Review

Mobile Apps for Women's Physiology Tracking: Structured Narrative Review

Tomasz Cedro[1,2,3*], BSc, MSc; Wojciech Glinkowski[3,4*], MD, PhD

[1]IQCREDO, Warsaw, Poland
[2]CeDeROM, Warsaw, Poland
[3]Polish Telemedicine and eHealth Society, Warsaw, Poland
[4]Department of Medical Informatics and Telemedicine, Medical University of Warsaw, Warsaw, Poland
*all authors contributed equally

Corresponding Author:

Tomasz Cedro, BSc, MSc
IQCREDO
Twarda 18
Warsaw 00-105
Poland
Email: tomasz.cedro@telemedycyna.org

Abstract

Background: Mobile health (mHealth) apps for menstrual cycle and fertility tracking are widely used to support self-monitoring, reproductive planning, and health awareness among women. These tools also promise personalized predictions and convenient access to reproductive health information. Still, concerns persist regarding their clinical accuracy, adaptability to irregular cycles, algorithmic transparency, and real-world user experience.

Objective: This structured review aimed to evaluate the features, physiological integration, predictive performance, validation practices, and user-reported outcomes of mobile apps designed for menstrual and fertility tracking, and to contextualize the current evidence using the COSMIN (Consensus-based Standards for the Selection of Health Measurement Instruments) and ISPOR (International Society for Health Economics and Outcomes Research) evaluation frameworks.

Methods: A structured narrative review of studies evaluating mobile apps for menstrual cycle and fertility tracking was conducted. Literature published between January 2013 and October 2025 was identified through searches of PubMed, Embase, Scopus, and Web of Science, supplemented by semantic and citation-based searches in Semantic Scholar, OpenAlex, and Google Scholar. AI-assisted prioritization was used only to support screening efficiency, whereas human reviewers performed all eligibility decisions and data verification. Forty studies met the predefined eligibility criteria and were included in the qualitative synthesis. Because of substantial heterogeneity in study designs, outcomes, and validation methods, no meta-analysis was performed.

Results: Of the 40 included studies, most were observational and relied on self-reported data from a predominantly high-income, technologically literate population. Twenty-four apps incorporated physiological inputs, such as basal body temperature, luteinizing hormone measurements, or wearable-derived metrics, whereas others relied primarily on calendar-based predictions. Multiparameter and sensor-augmented approaches generally demonstrated higher agreement with biological or clinical reference standards than calendar-only methods, with selected studies reporting fertile-window prediction accuracies approaching 85% to 90% under optimal adherence conditions. However, only a small subset of apps has reported formal clinical validation or regulatory clearance. User satisfaction was strongly associated with perceived accuracy, personalization, and usability, whereas inaccurate predictions, particularly among users with irregular cycles, were linked to frustration, anxiety, and high attrition.

Conclusions: Menstrual and fertility tracking apps that integrate physiological signals outperform calendar-based approaches in terms of predictive performance; however, robust clinical validation, transparency, and inclusivity remain limited. Reported accuracy metrics should be interpreted cautiously because real-world adherence, irregular cycle patterns, and algorithmic bias substantially affect reliability. These tools are best positioned as decision-support and self-awareness technologies, rather than as autonomous diagnostic instruments. Future evaluations should apply standardized frameworks, such as COSMIN and ISPOR, explicitly communicate uncertainty, and prioritize diverse and irregular-cycle populations to ensure equitable and clinically meaningful digital reproductive health solutions.

JMIR Mhealth Uhealth 2026;14:e92968; doi: 10.2196/92968

Keywords: mobile health; mHealth; fertility; fertility awareness; physiological monitoring; wearable sensors; wearable health monitoring; user experience; decision-support tools

Introduction

Mobile health (mHealth) apps have become integral tools for reproductive self-care, particularly for menstrual and fertility tracking, cycle awareness, and contraception- or conception-related decision-making [1-3]. Mobile apps for menstrual cycle and fertility tracking have been downloaded hundreds of millions of times worldwide, with leading apps such as Flo, Clue, and Period Tracker ranking among the most widely used digital reproductive health platforms [4-10]. Many menstruating users now use these tools to manage their cycles, plan or prevent pregnancy, and monitor symptoms such as pain, mood, and energy as part of their everyday health practices [3,5].

During the COVID-19 pandemic, disruptions in access to in-person sexual and reproductive health services coincided with increased reliance on remote and digital solutions, further embedding menstrual and fertility-tracking apps within broader telehealth and mHealth ecosystems [11,12]. Despite their promise, many menstrual and fertility-tracking apps lack robust clinical validation, are poorly tailored for users with irregular cycles, and show variable usability and information quality across populations and platforms [13-17].

There is a clear need for a comprehensive critical evaluation of these digital tools, not only in terms of predictive accuracy but also in terms of equity, transparency, and clinical relevance. This review assesses the features, physiological integration, validation practices, and user experiences of menstrual and fertility-tracking apps using structured review methods. It situates the findings within internationally recognized frameworks, including the COSMIN (Consensus-based Standards for the Selection of Health Measurement Instruments) [18-20] and ISPOR (International Society for Health Economics and Outcomes Research) [21,22].

Methods

Study Design

This study was conducted as a structured narrative review of mobile apps for menstrual cycle and fertility tracking. A structured approach was used for study identification, screening, eligibility assessment, and inclusion. However, it was not designed or registered as a fully systematic review, and no meta-analysis was planned or performed. This methodological framing was chosen because the available evidence was highly heterogeneous in terms of study designs, app types, tracked parameters, validation approaches, and reported outcomes.

Data Sources and Search Strategy

Literature searches were conducted in PubMed, Embase, Scopus, and Web of Science to identify peer-reviewed articles published between January 1, 2013, and October 30, 2025. Search strategies combined controlled vocabulary and free-text terms related to mHealth, menstrual cycle tracking, fertility tracking, ovulation prediction, women's health, and physiological parameters such as basal body temperature (BBT), luteinizing hormone (LH), and wearable-derived metrics.

To broaden coverage and reduce reliance on individual database indexing, additional semantic and citation-based searches were performed using Semantic Scholar, OpenAlex, and Google Scholar. Because the exploratory searches identified a large and heterogeneous body of literature, relevance-focused prioritization and structured screening procedures were subsequently applied.

Eligibility Criteria and Study Selection

Studies were eligible for inclusion if they (1) were published in English between 2013 and 2025, (2) evaluated a mobile app primarily intended for menstrual cycle or fertility tracking, (3) included users of reproductive age, and (4) reported at least one of the following outcomes: app features or tracked parameters, predictive performance or agreement with a biological or clinical reference, user experience or satisfaction, or validation methodology.

Studies were excluded if they were not focused on mobile apps, lacked primary data, evaluated nondigital or desktop-only tools, or were published as conference abstracts, editorials, or other nonpeer-reviewed sources.

After clearly removing irrelevant and duplicate records, titles and abstracts were screened using a structured, relevance-focused approach. Structured prioritization methods were used to support screening efficiency and record organization; however, all inclusion and exclusion decisions were exclusively made by human reviewers. Full-text articles considered potentially relevant were subsequently assessed against the predefined eligibility criteria. Disagreements were resolved through discussion and consensus.

Exploratory searches initially identified a broad and heterogeneous literature base. Following structured screening and eligibility assessment, 68 full-text articles were reviewed in detail; 40 met the inclusion criteria and were retained for qualitative synthesis.

Data Extraction and Synthesis

Data were extracted using a structured template that included study design, population characteristics, app type, tracked physiological parameters, prediction approach, validation status, reported performance measures, and user-reported outcomes.

Structured extraction support tools were used only to facilitate the organization of study information. All extracted data were subsequently reviewed, verified, and corrected by human reviewers before synthesis. AI tools were not used to independently determine study inclusion, study quality, or interpretation of findings.

Because of substantial heterogeneity in study designs, measured parameters, validation approaches, and reported outcomes, a quantitative meta-analysis was deemed inappropriate. Instead, a narrative thematic synthesis was performed. Apps were broadly grouped into calendar-based and physiologically informed approaches, with additional consideration of single-parameter and multiparameter models.

The interpretation of reporting quality and methodological transparency was informed by selected domains derived from the COSMIN and ISPOR frameworks, including validity, reliability, usability, engagement, clinical relevance, and regulatory considerations.

Results

Characteristics of Included Studies

Exploratory literature searches identified a broad and heterogeneous body of publications related to mHealth apps for menstrual cycle and fertility tracking. Following structured screening and eligibility assessment, 40 studies met the predefined inclusion criteria and were included in the qualitative synthesis. Figure 1 illustrates the study identification, screening, eligibility assessment, and inclusion processes used in this structured narrative review.

The main characteristics of the included studies are summarized in Table 1, whereas the detailed study-level information is provided in Multimedia Appendix 1.

Figure 1. Flow diagram of study identification, screening, eligibility assessment, and inclusion.

