascertainment and age

PCOS criteria and phenotype composition vary among clinics, so an administrative PCOS code cannot be assumed to indicate objectively verified anovulation. Diagnosis may occur during infertility care, dermatologic referral or obesity evaluation; each pathway produces different follow-up and biopsy probabilities. In adolescents and younger reproductive-age women, cycle irregularity may have different clinical implications than in longstanding adult ovulatory dysfunction. Cohorts should avoid assuming that the same number of missed periods translates to identical tissue risk across ages. An informative minimum dataset records diagnostic criteria, age at cycle irregularity onset, parity, metabolic history, treatments and the reason for endometrial assessment.

3.46 Weight change and reverse causation

Weight reduction is often recommended for overall metabolic health, yet studies correlating weight change with EH remission face difficult time ordering. Diagnosis can trigger treatment, dietary changes and more frequent follow-up; women with favorable early responses may also be more able to maintain care. It would be

inappropriate to attribute a later biopsy result exclusively to reduced weight without accounting for progestogens and patient selection. A rigorous analysis would define baseline before counseling and treatment, measure both weight and therapy longitudinally, and use methods for treatment-confounder feedback. Results should distinguish the goal of metabolic improvement from the specific causal claim of histological regression.

3.47 Metformin and adjunctive metabolic therapy

Metformin is sometimes prescribed in the context of diabetes or PCOS, creating a clinical hypothesis about improving the metabolic environment accompanying EH. Nevertheless, evidence of improved glycemia or cycle regularity is not the same as evidence that the drug alone prevents or reverses premalignant histology. A comparison of metformin users and nonusers is also susceptible to strong confounding: baseline disease severity and indication differ. An appropriate target trial would define eligibility, concurrent progestogen, treatment strategies, timing of biopsy and meaningful endpoints before observational analysis. Pharmacologic benefit should not be asserted solely on the basis of a plausible insulin pathway.

3.48 Distinguishing EH from carcinoma

The possibility of concurrent carcinoma makes the initial pathological and imaging workup especially consequential for EIN/AEH. A series limited to conservatively managed patients may underestimate cancer prevalence if women with suspicious features were triaged immediately to surgery. Surgical pathology and office biopsy are different verification pathways, each with distinct selection processes. A responsible synthesis therefore does not extrapolate concurrent-cancer proportions from an operated sample to all reproductive-age women presenting with irregular bleeding. Clinical action must be governed by established recommendations and individualized assessment, with transparent recognition of sampling uncertainty.

3.49 Economic considerations

Healthcare systems must compare the costs of repeated histology, imaging, long-term progestogen therapy and surgery against the consequences of delayed diagnosis and avoidable procedures. An algorithm that reduces biopsies may appear efficient while concealing an unacceptable number of missed EIN/AEH cases. Economic evaluations should report their time horizon, payer or societal perspective, resource inputs and assumptions about access to pathology. In low-resource settings, cost-effectiveness depends on service availability as much as discrimination. Models should avoid pricing a technology deployment without including clinician training, cybersecurity, maintenance and quality monitoring.

3.50 Prospective validation design

A practical prospective study could enroll consecutive reproductive-age women presenting with predefined abnormal uterine bleeding and collect a standardized questionnaire before diagnostic decisions. Variables would include age, menstrual pattern, pregnancy possibility, reproductive plans, body measurements, diabetes markers and hormone exposure. A prespecified pathology protocol would document specimen technique and expert classification. Investigators should register outcomes and statistical plans before examining associations, and should assess whether missing tissue samples differ systematically from completed assessments. This approach is slower than mining unrestricted records but much stronger for testing transportability and reducing diagnostic verification bias.

3.51 Testing interaction effects

Biological arguments frequently assume that obesity and anovulation amplify one another, but identifying statistical interaction requires more than comparing subgroup P values. Interaction may differ on additive and multiplicative scales, and adequate sample size is needed for clinically meaningful estimates. Researchers should specify the scale, covariate set and biological hypothesis before analysis. Apparent sex-hormone or BMI interactions can arise from measurement error or selective biopsy. When data are sparse, a clearly labeled exploratory description is more credible than a highly parameterized model with unstable estimates and narrow-looking but misleading intervals.

3.52 Longitudinal recurrence and treatment switching

Many patients change therapy because of bleeding, side effects, plans for conception or inadequate histological response. These changes are informative. Naively classifying everyone by their final therapy uses future information and can introduce severe time-related bias. A study should choose between an initiation strategy and a sustained-use strategy at baseline; different questions may warrant separate analyses. Censoring after switching can itself be selective and may require longitudinal weighting. Outcomes need standardized recurrence criteria, appropriate accounting for intervals between biopsies and a plan for hysterectomy as an event that prevents later observation of the untreated endometrium.