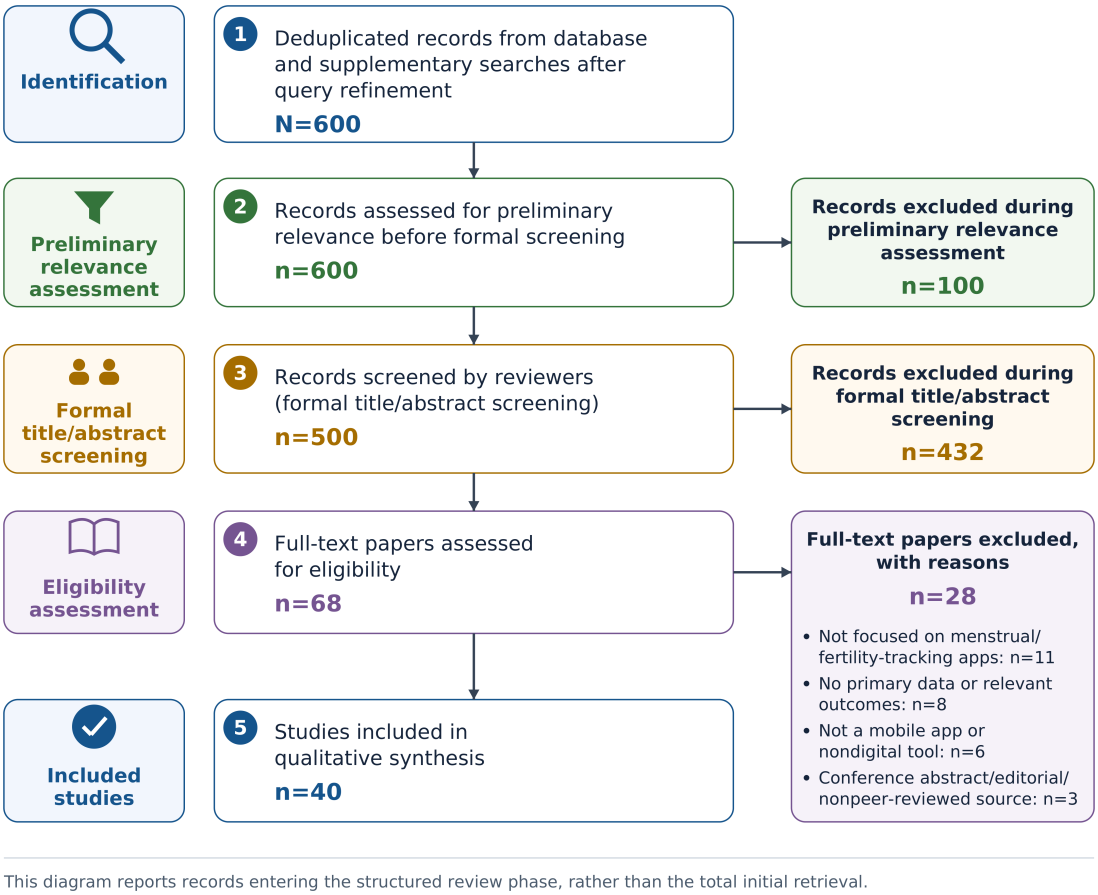


Table 1. Conceptual domains derived from the COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) and ISPOR (International Society for Health Economics and Outcomes Research) frameworks relevant to menstrual and fertility-tracking apps.

Framework	Domain	Application to menstrual/fertility apps
COSMIN	Validity (content and construct)	Are cycle-related symptoms or fertile window predictions based on valid constructs and user expectations?
COSMIN	Reliability (internal consistency and test-retest)	Do app-generated predictions or symptom scores remain stable and repeatable over successive cycles?

Framework	Domain	Application to menstrual/fertility apps
COSMIN	Responsiveness	Can the app detect changes in symptoms or physiological patterns across cycles or in response to interventions?
ISPOR	Clinical performance	Does the app improve clinical prediction accuracy, time-to-pregnancy estimates, or aid in decision-making?
ISPOR	Usability and engagement	Is the app intuitive and engaging? Does it support long-term use and provide personalized feedback?
ISPOR	Economic impact	Does the app reduce unnecessary clinical visits, enable self-care, or optimize the use of health resources?
ISPOR	Regulatory alignment	Is the app compliant with data protection laws (eg, GDPR[a] and HIPAA[b])? Does it meet the digital health standards?

[a]GDPR: General Data Protection Regulation [23].
[b]HIPAA: Health Insurance Portability and Accountability Act [24].

Among the 40 included studies, 30 (75%) used quantitative designs (retrospective, prospective, cohort, cross-sectional, or randomized controlled trial designs), 7 (17.5%) employed mixed methods approaches, and 2 (5%) were purely qualitative; one exploratory protocol was categorized as quantitative.

Most studies relied on menstrual cycle and symptom-tracking data, whereas a substantial subgroup incorporated physiological parameters such as BBT, LH monitoring, or wearable-derived metrics. Symptom tracking, including mood and premenstrual symptoms, was common, and several apps integrated wearable or biophysical parameters such as heart rate, heart rate variability (HRV), respiratory rate, sleep, activity, or skin temperature. Fewer studies included cervical mucus tracking, educational or algorithmic feedback features, or condition-specific modules, such as polycystic ovary syndrome (PCOS)–related tools.

The most commonly reported outcomes included fertile window or ovulation prediction, user experience and satisfaction, engagement, and symptom monitoring. Several studies also evaluated contraceptive effectiveness and fertility-related outcomes, such as time to pregnancy, menstrual health literacy, well-being, productivity, and cycle variability, in relation to factors including age, BMI, stress, and the COVID-19 pandemic. However, substantial heterogeneity was observed in reported outcomes, validation approaches, and performance metrics across the included literature.

Technical Features and Accuracy

Tracking Capabilities

Most apps track menstrual cycle dates. Many also incorporate physiological parameters, such as BBT, LH, cervical mucus, heart rate, and, increasingly, wearable-derived metrics (eg, skin temperature, steps, and HRV). Some apps (Ava, Natural Cycles, Flo, Dot, Cue, and Via) offer multiparameter tracking, whereas others primarily rely on calendar-based predictions.

Predictive Performance, Validation, and Reported Outcome Measures

Reported performance measures were highly heterogeneous and included predictive accuracy estimates, agreement measures, fertility-related outcomes, contraceptive effectiveness metrics, timing variability measures, physiological associations, and user-reported perceptions. Several included studies evaluated user-reported perceptions, usability, or health-related outcomes rather than direct predictive accuracy or biological validation. These outcomes should not be interpreted as directly comparable measures of clinical performance. Not all included studies reported predictive performance or validation outcomes; therefore, Table 2 summarizes only the subset of studies providing these data.

Table 2. Reported predictive performance, validation methods, and study limitations.

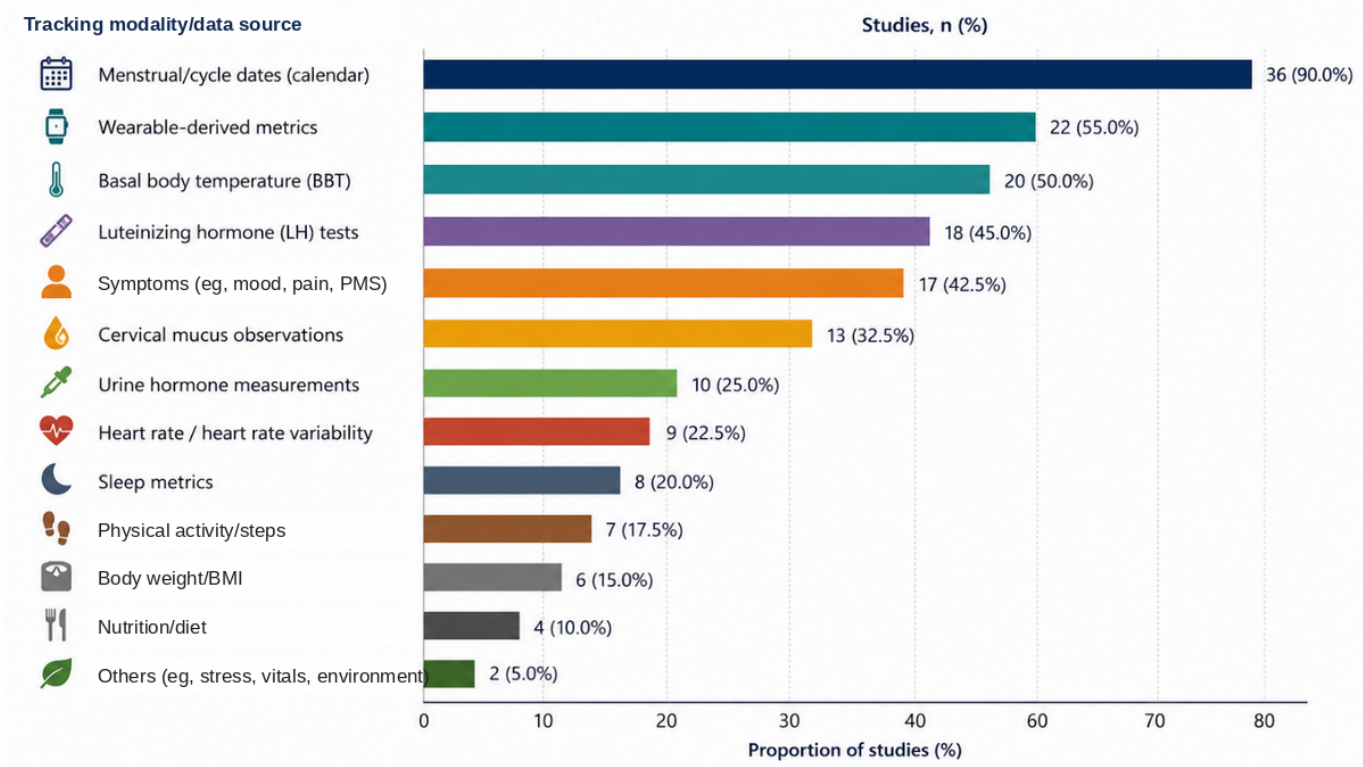
Study	Feature type	Performance metric type	Reported performance/ evaluation findings	Validation method	Limitations
Goodale et al [7] (2019)	Wrist skin temperature, heart rate, respiratory rate, heart rate variability, and skin perfusion	Predictive accuracy	90% accuracy (95% CI 0.89-0.92) for fertile window	Machine learning vs luteinizing hormone test	Based on compliant users; real-world adherence unknown
Stanford [25] (2019)	Cycle dates, luteinizing hormones, and symptoms	Fertility outcome association	Concordant with prior studies; day-specific conception probability up to 42%	App data vs clinical knowledge	Underreporting of intercourse; selection bias
Lang et al [26] (2024)	Heart rate, steps, and symptoms	Physiological association	Heart rate/steps varied by phase/ premenstrual syndrome	Wearable + questionnaire	Data completeness
Favaro et al [27] (2021)	Menstruation, basal body	Fertility outcomes	Median time to pregnancy 4 cycles;	App data, user-reported pregnancy	Dropout, underreporting of intercourse

Study	Feature type	Performance metric type	Reported performance/ evaluation findings	Validation method	Limitations
	temperature, and luteinizing hormone		6-cycle pregnancy probability 61%		
Mu and Fehring [28] (2023)	Luteinizing hormone (Premom/ Easy@Home), Clearblue Fertility Monitor	Agreement/ correlation	Correlation coefficient 0.99 with Clearblue Fertility Monitor for peak fertility	Comparison with Clearblue Fertility Monitor	Small, homogeneous sample
Mouriki et al [29] (2019)	Skin temperature, heart rate, and 5 others	Predictive accuracy	88.1% accuracy (wearable) vs 76.8% (calendar)	Luteinizing hormone test as reference	Small sample, regular cycles only
Berglund Scherwitzl et al [30] (2015)	Basal body temperature and ovulation test	Timing variability/ contraceptive reliability	1.9-day mean delay (basal body temperature vs luteinizing hormone); 0.05% false nonfertile days	App data vs ovulation test	Descriptive observational outcome only
Hills et al [31] (2023)	Luteinizing hormone and pregnanediol glucuronide	Phase prediction confidence	95% confidence for phase/day with age, cycle day, hormones	AI-powered urine test	No formal quantitative validation was reported
Manhart and Duane [32] (2022)	Basal body temperature and mucus	Fertile window agreement	12%-38% of cycles' fertile window within physiologic window	Comparison to Peak mucus	Small sample, regular cycles only
Wu et al [33] (2024)	Inertial measurement unit/gait	Prediction error	0.56 days average error for onset prediction	Time-series data, meta-learning	Small, short-term study
Pearson et al [34] (2021) (UK cohort)	Menstrual cycle data, basal body temperature, and optional luteinizing hormone tests	Contraceptive effectiveness	13-cycle cumulative pregnancy probability 7.1%; typical-use Pearl Index 6.1; perfect-use Pearl Index 2.0	Prospective real-world app cohort; Pearl Index and Kaplan-Meier life-table analysis	Self-selected app users; reliance on user-entered data; limited population diversity
Jennings et al [35] (2019)	Cycle dates	Contraceptive effectiveness	Perfect-use failure 1.0%, typical-use 5.8%	Life table, urine pregnancy test	Attrition, self-report
Sohda et al [36] (2017)	Cycle dates	Comparative predictive performance	Outperformed Ogino/mid-cycle methods	App data, linear modeling	No physiological metrics
Figueiredo et al [37] (2020)	Period, temperature, mucus, and ovulation	Timing variability	Fertile window start varied by 5 days; length 3-14 days	Simulated data, app review	No clinical validation
Bull et al [38] (2019)	Cycle, basal body temperature, and luteinizing hormone	Algorithmic prediction	Algorithmic ovulation/fertile window; no explicit accuracy	App data	No clinical validation
Symul et al [39] (2019)	Basal body temperature, mucus, and cervix	Predictive modeling	Hidden Markov Model for ovulation; 24% ovulate day 14-15	App data, modeling	Self-selection, incomplete cycles
Faust et al [40] (2019)	Cycle, ovulation, and symptoms	Fertility probability association	Highest conception probability day before ovulation	Bayesian models, app data	No explicit accuracy metrics were reported
Stanford et al [41] (2020)	Cycle, basal body temperature, luteinizing hormone, and mucus	Fertility outcome	12% to 20% increased fecundability with app use	Prospective cohort	No intercourse timing data
Liu et al [42] (2019)	Basal body temperature, heart rate, and symptoms	Predictive modeling outcome	Top 10% predicted 89% pregnancy in 6 cycles	Long short-term memory/logistic regression	Model interpretability, data completeness
Nulty et al [43] (2022)	Heart rate, sleep, and mood	Physiological agreement	Mean absolute percentage error 11.4% (heart rate), 8.5% (sleep) vs criterion	Polar/ActiGraph comparison	Small sample, pregnancy underpowered

Study	Feature type	Performance metric type	Reported performance/ evaluation findings	Validation method	Limitations
Song and Kanaoka [44] (2018)	Cycle dates	Health-related and productivity outcomes	No direct predictive accuracy assessment was reported	Randomized controlled trial, self-report	Self-selection bias and survey-based methodology
Grieger and Norman [45] (2020)	Cycle, flow, and ovulation	Algorithmic prediction outcomes	AI-based prediction; not quantified	App data	Subjective flow, incomplete BMI
Ponzo et al [46] (2022)	Symptoms, heart rate, and sleep	User-reported health impact	No explicit accuracy metrics were reported	Survey, wearable subset	Only premium users, no symptom-specific productivity
Lang et al [47] (2023)	Heart rate, steps, and symptoms	Physiological association	Heart rate/steps varied by premenstrual syndrome, phase	Wearable + questionnaire	Retention, small sample
Goodale and Blaivas [48] (2021)	5 biophysical parameters	User-reported usability/perceived value	No explicit accuracy metrics were reported	User survey	No accuracy data
Cunningham et al [49] (2024)	Cycle tracking	Health literacy and well-being outcomes	No explicit accuracy metrics were reported	Randomized controlled trial, self-report	No formal quantitative validation was reported
Zhaunova et al [50] (2023)	Cycle and symptoms	User-reported health and educational outcome	No explicit accuracy metrics were reported	Survey	Self-report, premium bias
Harper et al [51] (2022)	Cycle and ovulation	User-reported reliability	Not accurate; many periods earlier/ later than predicted	Survey	Selection bias, self-report
Broad et al [16] (2022)	Cycle and ovulation	User-reported reliability	Inaccurate for many users	Survey, thematic analysis	Social media recruitment
Gambier-Ross et al [52] (2018)	Cycle, mood, and symptoms	User-reported usability/perceived accuracy	Satisfactory for regular cycles; poor for irregular	Survey and interviews	Small, homogeneous sample
Grądzik et al [53] (2022)	Cycle, pregnancy, and parenting	User-reported usability and awareness outcomes	No explicit accuracy metrics were reported	Survey	Primarily usability-focused study
Haile et al [54] (2018)	Cycle dates	Contraceptive method implementation outcomes	Standard Days Method–based; not directly assessed	In-app survey	Self-report, language, no longitudinal data
Stujenske et al [55] (2023)	Urine hormone, app, and temperature	Fertility awareness and user satisfaction outcomes	No explicit accuracy metrics were reported	Survey	Fertility awareness focus
Arbeena et al [56] (2024)	Tracking, premenstrual syndrome, and polycystic ovary syndrome	User experience and symptom-management outcomes	No explicit accuracy metrics were reported	Interviews	Small, qualitative
Grenfell et al [57] (2021)	Basal body temperature and luteinizing hormone	User-reported reliability/perceived accuracy	Some users questioned basal body temperature accuracy	Interviews	Small, qualitative
Pearson et al [58] (2021) (US cohort)	Menstrual cycle data, basal body temperature, and optional luteinizing hormone tests	Contraceptive effectiveness outcome	13-cycle cumulative pregnancy probability 7.2%; typical-use Pearl Index 6.2; perfect-use Pearl Index 2.0	Prospective real-world US app cohort; Pearl Index and Kaplan-Meier life-table analysis	Self-selected annual subscribers; reliance on user-entered data; limited generalizability beyond current app users

Regarding input data and tracking modalities, most studies either used wearable or physiological sensor data (eg, heart rate, temperature, and steps) or relied solely on calendar-based cycle dates, with a smaller subset combining BBT, LH, urinary hormone tests, or cervical mucus observations. Several studies implemented multisymptom tracking that integrates menstrual dates, ovulation, and a range of symptoms, including mood and premenstrual complaints. The proportional distribution of the tracking modalities and data sources across the 40 studies is shown in Figure 2.

Figure 2. Tracking modalities and physiological data sources used across included studies (N=40). The figure presents the principal tracking modalities, physiological parameters, and data sources reported across studies evaluating menstrual cycle and fertility-tracking apps. Reported modalities include menstrual cycle dates, wearable-derived physiological metrics, basal body temperature, luteinizing hormone testing, symptom tracking, cervical mucus observations, sleep and activity monitoring, and additional biobehavioral parameters. Multiple modalities can be reported within a single study; therefore, percentages exceed 100%. Percentages represent the proportion of included studies reporting each modality or data source. PMS: premenstrual syndrome.