3.53 Cross-border evidence transport

Endometrial risk estimates from highly specialized centers or high-income insurance databases need not transfer directly to Uzbekistan. Population differences in obesity distribution, fertility patterns, health-seeking behavior, pathology access and medication availability can alter both baseline risks and observed intervention effectiveness. Validation must test clinical concepts rather than automatically transfer code lists. A local study could initially compare a small group of expert-verified cases against existing records to quantify agreement, then expand geographically once

reproducibility has been demonstrated. Such a staged approach is more defensible than immediate national AI deployment.

3.54 Governance and patient trust

Endometrial datasets are sensitive because they contain reproductive history, menstrual information, medication exposure and potential cancer diagnoses. Consent and lawful processing are not merely legal formalities; inadequate controls can deter participation and distort the dataset. A governance structure should include patient representatives, clinical custodians and independent oversight. Dataset minimization, role-based access, encryption and retention policies should accompany audit trails. Model providers should not receive more identifiable clinical material than necessary. Researchers should disclose conflicts and publish limitations, and all automated alerts should support, never silently replace, responsible clinical judgment.

3.55 Interpretive boundaries

This review integrates guidelines, biological concepts and methodological recommendations rather than generating a formally pooled statistical estimate. It does not establish independent effect sizes for individual exposures or quantify how much disease could be prevented by modifying one factor. The absence of a verified study denominator also precludes a PRISMA study-flow diagram or evidence-grade calculation for every assertion. Readers should treat proposed registry fields and causal designs as a research agenda. The strongest immediately actionable conclusions concern appropriate evaluation of persistent abnormal bleeding, attention to metabolic and ovulatory context, accurate histologic distinction and surveillance according to risk category.

3.56 Metabolic syndrome is not one exposure

Metabolic syndrome combines several correlated findings and may offer a convenient clinical shorthand, but different definitions can classify the same patient differently. Treating the syndrome as a singular mechanism also obscures whether adiposity, hypertension, hyperglycemia or dyslipidemia drives a specific association. For EH research, the components should be measured separately and modeled with explicit attention to collinearity. A composite index might improve prediction while offering less biological interpretability. When reporting individual components, avoid confusing statistical adjustment with evidence that an isolated intervention changes endometrial tissue risk. Clear definitions, reproducible cut-offs and temporally appropriate measurements are prerequisites to credible between-study comparisons.

3.57 The role of reproductive history

Nulliparity, infertility, pregnancy history and exogenous hormone use may capture duration of anovulatory exposure, healthcare contact and reproductive intentions. Yet their relationships are complex. Infertility can prompt more frequent ultrasound and

tissue sampling, increasing recorded lesions independently of underlying incidence. Pregnancy alters both the endocrine environment and the opportunity for detection. Analyses should clarify whether parity is an exposure, mediator or confounder under the causal question being asked. Clinic-based studies additionally need to compare patients entering through infertility services with those referred for heavy or irregular bleeding. A common regression adjustment cannot automatically resolve selection through different referral routes.

3.58 Differential diagnosis of bleeding

Not every bleeding disorder in reproductive-age women is caused by endometrial hyperplasia. Pregnancy complications, leiomyomas, polyps, adenomyosis, coagulopathy, thyroid dysfunction, medication effects and ovulatory disorders can produce overlapping presentations. A research cohort that assumes any irregular bleeding confirms EH will suffer major outcome misclassification. Conversely, a strategy focusing solely on metabolic disease may fail to recognize relevant structural pathology. Clinical protocols should articulate pregnancy exclusion, symptom assessment, physical examination, imaging when indicated and targeted histology. The purpose of metabolic risk stratification is to improve appropriate evaluation, not to bypass differential diagnosis or reproduce simplistic high-risk labels.

3.59 Interpreting a biopsy-negative result

A biopsy without evidence of atypia is reassuring only in relation to specimen adequacy, clinical suspicion and the method of sampling. Focal lesions may be missed, and persistent or recurrent symptoms justify reassessment rather than automatic reliance on a historical negative specimen. Pathology should document whether the sample is sufficient for evaluation, while clinical records should specify subsequent bleeding, medication use and imaging findings. Researchers who classify every negative or inadequate sample as absence of disease risk serious nondifferential and sometimes differential misclassification. Treatment studies should preserve distinct categories for resolved lesion, no sample, inadequate sample and unknown follow-up.

3.60 Patient-reported outcomes

For affected women, therapeutic success may include predictable bleeding, relief from anemia-related fatigue, acceptable side effects, reproductive autonomy and reduced anxiety. Histologic regression remains fundamental in precancer surveillance, but it is not the only valued outcome. Symptom instruments and reproductive-goal measures should be selected in advance, translated carefully and tested for comprehension in the intended population. Electronic reminders and app-based symptom capture might increase longitudinal completeness, but those refusing digital tools should not disappear from follow-up cohorts. Quality-of-life improvements ought to be analyzed separately from disease modification.