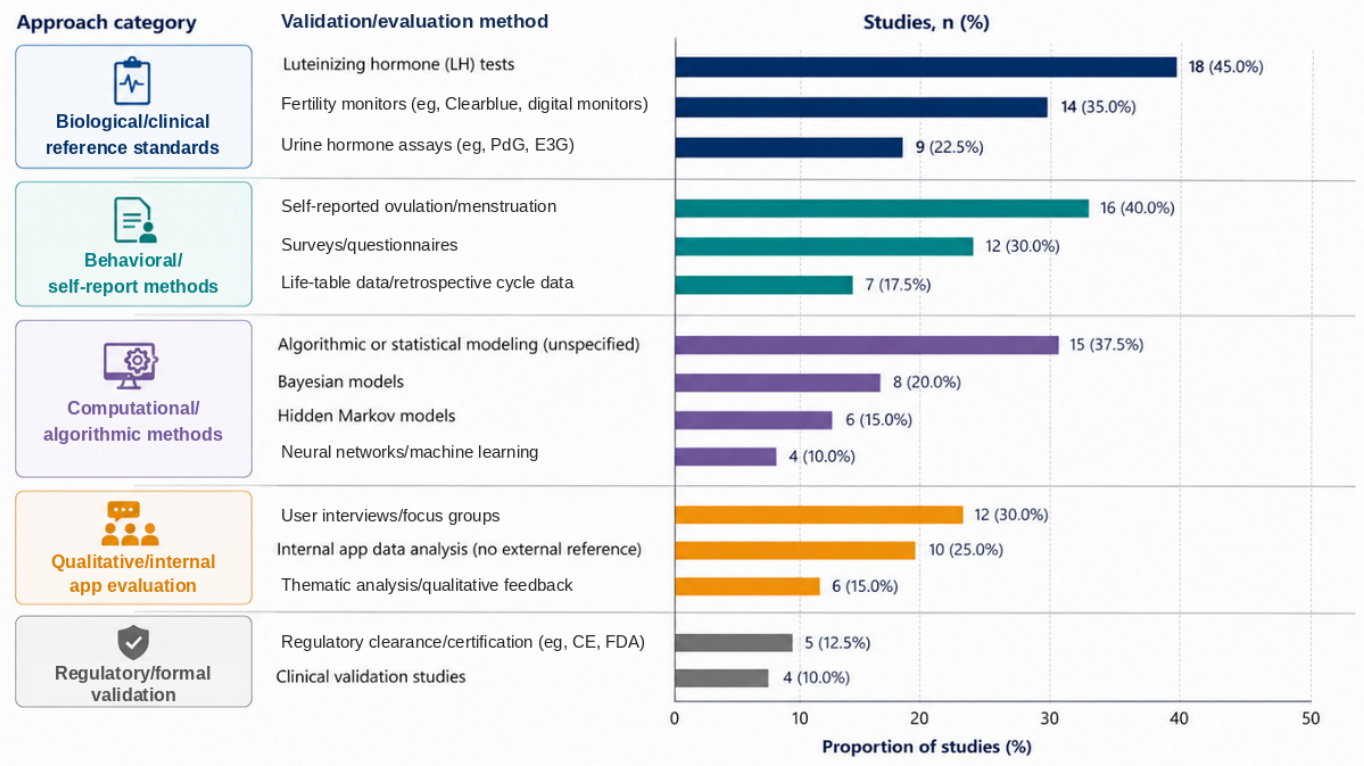
Only a minority of studies reported formal quantitative measures of predictive performance or biological validation. Five studies provided numerical estimates of prediction accuracy or error, 6 examined conception probability or time-to-pregnancy outcomes, and 3 quantified timing variability or prediction error related to fertile-window estimation. Individual studies additionally reported contraceptive failure rates under perfect- and typical-use conditions, correlations with physiological or clinical reference standards, overlap between predicted and biologically estimated fertile windows, and distributions of ovulation timing. Several reports also included qualitative observations about perceived inaccuracies, user dissatisfaction, uncertainty, or concerns regarding predictive reliability. In contrast, many included studies primarily focused on usability, educational value, behavioral engagement, health literacy, or implementation-related outcomes rather than the direct evaluation of predictive accuracy or physiological validation.

The validation approaches were highly heterogeneous. Some studies used clinical or laboratory references, such as LH tests, commercial fertility monitors, or urine ovulation tests. In contrast, others relied on self-reported outcomes (eg, pregnancies or life-table data), wearable sensor measurements combined with questionnaires, or purely algorithmic or model-based validation (including Bayesian, hidden Markov, or neural network approaches). Several reports drew only on internal app data or qualitative methods, such as surveys and interviews; in 12 studies, no explicit accuracy metrics or validation strategy could be identified. The distribution of validation methods and data sources is summarized in Figure 3.

A major trend identified in the reviewed literature is the transition from traditional calendar-based prediction models to increasingly complex, physiology-informed approaches. Earlier apps primarily relied on cycle averaging and retrospective menstrual date patterns, often assuming relatively stable ovulatory timing and regular cycle length. In contrast, more recent systems increasingly integrate multiparameter physiological data, including BBT, LH measurements, heart rate, HRV, sleep patterns, respiratory rate, wearable-derived metrics, and symptom-based monitoring. These developments reflect a broader shift from simple retrospective cycle estimation to dynamic, individualized tracking of reproductive physiology.

Although these newer approaches appear to improve prediction performance and biological plausibility, the available evidence suggests that substantial limitations remain. Menstrual physiology demonstrates considerable intraindividual and interindividual variability, influenced by stress, sleep, illness, physical activity, hormonal disorders, aging, medication use, and behavioral factors. Consequently, even advanced multiparameter models may provide an apparent level of precision that exceeds the true biological predictability of ovulation timing or menstrual phase transitions in real-world settings.

Figure 3. Validation and evaluation approaches reported across included studies (N=40). The figure summarizes the principal validation, assessment, and evaluation methods identified in the included studies investigating menstrual cycle and fertility-tracking apps. Approaches were grouped into biological or clinical reference standards, behavioral and self-report methods, computational or algorithmic models, qualitative or internal app evaluations, and formal regulatory or clinical validation pathways. Multiple approaches could be reported within a single study; therefore, percentages exceed 100%. Percentages represent the proportion of included studies reporting each approach category or method. CE: Conformité Européenne/European Conformity; FDA: US Food and Drug Administration.



These limitations may be particularly relevant for individuals with irregular menstrual cycles, PCOS, those in the perimenopausal transition, postpartum hormonal variability, or inconsistent adherence to physiological tracking. Importantly, many currently available apps were initially developed and validated predominantly in populations with relatively regular menstrual cycles, high user engagement, and relatively stable tracking behavior. This may limit generalizability and reduce predictive reliability in broader real-world populations characterized by more heterogeneous physiological and behavioral patterns.

The findings, therefore, suggest that physiology-informed and wearable-supported apps should be interpreted as tools that may improve reproductive awareness and cycle-related self-management, rather than as systems capable of eliminating the inherent biological uncertainty associated with menstrual and fertility prediction.

User Experience and Implementation

Use Patterns

The use of eHealth applications is widespread worldwide, with a user base ranging from hundreds of thousands to millions of users [4,38,59,60]. Motivation includes conception, contraception, health awareness, symptom management, and education. Use patterns vary by demographic factors (age, education, income, and geography), with higher uptake among younger, educated, and high-income users.

User Satisfaction

User satisfaction was most consistently associated with perceived accuracy, ease of use, personalized feedback, and improved awareness of menstrual or fertility patterns. Several studies reported that users valued apps for increasing symptom awareness, enhancing reproductive autonomy and preparedness, and improving communication about fertility or menstrual health [16,49,50,52].

At the same time, dissatisfaction was commonly linked to inaccurate cycle or ovulation predictions, particularly among users with irregular menstrual cycles, PCOS, stress-related variability, or changing hormonal patterns [16,52,56]. Some qualitative studies described frustration, confusion, anxiety, and reduced trust when predicted menstruation or fertile windows differed substantially from real-world experiences [16,37,57].

Several reports also highlighted concerns regarding limited algorithmic transparency, uncertainty communication, and overreliance on predictive outputs [37,57]. Although some users perceived these apps as highly reliable, others questioned whether simplified calendar-based predictions adequately reflected physiological complexity, especially in nonregular cycles [37,52,57].

Demographic and implementation factors, including education level, health literacy, socioeconomic status, digital engagement, and access to premium features, also influenced

user-experience findings [46,50]. Overall, user satisfaction appeared closely linked not only to predictive performance but also to transparency, personalization, and realistic communication of uncertainty.

User-Reported Experiences, Perceived Benefits, and Implementation Challenges

User-reported outcomes were heterogeneous and included perceived usefulness, satisfaction, reproductive autonomy, symptom awareness, educational value, emotional responses, and implementation-related concerns. Table 3 summarizes only studies reporting user-related outcomes, perceived benefits, or implementation findings.

Table 3. User-reported experiences, perceived benefits, and implementation challenges of menstrual and fertility-tracking apps.

Study	User-reported outcome domain	Reported user experience	Reported benefits and concerns	Study limitations
Goodale et al [7] (2019)	Real-time insights and education	Positive	Accurate, convenient, and educational real-time fertility insights	Adherence required; real-world compliance uncertain
Stanford [25] (2019)	Symptom tracking and fertility awareness	Positive	Large-scale symptom awareness data and conception-related insights	Underreporting of intercourse; potential selection bias
Mahalingaiah et al [59] (2021)	Tracking and retention	Mixed	Diverse cohort participation and long-term engagement assessment	Low response rate and participant dropout
Lang et al [26] (2024)	Symptom-physiology associations	Positive	Improved biological insights into the menstrual cycle and premenstrual symptom patterns	Data completeness issues and participant engagement variability
Favaro et al [27] (2021)	Reproductive autonomy and fertility awareness	Positive	Improved learning, fertility-related communication, and shared decision-making	Participant dropout and potential underreporting of intercourse
Mu and Fehring [28] (2023)	Ease of use and ovulation monitoring	Positive	Preference for the Clearblue Fertility Monitor and accurate ovulation detection	Small and relatively homogeneous study population
Nulty et al [43] (2022)	Acceptability of fertility and physiological tracking	Mixed/positive	Sleep and heart-rate tracking were generally well accepted	Pregnancy tracking was less accepted; small sample size and limited pregnancy-related data
Hills et al [31] (2023)	Cycle phase and fertility monitoring	Positive	Precise identification of menstrual cycle phase and fertile days using hormone-based assessment	Limited generalizability and absence of broader real-world validation data
Mouriki et al [29] (2019)	Fertile window prediction and wearable-based cycle tracking	Positive	Higher predictive performance compared with calendar-based approaches	Small sample size and inclusion primarily of users with regular menstrual cycles
Goodale and Blaivas [48] (2021)	Employer-supported fertility and health benefits	Positive	Easy-to-use app perceived as a valued employee health benefit	Limited clinical outcome data and lack of formal accuracy assessment
Lang et al [47] (2023)	Feasibility of combined physiological and symptom tracking	Positive	Successful integration of wearable and symptom-based data collection	Participant retention issues and study dropout
Berglund Scherwitzl et al [30] (2015)	Fertility awareness and ovulation tracking	Positive	Low false-positive rate and improved fertile window identification	Limited reporting of broader real-world accuracy and user diversity
Gambier-Ross et al [52] (2018)	Education, menstrual tracking, and symptom awareness	Highly positive	Improved symptom tracking, learning, and reproductive awareness	Privacy concerns and limited customization for diverse user needs
Symul et al [39] (2019)	User engagement and fertility awareness tracking	Positive	Frequent tracking and high user engagement with fertility-monitoring features	Self-selection bias and incomplete cycle reporting
Faust et al [40] (2019)	Symptom awareness and fertility monitoring	Positive	Increased awareness of symptom patterns and cervical changes associated with fertility	Limited clinical validation and incomplete physiological confirmation
Broad et al [16] (2022)	Preparedness, reproductive control, and cycle awareness	Mixed	Improved preparedness and menstrual awareness, but also frustration and anxiety related to inaccurate predictions	Social media recruitment and potential selection bias
Bull et al [38] (2019)	Ovulation and fertile window tracking	Not reported	Large user base and broad adoption of fertility-tracking functionality	No direct user-experience assessment or clinical validation data