3.61 Multidisciplinary clinical workflow

High-quality care requires gynecologists, histopathologists, endocrinologists, primary-care practitioners and clinical pharmacologists to share accurate information. The pathologist determines the tissue category; the treating team evaluates bleeding, risk factors, fertility preferences and therapeutic options. Endocrine input may help manage poorly controlled diabetes or ovulatory dysfunction, but does not replace histologic assessment when clinically indicated. Clinical pharmacologists can review progestogen regimens, interactions and adherence. A shared dashboard should show the latest tissue classification, treatment start, scheduled surveillance, safety events and missed visits, with responsibility clearly assigned for acting on overdue follow-up.

3.62 Transparency of clinical thresholds

Decision thresholds should reflect consequences, not merely statistical convenience. If a risk model aims to recommend sampling, a false negative involving EIN/AEH may carry much greater harm than a false positive leading to a careful diagnostic conversation. Yet routine biopsy for every patient regardless of indication may also cause pain, cost and unnecessary intervention. Threshold selection should therefore combine local prevalence, clinical utility, patient values and available diagnostic capacity. Performance metrics must include sensitivity at clinically selected operating points, positive predictive value under realistic prevalence, calibration and consequences of referral. A high overall accuracy can be misleading when the clinically critical category is infrequent.

3.63 Implementation evaluation

Before scale-up, a prospective pilot should measure whether the proposed registry or decision-support workflow improves process outcomes: valid pathology coding, fewer lost biopsy results, timely follow-up appointments and documented discussion of treatment choice. Later evaluation may consider recurrence ascertainment, unnecessary procedures, equity and cost. An interrupted time-series or stepped implementation design might evaluate system changes while accounting for secular trends, although randomization of sites can offer stronger internal validity when feasible. Implementation reports should distinguish technical uptime and record completeness from improved clinical outcomes; success in data capture is not itself proof of reduced cancer risk.

4. DISCUSSION

4. Synthesis and clinical hierarchy

Metabolic risk profiles are clinically important but cannot supplant histologic classification. The central mechanistic theme is sustained proliferative signaling under prolonged ovulatory dysfunction, reinforced in some women by adipose endocrine changes and hyperinsulinemic states. This is a model with multiple interacting

pathways, not a single causal marker. The heterogeneity of PCOS, obesity and diabetes argues against blanket screening rules. Evidence relating to EIN/AEH management must not be extrapolated without qualification to all cases of benign hyperplasia [1–4].

4. Evidence limits and bias

Most routinely collected datasets measure care rather than biology. There is risk of confounding by indication, differential biopsy intensity, immortal-time bias, reverse causality, omitted reproductive history and differential losses to follow-up. Statistical adjustment addresses only a subset of these problems. A transparent study should retain negative findings, measure covariate overlap, specify missingness assumptions, document pathology verification and identify when results are too uncertain to support clinical change.

4. AI performance and safety

Machine learning may improve triage and surveillance logistics, but the relevant endpoint is safer and more equitable care, not merely a high area under the curve in a retrospective file. External validation should monitor missed atypia, calibration and clinical net benefit across care sites and socioeconomic strata. WHO emphasizes governance, human oversight and accountability; an LLM-generated clinical summary should not replace confirmatory pathology [9–10].

4. Regulatory science

The FDA and EMA approaches to RWE illustrate a distinction between data volume and evidentiary validity. Large networks can still produce unreliable answers when diagnoses are not validated or the target trial is poorly specified. For Uzbekistan, data interoperability and standardized pathology vocabulary would therefore provide greater scientific value than a large but inconsistently coded repository [5–8].

4. Patient-centered interpretation

The same histological label can lead to different acceptable treatment pathways depending on fertility goals, medical comorbidity, patient preferences and access to repeat sampling. Counseling should address what is known, uncertain and subject to surveillance. Decisions about conservative management of EIN/AEH require particular care to avoid undertreating concurrent carcinoma. Neither molecular biomarkers nor digital models currently justify bypassing established diagnostic safeguards [1].

5. CONCLUSION

Endometrial hyperplasia in reproductive-age women is linked to prolonged unopposed estrogen action and to interrelated endocrine-metabolic factors, including chronic anovulation, PCOS, adiposity and dysglycemia. Their interpretation requires separation of prediction, biological plausibility and causal effects. Clinical management turns on tissue-defined nonatypical hyperplasia versus EIN/AEH, cancer exclusion where appropriate, reproductive goals and robust follow-up. Well-specified

RWD analyses and validated AI may improve implementation, but neither substitutes for representative cohorts, pathology quality or accountable clinical assessment. Multicenter registry development in Uzbekistan should begin with standardized diagnoses, linked pathology, treatment histories and prespecified methodological safeguards.