Study	User-reported outcome domain	Reported user experience	Reported benefits and concerns	Study limitations
Figueiredo et al [37] (2020)	Trust, confusion, and perceived reliability of cycle predictions	Mixed	Some users perceived the app as trustworthy and useful, whereas others reported confusion related to prediction variability and limited transparency	Lack of algorithmic transparency and limited explanation of predictive uncertainty
Stujenske et al [55] (2023)	Satisfaction, fertility knowledge, and diagnostic support	Highly positive	Improved fertility knowledge, perceived diagnostic support, and increased user engagement	Limited generalizability due to a fertility awareness–focused population
Zhaunova et al [50] (2023)	Health knowledge, symptom awareness, and reproductive goal tracking	Positive	Improved health awareness, goal tracking, and engagement with menstrual health management	Self-reported data and potential premium-user selection bias
Simmons et al [61] (2017)	Feasibility of fertility awareness–based contraceptive use	Not reported	Feasibility and implementation insights into fertility-awareness apps	Protocol-based study without completed real-world outcome data
Stanford et al [41] (2020)	Fertility awareness and fecundability support	Positive	Increased fecundability and improved fertility awareness through app-supported cycle tracking	Lack of detailed intercourse timing data and potential behavioral confounding
Sohda et al [36] (2017)	Simplicity and usability of menstrual cycle tracking	Positive	Simple and accessible fertility tracking approach for general users	Lack of physiological measurements and limited biological validation
Ponzo et al [46] (2022)	Symptom management and menstrual health support	Positive	Improved preparedness, symptom support, and perceived emotional reassurance	Predominantly premium-user population and limited generalizability
Grieger and Norman [45] (2020)	Global reach and accessibility of menstrual-tracking apps	Positive	Broad international reach and increased accessibility of menstrual health–tracking tools	Subjective self-reported data and incomplete physiological validation
Grądzik et al [53] (2022)	Pregnancy awareness, enjoyment, and reproductive engagement	Positive	Increased pregnancy awareness and user enjoyment of app-based tracking features	Parenting-related functions were perceived as less useful or less frequently used
Wu et al [33] (2024)	Menstrual onset prediction and wearable-based cycle monitoring	Positive	Early prediction of menstrual onset using gait and wearable-derived data	Small sample size and short-term observational design
Liu et al [42] (2019)	Predictive stratification and personalized fertility prediction	Positive	Personalized prediction models enabled stratification of pregnancy likelihood and cycle-related outcomes.	Model interpretability limitations and incomplete real-world data consistency
Cunningham et al [49] (2024)	Health literacy, well-being, and reproductive communication	Positive	Improved sense of control, communication, and menstrual health awareness	Limited reporting of objective behavioral or long-term health outcomes
Arbeena et al [56] (2024)	Scheduling support, stigma reduction, and symptom understanding	Positive	Improved understanding of premenstrual symptoms and perceived support for cycle-related self-management	Small qualitative sample and limited generalizability
Jennings et al [35] (2019)	Contraceptive effectiveness and usability across demographic groups	Positive	Low contraceptive failure rates and comparable usability across age and racial groups	Attrition and reliance on self-reported data
Haile et al [54] (2018)	Satisfaction, recommendation, and accessibility of fertility-awareness apps	Positive	Improved accessibility and acceptability as a modern nonhormonal fertility management method	Self-reported data, language-related limitations, and lack of longitudinal follow-up
Grenfell et al [57] (2021)	Reproductive control, fertility knowledge, and relationship-related experiences	Mixed	Users appreciated the natural and nonintrusive approach to fertility-awareness tracking.	Small qualitative sample; reports of disrupted intimacy and increased gender-related responsibility burden
Nguyen et al [11] (2021)	Stress monitoring and menstrual cycle stability tracking	Positive	Reliable self-reporting of stress and perceived cycle stability patterns	Reliance on self-reported data and limited physiological confirmation
Song and Kanaoka [44] (2018)	Depression, dysmenorrhea, and menstrual health–related quality of life	Positive	Improved quality-adjusted life-year estimates and perceived cost-effectiveness of menstrual health support	Self-selection bias and survey-based methodology
Harper et al [51] (2022)	Frustration, anxiety, and perceived prediction reliability	Mixed/negative	Increased menstrual awareness despite concerns about prediction inaccuracies	Selection bias and reliance on self-reported user experiences
Manhart and Duane [32] (2022)	Fertility awareness–based fertile window estimation	Mixed/positive	Potential for improved fertile window prediction and nonhormonal fertility management	Small sample size and limited validation across diverse cycle patterns
Pearson et al	Contraceptive effectiveness and demographic characteristics of app users	Not reported	Demonstrated real-world contraceptive effectiveness of an FDA[a]-cleared fertility-awareness app in a US cohort	Self-selected annual subscribers; reliance on app-entered data; limited generalizability beyond current users

Study	User-reported outcome domain	Reported user experience	Reported benefits and concerns	Study limitations
[58] (2021) (US cohort)				
Pearson et al [34] (2021) (UK cohort)	Contraceptive effectiveness, user demographics, and continuation of app use	Not reported	Provided real-world evidence on contraceptive outcomes and continued use of a fertility-awareness app in a UK cohort	Data-entry burden and self-selected app users; reliance on user-entered data; substantial attrition over follow-up

[a]FDA: US Food and Drug Administration.

Across the included studies (Table 3), a clearly positive satisfaction level was reported in 10 studies, with 3 describing high satisfaction (eg, "high for fertility," "higher than for the Clearblue Fertility Monitor," or "98% useful"). Mixed or ambivalent satisfaction emerged in 4 studies, while 8 did not specify a satisfaction level despite mentioning user experience, and 12 did not report satisfaction at all.

Reported user-centered outcomes spanned multiple domains. Education, learning, or knowledge-related outcomes appeared in 9 out of 40 (22.5%) studies; tracking, engagement, retention, or acceptability-related outcomes in 7 (17.5%); and awareness, preparedness, control, or enjoyment in 9 (22.5%). Health, diagnostic, or well-being outcomes were reported in 4 (10%) studies; explicit satisfaction as a primary outcome in 3 (7.5%) studies; and symptom management in 1 (2.5%) study. Individual reports described unique outcomes, including real-time insights, employer benefits, trust or confusion, stress and cycle stability, frustration or anxiety, and no age or race differences (each in 1 [2.5%] study). In contrast, 12 (30%) studies did not provide user outcome data.

Education, learning, and knowledge benefits were reported in 10 (25%) studies, while accuracy, precision, or low false-positive rates were reported in 5 (12.5%). Awareness, preparedness, control, or pregnancy awareness benefits appeared in 8 (20%) studies; ease of use or accessibility in 3 (7.5%); and diagnostic support or well-being benefits in 5 (12.5%) studies. User bases or large-scale data were beneficial in 2 (5%) studies and symptom tracking or symptom-management benefits in 3 (7.5%) studies. Various unique benefits were reported in 1 (2.5%) study. Only 1 (2.5%) study did not describe benefits. Adherence, engagement, retention, dropout, attrition, and issues were reported as challenges in 13 (32.5%) studies. Underreporting, selection bias, self-selection, or "premium user" bias were mentioned in 8 (20%) studies, and data limitations, small samples, and restricted generalizability in 7 (17.5%) studies. Privacy or customization challenges were reported in 2 (5%) studies and inaccuracy in 1 (2.5%) study. User- and context-related issues were reported in a small number of studies. In contrast, 9 (22.5%) studies did not describe specific challenges.

Discussion

Principal Findings

This structured narrative review synthesizes evidence from 40 studies evaluating mobile apps for menstrual cycle and fertility tracking. The principal findings indicate that apps integrating physiological signals, such as BBT, LH measurements, and wearable-derived metrics, generally demonstrate superior predictive performance compared with calendar-based approaches [6,29,43]. At the same time, the review highlights persistent limitations, including limited clinical validation, heterogeneous evaluation methodologies, underrepresentation of users with irregular menstrual cycles, and insufficient transparency regarding algorithmic performance [1,13,37,60]. Together, these findings suggest that current apps are evolving toward more physiologically informed digital health solutions, although important methodological and clinical gaps remain.

Apps (Multimedia Appendix 2) that use physiological inputs, such as BBT, LH, and wearable signals, estimate ovulation or the fertile window more accurately than calendar-only methods. Multiparameter models achieved high agreement with clinical references, whereas simple averaging algorithms performed variably and were unreliable for irregular cycles [7,32,37]. Despite this growing algorithmic sophistication, formal clinical validation remains the exception rather than the rule [13,30,34,58]. Only a small number of fertility-tracking products, most prominently Natural Cycles and a few comparable regulated apps, have obtained Conformité Européenne (European Conformity) marks and/or US Food and Drug Administration clearance as medical devices based on clinical data [30,58,62]. At the same time, the vast majority rely on proprietary, nontransparent models that have not been tested in rigorous, peer-reviewed trials [13,37,60].

This review also highlighted several methodological challenges that limit the generalizability of the current evidence. Most studies were observational or retrospective and often lacked control groups or comparator tools, relying heavily on self-reported data and offering little or no objective validation [1,13,37,60]. Participant samples were predominantly drawn from high-income, technology-literate populations, providing limited insight into how these apps perform in underserved settings, among users with low digital literacy, or in groups with irregular menstrual cycles and reproductive health disorders [4,5,17,32,60,63].

Moreover, app performance has rarely been benchmarked against established clinical standards, such as LH surge kits, clinician-assessed ovulation timing, or structured fertility awareness protocols [13,27,28].

Evaluation studies should systematically apply established frameworks, such as the COSMIN [18] checklist and ISPOR [21] good practice guidelines, to strengthen the methodological rigor of future work. COSMIN offers structured criteria for assessing the reliability, validity, responsiveness, and interpretability of health measurement instruments, including app-generated symptom scores and digital biomarkers (eg, pain, mood, premenstrual syndrome severity, and cycle regularity). The ISPOR recommendations for digital health technologies emphasize the multidimensional appraisal of clinical performance, patient experience, economic value, and integration into health systems, providing a complementary lens for assessing menstrual and fertility apps [22,63]. The consistent use of these frameworks would standardize evaluation approaches, improve comparability across studies, and support more inclusive research that intentionally enrolls users with irregular cycles, reproductive health conditions, and those from low-resource settings [1,17,60].

Clinical Usefulness vs the Illusion of Precision

High predictive accuracy is frequently presented as a key performance indicator for menstrual and fertility-tracking apps; however, numerical accuracy alone does not necessarily translate into clinical usefulness [13,37,64]. Multiparameter apps integrating physiological signals such as BBT, LH measurements, HRV, sleep metrics, and wearable-derived data may achieve fertile-window prediction accuracies approaching 85% to 90% under optimal adherence conditions [6,29,43]. Nevertheless, real-world performance is influenced by incomplete data entry, inconsistent sensor use, behavioral variability, and the inherent biological variability of menstrual physiology [11,15,16,38,65].

As a result, many commercially available apps may convey a level of certainty that exceeds the biological predictability of ovulation timing or menstrual phase transitions [2,13,37]. This creates a potential "illusion of precision," whereby users may overestimate the reliability of algorithm-generated predictions when making reproductive or contraceptive decisions [16,66,67].

Consequently, menstrual and fertility-tracking apps should primarily be regarded as decision-support and self-management tools rather than autonomous diagnostic systems [32,68,69]. Future digital health solutions should therefore communicate uncertainty explicitly, for example, by presenting confidence ranges or prediction variability instead of single deterministic dates [65,66].

Irregular Cycles as a Stress Test for Algorithms

Across the reviewed studies, users with irregular menstrual cycles, such as those with PCOS, perimenopausal changes, or stress-related variability, were markedly underrepresented. Although menstrual irregularities are common among women of reproductive age, most evaluated algorithms were developed and validated primarily using datasets representing relatively regular ovulatory cycles [13,32,34,58].

This selection bias limits the generalizability of reported performance metrics, which are likely to be most representative of users with relatively regular menstrual patterns rather than populations with greater physiological variability [32,37,60].

Addressing this gap will require the purposeful inclusion of irregular-cycle datasets and adaptive modeling approaches that learn from intraindividual variability rather than primarily relying on population averages [6,11,70].

Lifestyle Trackers vs Regulated Devices

The findings highlight the ambiguity between lifestyle-tracking apps and regulated digital medical devices [13,37,60]. Notably, few products, such as Natural Cycles, have clearance, despite marketing suggesting clinical reliability [30,34,58].

Apps positioned as medical devices should be subject to transparent validation standards, postmarket surveillance, and explicit risk communication strategies [66,69]. In contrast, lifestyle tools may prioritize usability but must avoid presenting outputs as equivalent to medical advice [16,17,65]. Therefore, classification and labeling are critical for regulatory compliance and informed decisions by users and clinicians [66,71].

AI-Driven Cycle Modeling: Opportunity and Responsibility

AI and machine learning transform menstrual and fertility tracking into multiparameter models integrating BBT, hormone assays, heart rate–derived metrics, and wearable data [6,7,70]. These systems increasingly resemble adaptive learning health technologies with the potential to uncover subtle patterns in cycle physiology and support individualized predictions [6,64,70]. Simultaneously, the increasing complexity of models raises concerns regarding transparency, explainability, and trust, particularly when black-box approaches obscure assumptions and failure modes [66,71]. For apps in reproductive health, explainable AI and user-facing interpretability should be treated as core design requirements rather than optional enhancements, ensuring that users and clinicians can understand the basis and limitations of the predictions [66,70,71].

Integration Into Clinical and Telemedical Workflows

Finally, the value of menstrual and fertility apps extends beyond self-tracking when data are integrated into clinical and telemedicine workflows. Structured, longitudinal app-generated information could support shared decision-making, symptom trending, and patient-reported outcome monitoring, especially if aligned with established patient-reported outcome measures and digital health evaluation frameworks. Realizing that this potential will require

standardized data structures, interoperability with electronic health records, and clear definitions of clinical responsibility for interpreting and acting on app data. Without such infrastructure, the translation of app-based insights into meaningful health care actions is likely to remain fragmented and opportunistic rather than systematic [17,59,72].

Limitations

This review has several limitations that should be considered when interpreting its findings. Although a structured reporting framework, informed by selected PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) principles [73], was used to guide study identification, screening, and inclusion, the review was not designed as a fully PRISMA-compliant systematic review. The exploratory search strategy intentionally prioritized sensitivity and broad coverage, yielding a large, heterogeneous corpus of records that required structured prioritization and relevance-focused refinement before formal screening. Supportive prioritization tools, including AI-assisted relevance ranking at the title and abstract levels, were used only to improve screening feasibility and organization; however, all inclusion, exclusion, and data verification decisions were made exclusively by human reviewers. Consequently, the study selection workflow reflected the records that entered the structured review phase rather than the total number of records initially retrieved. Furthermore, the use of supportive prioritization methods may theoretically have contributed to the omission of some relevant studies, particularly those using unconventional terminology or less common indexing terms.

Second, substantial heterogeneity among the included studies precluded quantitative synthesis. The reviewed literature varied widely in study design, population characteristics, physiological parameters tracked, outcome definitions, validation methods, and performance metrics. Therefore, accuracy estimates and predictive performance indicators were reported descriptively and could not be interpreted as pooled or directly comparable measures. As a result, the reported ranges of prediction accuracy should be understood as study-specific rather than generalizable benchmarks.

Third, the evidence base was dominated by observational and retrospective studies, with frequent reliance on self-reported data and limited use of standardized clinical reference standards. Only a minority of apps have undergone formal clinical validation or regulatory assessment, and many studies lack transparency about the underlying algorithms or data-processing methods. These factors limit the ability to draw firm conclusions regarding clinical-grade reliability.

Finally, participant samples in many studies were predominantly drawn from high-income, technologically literate populations, with underrepresentation of users with irregular menstrual cycles, reproductive health disorders (such as PCOS and perimenopausal changes), and populations from low-resource settings. This limits the external validity of the reported performance and user-experience outcomes and may contribute to algorithmic bias favoring regular cycle patterns.

Despite these limitations, the structured and transparent synthesis presented in this review provides a critical overview of the current evidence, highlights consistent patterns across heterogeneous studies, and identifies key methodological and clinical gaps that should inform future research and evaluation of menstrual and fertility-tracking apps.

These limitations should be considered when interpreting the conclusions, which emphasize relative rather than absolute performance differences between the app types.

Conclusions

This structured narrative review demonstrates that menstrual and fertility-tracking apps that integrate physiological signals, such as BBT, LH measurements, and wearable-derived metrics, generally provide more robust predictions of ovulation and fertile windows than calendar-based approaches. However, the review also shows that formal clinical validation, transparent reporting of algorithmic performance, and regulatory evaluation remain limited in the currently available evidence.

The findings further indicate that predictive performance should be interpreted in the context of substantial biological variability, particularly among individuals with irregular menstrual cycles, PCOS, perimenopausal transition, or inconsistent tracking behavior. Consequently, these apps are currently best regarded as decision-support and self-management technologies that enhance reproductive awareness rather than as autonomous diagnostic systems.

This review also identifies important methodological gaps that should guide future research, including the need for standardized evaluation frameworks, transparent reporting of analytical methods, inclusion of more diverse user populations, and rigorous prospective validation. Addressing these priorities will improve the clinical relevance, interpretability, and equitable implementation of digital reproductive health technologies while supporting their future integration into routine health care and telemedicine.

Acknowledgments

The authors used ChatGPT (OpenAI) during manuscript preparation for language editing, structural refinement, and assistance with figure preparation and formatting. AI-assisted tools were also used to support relevance prioritization during title and abstract screening and the organization of extracted information, as described in the Methods section. All eligibility decisions, data verification, interpretation of findings, and final manuscript content were reviewed and approved by the authors. The authors take full responsibility for the content of the manuscript.

Funding

The study was self-funded.