REFERENCES

1. American College of Obstetricians and Gynecologists' Committee on Clinical Consensus–Gynecology. Management of endometrial intraepithelial neoplasia or atypical endometrial hyperplasia: ACOG Clinical Consensus No. 5. *Obstet Gynecol.* 2023;142(3):735–744. doi:10.1097/AOG.0000000000005297.
2. Ring KL, Mills AM, Modesitt SC. Endometrial hyperplasia. *Obstet Gynecol.* 2022;140(6):1061–1075. doi:10.1097/AOG.0000000000004989.
3. Royal College of Obstetricians and Gynaecologists. Management of endometrial hyperplasia. Green-top Guideline No. 67. London: RCOG; 2016. Available from: <https://www.rcog.org.uk/guidance/browse-all-guidance/green-top-guidelines/management-of-endometrial-hyperplasia-green-top-guideline-no-67/>
4. Mittermeier T, Farrant C, Wise MR. Levonorgestrel-releasing intrauterine system for endometrial hyperplasia. *Cochrane Database Syst Rev.* 2020;(9):CD012658. doi:10.1002/14651858.CD012658.pub2.
5. US Food and Drug Administration. Real-world evidence. Silver Spring: FDA. Available from: <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>
6. US Food and Drug Administration. Real-world data: assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products. Guidance for industry. 2024. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory>
7. US Food and Drug Administration. Considerations for the use of real-world data and real-world evidence to support regulatory decision-making for drug and biological products. 2023. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug>
8. European Medicines Agency. Real-world evidence. Amsterdam: EMA. Available from: <https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/real-world-evidence>

9. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO; 2021. ISBN: 9789240029200.
10. World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Geneva: WHO; 2024. ISBN: 9789240084759.
11. European Medicines Agency. Real-world evidence framework to support EU regulatory decision-making: third report on experience gained with regulator-led studies, February 2024–February 2025. Amsterdam: EMA; 2025.
12. US Food and Drug Administration. Use of real-world evidence to support regulatory decision-making for medical devices. Guidance. 2025. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-real-world-evidence-support-regulatory-decision-making-medical-devices>
13. European Medicines Agency. EMA and FDA set common principles for AI in medicine development. EMA; 2026. Available from: <https://www.ema.europa.eu/en/news/ema-fda-set-common-principles-ai-medicine-development-0>
14. World Health Organization. WHO releases AI ethics and governance guidance for large multi-modal models. News release. 2024 Jan 18. Available from: <https://www.who.int/news/item/18-01-2024-who-releases-ai-ethics-and-governance-guidance-for-large-multi-modal-models>
15. International Council for Harmonisation. E9(R1): Addendum on estimands and sensitivity analysis in clinical trials. Geneva: ICH; 2019.
16. International Council for Harmonisation. E6(R3): Guideline for good clinical practice. Geneva: ICH; 2025.
17. World Health Organization. Global strategy on digital health 2020–2025. Geneva: WHO; 2021.
18. World Health Organization. Classification of female genital tumours. WHO Classification of Tumours. 5th ed. Lyon: International Agency for Research on Cancer; 2020.
19. European Medicines Agency. Data Analysis and Real World Interrogation Network (DARWIN EU). Amsterdam: EMA. Available from: <https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/real-world-evidence>

20. National Institute for Health and Care Excellence. Heavy menstrual bleeding: assessment and management. NICE guideline NG88. London: NICE; 2018 (updated). Available from: <https://www.nice.org.uk/guidance/ng88>
21. US National Library of Medicine. PubMed: biomedical literature database. Bethesda: National Library of Medicine. Available from: <https://pubmed.ncbi.nlm.nih.gov/>
22. Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. doi:10.1136/bmj.i4919.
23. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758–764. doi:10.1093/aje/kwv254.
24. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41–55. doi:10.1093/biomet/70.1.41.
25. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399–424. doi:10.1080/00273171.2011.568786.
26. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. *PLoS Med*. 2007;4(10):e296. doi:10.1371/journal.pmed.0040296.
27. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55–63. doi:10.7326/M14-0697.
28. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565–574. doi:10.1177/0272989X06295361.
29. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
30. Little RJA, Rubin DB. Statistical analysis with missing data. 3rd ed. Hoboken: Wiley; 2019.
31. Pearl J. Causality: models, reasoning, and inference. 2nd ed. Cambridge: Cambridge University Press; 2009.
32. Hernán MA, Robins JM. Causal inference: what if. Boca Raton: Chapman & Hall/CRC; 2020.

33. Riley RD, van der Windt D, Croft P, Moons KGM, editors. Prognosis research in healthcare: concepts, methods, and impact. Oxford: Oxford University Press; 2019.
34. World Health Organization. Recommendations on digital interventions for health system strengthening. Geneva: WHO; 2019.
35. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170(1):51–58. doi:10.7326/M18-1376.