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

Both authors contributed equally to this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Characteristics of included studies.
[DOCX File (Microsoft Word File), 61 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Comparison of selected menstrual and fertility-tracking apps.
[DOCX File (Microsoft Word File), 16 KB-Multimedia Appendix 2]

References

1. Kalampalikis A, Chatziioannou SS, Protopapas A, Gerakini AM, Michala L. mHealth and its application in menstrual related issues: a systematic review. Eur J Contracept Reprod Health Care. Feb 2022;27(1):53-60. [doi: 10.1080/13625187.2021.1980873] [Medline: 34615425]
2. Earle S, Marston HR, Hadley R, Banks D. Use of menstruation and fertility app trackers: a scoping review of the evidence. BMJ Sex Reprod Health. Apr 2021;47(2):90-101. [doi: 10.1136/bmjsrh-2019-200488] [Medline: 32253280]
3. Kazakoff A, Doroshuk ML, Ganshorn H, Doyle-Baker PK. Motivations for use, user experience and quality of reproductive health mobile applications in a pre-menopausal user base: a scoping review. Health Care (Don Mills). 2025;13(8):877. [doi: 10.3390/healthcare13080877]
4. Rampazzo F, Raybould A, Rampazzo P, Barker R, Leasure D. "UPDATE: I'm pregnant!": inferring global downloads and reasons for using menstrual tracking apps. Digit Health. 2024;10:20552076241298315. [doi: 10.1177/20552076241298315] [Medline: 39582950]
5. Hong M, Rajaguru V, Kim K, Jang SY, Lee SG. Menstrual cycle management and period tracker app use in millennial and generation Z individuals: mixed methods study. J Med Internet Res. Oct 10, 2024;26:e53146. [doi: 10.2196/53146] [Medline: 39388687]
6. Lyzwinski L, Elgendi M, Menon C. Innovative approaches to menstruation and fertility tracking using wearable reproductive health technology: systematic review. J Med Internet Res. Feb 15, 2024;26:e45139. [doi: 10.2196/45139] [Medline: 38358798]
7. Goodale BM, Shilaih M, Falco L, Dammeier F, Hamvas G, Leeners B. Wearable sensors reveal menses-driven changes in physiology and enable prediction of the fertile window: observational study. J Med Internet Res. Apr 18, 2019;21(4):e13404. [doi: 10.2196/13404] [Medline: 30998226]
8. Menstrual health apps market expected to reach USD 7.52 billion by 2032, driven by smartphone adoption & personalized digital health solutions–SNS Insider. Yahoo Finance. 2025. URL: https://finance.yahoo.com/news/menstrual-health-apps-market-expected-100400856.html? [Accessed 2026-09-18]
9. Flo Health Inc. Ovulation & period tracker Flo. Google Play. 2026. URL: https://play.google.com/store/apps/details/Flo_Ovulation_Period_Tracker?id=org.iggymedia.periodtracker&hl=en_IN [Accessed 2026-09-18]
10. Clue: period & cycle tracker. Google Play. 2026. URL: https://play.google.com/store/apps/details/Clue_Cycle_Period_Tracker?id=com.clue.android&hl=uz [Accessed 2026-09-18]
11. Nguyen BT, Pang RD, Nelson AL, et al. Detecting variations in ovulation and menstruation during the COVID-19 pandemic, using real-world mobile app data. PLOS ONE. 2021;16(10):e0258314. [doi: 10.1371/journal.pone.0258314] [Medline: 34669726]
12. VanBenschoten H, Kuganantham H, Larsson EC, et al. Impact of the COVID-19 pandemic on access to and utilisation of services for sexual and reproductive health: a scoping review. BMJ Glob Health. Oct 2022;7(10):e009594. [doi: 10.1136/bmjgh-2022-009594] [Medline: 36202429]
13. Saugar EE, Katsoulos S, Kim HS, et al. Factors used by mobile applications to predict female fertility status and their reported effectiveness: a scoping review. Cureus. Nov 2023;15(11):e48847. [doi: 10.7759/cureus.48847] [Medline: 38106802]

14. Ramaiyer M, El Sabeh M, Zhu J, et al. The association of COVID-19 vaccination and menstrual health: a period-tracking app-based cohort study. Vaccine X. Aug 2024;19:100501. [doi: 10.1016/j.jvacx.2024.100501] [Medline: 38832342]
15. Karasneh RA, Al-Azzam SI, Alzoubi KH, Muflih SM, Hawamdeh SS. Smartphone applications for period tracking: rating and behavioral change among women users. Obstet Gynecol Int. 2020;2020:2192387. [doi: 10.1155/2020/2192387] [Medline: 32952563]
16. Broad A, Biswakarma R, Harper JC. A survey of women's experiences of using period tracker applications: attitudes, ovulation prediction and how the accuracy of the app in predicting period start dates affects their feelings and behaviours. Womens Health (Lond). 2022;18:17455057221095246. [doi: 10.1177/17455057221095246] [Medline: 35465788]
17. Sheridan Clay K, Ford T, Mackay R, Keating S, Ziebland S, Powell J. Implications of digital fertility tracking for clinical care: a qualitative systematic review. Reprod Health. Oct 6, 2025;22(1):183. [doi: 10.1186/s12978-025-02083-1] [Medline: 41053790]
18. Mokkink LB, Prinsen CAC, Bouter LM, Vet H de, Terwee CB. The Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) and how to select an outcome measurement instrument. Braz J Phys Ther. Jan 19, 2016;20(2):105-113. [doi: 10.1590/bjpt-rbf.2014.0143] [Medline: 26786084]
19. Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN). URL: https://www.cosmin.nl/ [Accessed 2026-09-18]
20. Muro-Culebras A, Escriche-Escuder A, Martin-Martin J, et al. Tools for evaluating the content, efficacy, and usability of mobile health apps according to the consensus-based standards for the selection of health measurement instruments: systematic review. JMIR Mhealth Uhealth. Dec 1, 2021;9(12):e15433. [doi: 10.2196/15433] [Medline: 34855618]
21. The Professional Society for Health Economics and Outcomes Research (ISPOR). URL: https://www.ispor.org/ [Accessed 2026-09-18]
22. Burrell A, Zrubka Z, Champion A, et al. How useful are digital health terms for outcomes research? An ISPOR special interest group report. Value Health. Sep 2022;25(9):1469-1479. [doi: 10.1016/j.jval.2022.04.1730] [Medline: 36049797]
23. General Data Protection Regulation. URL: https://gdpr-info.eu/ [Accessed 2026-09-18]
24. HIPAA and administrative simplification. Centers for Medicare & Medicaid Services (CMS). URL: https://www.cms.gov/priorities/key-initiatives/burden-reduction/administrative-simplification/hipaa [Accessed 2026-10-14]
25. Stanford JB. Big data meets the menstrual cycle. Fertil Steril. Sep 2019;112(3):464-465. [doi: 10.1016/j.fertnstert.2019.05.035] [Medline: 31446905]
26. Lang AL, Bruhn RL, Fehling M, et al. Feasibility study on menstrual cycles with Fitbit device (FEMFIT): prospective observational cohort study. JMIR Mhealth Uhealth. Mar 12, 2024;12:e50135. [doi: 10.2196/50135] [Medline: 38470472]
27. Favaro C, Pearson JT, Rowland SP, et al. Time to pregnancy for women using a fertility awareness based mobile application to plan a pregnancy. J Womens Health (Larchmt). Nov 2021;30(11):1538-1545. [doi: 10.1089/jwh.2021.0026] [Medline: 34495761]
28. Mu Q, Fehring RJ. A comparison of two hormonal fertility monitoring systems for ovulation detection: a pilot study. Medicina (Kaunas). Feb 18, 2023;59(2):400. [doi: 10.3390/medicina59020400] [Medline: 36837601]
29. Mouriki E, Bilic A, Goodale BM, et al. Detection of the fertile window using a wearable medical device and the calendar method: a comparative study. Fertil Steril. Sep 2019;112(3):e268. [doi: 10.1016/j.fertnstert.2019.08.035]
30. Berglund Scherwitzl E, Lindén Hirschberg A, Scherwitzl R. Identification and prediction of the fertile window using NaturalCycles. Eur J Contracept Reprod Health Care. 2015;20(5):403-408. [doi: 10.3109/13625187.2014.988210] [Medline: 25592280]
31. Hills E, Woodland MB, Divaraniya A. Using hormone data and age to pinpoint cycle day within the menstrual cycle. Medicina (Kaunas). Jul 23, 2023;59(7):1348. [doi: 10.3390/medicina59071348] [Medline: 37512159]
32. Manhart MD, Duane M. A comparison of app-defined fertile days from two fertility tracking apps using identical cycle data. Contraception. Nov 2022;115:12-16. [doi: 10.1016/j.contraception.2022.07.007] [Medline: 35901971]
33. Wu Y, Zhang J, Chen Y, Shi W, Tan H. MC-tracking: towards ubiquitous menstrual cycle tracking using the smartphone. IEEE Trans Mobile Comput. 2024;23(5):5604-5615. [doi: 10.1109/TMC.2023.3310941]
34. Pearson JT, Chelstowska M, Rowland SP, et al. Natural Cycles app: contraceptive outcomes and demographic analysis of UK users. Eur J Contracept Reprod Health Care. Apr 2021;26(2):105-110. [doi: 10.1080/13625187.2020.1867844] [Medline: 33539252]
35. Jennings V, Haile LT, Simmons RG, Spieler J, Shattuck D. Perfect- and typical-use effectiveness of the Dot fertility app over 13 cycles: results from a prospective contraceptive effectiveness trial. Eur J Contracept Reprod Health Care. Apr 2019;24(2):148-153. [doi: 10.1080/13625187.2019.1581164] [Medline: 30880509]

36. Sohda S, Suzuki K, Igari I. Relationship between the menstrual cycle and timing of ovulation revealed by new protocols: analysis of data from a self-tracking health app. J Med Internet Res. Nov 27, 2017;19(11):e391. [doi: 10.2196/jmir.7468] [Medline: 29180346]
37. Figueiredo MC, Su HI, Chen Y. An evaluation of commercial fertility apps: algorithmic predictions and users' perceptions. Fertil Steril. Sep 2020;114(3):e552-e553. [doi: 10.1016/j.fertnstert.2020.09.091]
38. Bull JR, Rowland SP, Scherwitzl EB, Scherwitzl R, Danielsson KG, Harper J. Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. NPJ Digit Med. 2019;2(1):83. [doi: 10.1038/s41746-019-0152-7] [Medline: 31482137]
39. Symul L, Wac K, Hillard P, Salathé M. Assessment of menstrual health status and evolution through mobile apps for fertility awareness. NPJ Digit Med. 2019;2(1):64. [doi: 10.1038/s41746-019-0139-4] [Medline: 31341953]
40. Faust L, Bradley D, Landau E, et al. Findings from a mobile application-based cohort are consistent with established knowledge of the menstrual cycle, fertile window, and conception. Fertil Steril. Sep 2019;112(3):450-457.e3. [doi: 10.1016/j.fertnstert.2019.05.008] [Medline: 31272722]
41. Stanford JB, Willis SK, Hatch EE, Rothman KJ, Wise LA. Fecundability in relation to use of mobile computing apps to track the menstrual cycle. Hum Reprod. Oct 1, 2020;35(10):2245-2252. [doi: 10.1093/humrep/deaa176] [Medline: 32910202]
42. Liu B, Shi S, Wu Y. Predicting pregnancy using large-scale data from a women's health tracking mobile application. Presented at: WWW '19: Proceedings of the 2019 World Wide Web Conference; May 13-17, 2019; San Francisco, CA, USA. [doi: 10.1145/3308558.3313512]
43. Nulty AK, Chen E, Thompson AL. The Ava bracelet for collection of fertility and pregnancy data in free-living conditions: an exploratory validity and acceptability study. Digit Health. 2022;8:20552076221084461. [doi: 10.1177/20552076221084461] [Medline: 35295766]
44. Song M, Kanaoka H. Effectiveness of mobile application for menstrual management of working women in Japan: randomized controlled trial and medical economic evaluation. J Med Econ. Nov 2018;21(11):1131-1138. [doi: 10.1080/13696998.2018.1515082] [Medline: 30130990]
45. Grieger JA, Norman RJ. Menstrual cycle length and patterns in a global cohort of women using a mobile phone app: retrospective cohort study. J Med Internet Res. Jun 24, 2020;22(6):e17109. [doi: 10.2196/17109] [Medline: 32442161]
46. Ponzo S, Wickham A, Bamford R, et al. Menstrual cycle-associated symptoms and workplace productivity in US employees: a cross-sectional survey of users of the Flo mobile phone app. Digit Health. 2022;8:20552076221145852. [doi: 10.1177/20552076221145852] [Medline: 36544535]
47. Lang A, Bruhn R, Fehling M, et al. FeMFit: feasibility of tracking the menstrual cycle using digital questionnaires and wearable devices. Eur J Public Health. Oct 24, 2023;33(Supplement_2):ckad160.1211. [doi: 10.1093/eurpub/ckad160.1211]
48. Goodale BM, Blaivas M. wReal-world perspective on wearable sensor technology as a complement to traditional fertility services and an employer-covered benefit. Fertil Steril. Sep 2021;116(3):e268. [doi: 10.1016/j.fertnstert.2021.07.718]
49. Cunningham AC, Prentice C, Peven K, et al. Efficacy of the Flo app in improving health literacy, menstrual and general health, and well-being in women: pilot randomized controlled trial. JMIR Mhealth Uhealth. May 2, 2024;12:e54124. [doi: 10.2196/54124] [Medline: 38696773]
50. Zhaunova L, Bamford R, Radovic T, et al. Characterization of self-reported improvements in knowledge and health among users of Flo period tracking app: cross-sectional survey. JMIR Mhealth Uhealth. Apr 26, 2023;11:e40427. [doi: 10.2196/40427] [Medline: 37099370]
51. Harper J, Broad A, Biswakarma R. P-494 A survey of women's experiences, behaviours and attitudes of using period tracker applications (apps). Hum Reprod. Jun 29, 2022;37(Supplement_1):deac107.460. [doi: 10.1093/humrep/deac107.460]
52. Gambier-Ross K, McLernon DJ, Morgan HM. A mixed methods exploratory study of women's relationships with and uses of fertility tracking apps. Digit Health. 2018;4:2055207618785077. [doi: 10.1177/2055207618785077] [Medline: 30225095]
53. Grądzik A, Gorajek A, Osiejewska A, Wojtachnio D, Nowakowska I, Bartoszewicz J. The popularity and usability of mobile applications for women and expecting mothers. J Educ Health Sport. 2022;12(8):378-382. [doi: 10.12775/JEHS.2022.12.08.039]
54. Haile LT, Fultz HM, Simmons RG, Shelus V. Market-testing a smartphone application for family planning: assessing potential of the CycleBeads app in seven countries through digital monitoring. Mhealth. 2018;4:27. [doi: 10.21037/mhealth.2018.06.07] [Medline: 30148140]
55. Stujenske TM, Mu Q, Pérez Capotosto M, Bouchard TP. Survey analysis of quantitative and qualitative menstrual cycle tracking technologies. Medicina (Kaunas). Aug 22, 2023;59(9):1509. [doi: 10.3390/medicina59091509] [Medline: 37763628]

56. Arbeena S, Asif M, Wani MI. Menstrual health in the digital age: a qualitative study of motivations for using period-tracking apps by women in Kashmir. Women's Reprod Health. 2025;12:81-99. [doi: 10.1080/23293691.2024.2381036]
57. Grenfell P, Tilouche N, Shawe J, French RS. Fertility and digital technology: narratives of using smartphone app 'Natural Cycles' while trying to conceive. Sociol Health Illn. Jan 2021;43(1):116-132. [doi: 10.1111/1467-9566.13199] [Medline: 33147647]
58. Pearson JT, Chelstowska M, Rowland SP, et al. Contraceptive effectiveness of an FDA-cleared birth control app: results from the Natural Cycles US cohort. J Womens Health (Larchmt). Jun 2021;30(6):782-788. [doi: 10.1089/jwh.2020.8547] [Medline: 33370220]
59. Mahalingaiah S, Fruh V, Rodriguez E, et al. Design and methods of the Apple Women's Health Study: a digital longitudinal cohort study. Am J Obstet Gynecol. Apr 2022;226(4):545. [doi: 10.1016/j.ajog.2021.09.041] [Medline: 34610322]
60. Schantz JS, Fernandez CSP, Anne Marie ZJ. Menstrual cycle tracking applications and the potential for epidemiological research: a comprehensive review of the literature. Curr Epidemiol Rep. Mar 2021;8(1):9-19. [doi: 10.1007/s40471-020-00260-3] [Medline: 34055569]
61. Simmons RG, Shattuck DC, Jennings VH. Assessing the efficacy of an app-based method of family planning: the dot study protocol. JMIR Res Protoc. Jan 18, 2017;6(1):e5. [doi: 10.2196/resprot.6886] [Medline: 28100441]
62. Lovett L. Natural Cycles snags FDA clearance to use wearable temperature data in its fertility algorithm. MobiHealthNews. 2021. URL: https://www.mobihealthnews.com/news/natural-cycles-snags-fda-clearance-use-wearable-temperature-data-its-fertility-algorithm [Accessed 2026-09-18]
63. Karanikolos M, Adib K, Azzopardi Muscat N, et al. Assessing health system performance: proof of concept for a HSPA dashboard of key indicators. OECD Publishing; 2023. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/12/assessing-health-system-performance_d6998b2c/4e6b28c0-en.pdf [Accessed 2026-09-18]
64. Milad HC, Cromack SC, Walter JR. Menstrual tracking technologies and fertility: evaluating accuracy, utility, and impact on time to pregnancy. Fertil Steril. Jun 2026;125(6):983-992. [doi: 10.1016/j.fertnstert.2026.02.027] [Medline: 41747954]
65. Patel U, Broad A, Biswakarma R, Harper JC. Experiences of users of period tracking apps: which app, frequency of use, data input and output and attitudes. Reprod Biomed Online. Mar 2024;48(3):103599. [doi: 10.1016/j.rbmo.2023.103599] [Medline: 38295553]
66. Punzi MC, Thuis T. Mapping ethical concerns in algorithm-driven period and fertility tracking technologies. Contraception. May 2026;157:110837. [doi: 10.1016/j.contraception.2025.110837] [Medline: 39923910]
67. Levy J, Romo-Avilés N. "A good little tool to get to know yourself a bit better": a qualitative study on users' experiences of app-supported menstrual tracking in Europe. BMC Public Health. Sep 3, 2019;19(1):1213. [doi: 10.1186/s12889-019-7549-8] [Medline: 31481043]
68. Duane M, Stanford JB, Porucznik CA, Vigil P. Fertility awareness-based methods for women's health and family planning. Front Med (Lausanne). 2022;9:858977. [doi: 10.3389/fmed.2022.858977] [Medline: 35685421]
69. FSRH clinical guideline: fertility awareness methods. Faculty of Sexual & Reproductive Healthcare; 2015. URL: https://www.cosrh.org/Common/Uploaded%20files/documents/ceuguidancefertilityawarenessmethods.pdf [Accessed 2026-09-18]
70. Moghimikandelousi S, Najm L, Lee Y, et al. Advances in biomonitoring technologies for women's health. Nat Commun. Sep 26, 2025;16(1):8507. [doi: 10.1038/s41467-025-63501-3] [Medline: 41006271]
71. Anokwuru CC, Nwuzoh MI, Eneh S, et al. Regulating algorithmic tools in reproductive health: ethical and legal challenges. Front Reprod Health. 2026;8:1771550. [doi: 10.3389/frph.2026.1771550] [Medline: 42179720]
72. Adnan T, Coull BA, Jukic AM, Mahalingaiah S. The real-world applications of the symptom tracking functionality available to menstrual health tracking apps. Curr Opin Endocrinol Diabetes Obes. Dec 1, 2021;28(6):574-586. [doi: 10.1097/MED.0000000000000682] [Medline: 34560714]
73. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). URL: https://www.prisma-statement.org/? [Accessed 2026-09-18]

Abbreviations

BBT : basal body temperature
COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
HRV : heart rate variability
ISPOR: International Society for Health Economics and Outcomes Research
LH : luteinizing hormone
mHealth: mobile health
PCOS: polycystic ovary syndrome

PRISMA : Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by Nazlena Mohamad Ali; peer-reviewed by Achmad Hamdan, Ekaterina Andreevna Narodova, Tirumala Ashish Kumar Manne; submitted 05.Feb.2026; initial comments to author 30.Mar.2026; final revised version received 26.Jul.2026; accepted 10.Aug.2026; published 19.Nov.2026

Please cite as:
Cedro T, Glinkowski W
Mobile Apps for Women's Physiology Tracking: Structured Narrative Review
JMIR Mhealth Uhealth 2026;14:e92968
URL: https://mhealth.jmir.org/2026/1/e92968
doi: 10.2196/92968

© Tomasz Cedro, Wojciech Glinkowski. Originally published in JMIR mHealth and uHealth (https://mhealth.jmir.org), 19.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on https://mhealth.jmir.org/, as well as this copyright and license information must be included.