

AI and Wearable Technologies in High-Risk Pregnancy Monitoring: A Review of the Gap between Technological Innovation and Clinical Application

Mehri Layegh

Affiliation: Counseling and Reproductive Health Research Center, Golestan University of
Medical Sciences, Gorgan, Iran.

Email: MehriLayegh882@gmail.com

Milad Panahandeh*

Affiliation: Clinical Research Development Unit (CRDU), 5Azar Hospital, Golestan University of
Medical Sciences, Gorgan, Iran.

Email: Panahandeh.milad@gmail.com

Khadije Layegh

Affiliation: Department of Research & Technology, Golestan University of Medical Sciences,
Gorgan, Iran.

Email: khadije.layegh@gmail.com

Abstract

High-risk pregnancy accounts for a disproportionate share of adverse maternal and perinatal outcomes worldwide, and continuous, reliable monitoring remains central to reducing this burden. Recent engineering and computer science research has produced wearable foetal and maternal sensors, machine learning models for obstetric risk prediction, digital twin frameworks, and integrated remote monitoring platforms, each reported with encouraging technical performance. Whether these systems have moved from laboratory validation into routine obstetric care is a separate question, and one this review addresses directly. Twenty-one recent studies were narratively synthesised across five technology clusters (wearable sensing, AI-based risk prediction, telehealth, IoT ecosystems, and equity-focused interventions), and this purposive synthesis identifies a recurring pattern across that set: technical feasibility is repeatedly demonstrated, while clinical validation, regulatory clearance, and workflow integration lag behind. Remote foetal and blood-pressure monitoring show measurable benefit in trial settings, yet most wearable sensing and AI risk-prediction studies remain single-centre, short-term, or unvalidated across diverse populations. Equity-focused and community-based interventions illustrate what implementation-aware design can achieve, but such examples remain the exception rather than the norm within the sources reviewed. Because the review is narrative rather than systematic, this pattern should be read as characteristic of the literature examined rather than as a comprehensive census of the field. The review concludes that closing the innovation-clinical gap will depend less on further algorithmic refinement than on prospective validation, regulatory clarity, interoperable systems, and clinician involvement from the design stage onward. Without these, promising technologies risk remaining permanently at the pilot stage.

Keywords: Wearable Sensors; Artificial Intelligence; High-risk Pregnancy; Digital Twin; Clinical Translation

1 .Introduction

Preterm birth, pre-eclampsia, gestational diabetes, and foetal growth restriction remain leading contributors to maternal and neonatal morbidity, and the pregnancies classed as high-risk require monitoring intensity that conventional antenatal scheduling struggles to sustain. Intermittent clinic visits capture only a narrow window of a physiological state that can change within hours. This mismatch between monitoring frequency and clinical need has driven a wave of technical development.

Wearable sensors now record foetal movement, foetal heart rate, and uterine electrical activity outside the clinic (Yap et al., 2025; Likitalo et al., 2026; Barnova et al., 2025). Machine learning models are being trained to flag preterm birth risk, premature rupture of membranes (PROM), and complications of gestational diabetes and hypertensive disorders before they become clinically obvious (Zheng et al., 2025; Kang et al., 2025; Kim et al., 2025; Taheri et al., 2026). Patient-specific computational simulations and population-scale AI cohorts propose to model an individual pregnancy's physiology directly, rather than relying only on discrete clinic visits (Scott and Oyen, 2024; Jayaratne et al., 2026). Each of these developments answers a real clinical limitation. None of them, on its own, has yet reached routine obstetric practice.

That distinction — between what has been engineered and what has been adopted — is the argument this review pursues. A sensor that performs well in a validation study is not the same thing as a sensor a midwifery team trusts, integrates into a shift pattern, and acts on at three in the morning. An algorithm with strong discriminative accuracy is not the same thing as a tool that has cleared a regulatory pathway, been reimbursed, and been tested across the population diversity a maternity unit actually serves. This gap between technological innovation and clinical application is not a footnote to the digital health literature on high-risk pregnancy. It is close to the central finding.

This review examines that gap across five specific angles: evidence maturity against bedside implementation, algorithmic performance against clinical utility, technical feasibility against health-system readiness, equity and access barriers, and the lag between regulatory approval and reimbursement mechanisms. This study is a narrative review, not a systematic review or meta-analysis; it draws on 21 sources selected for their direct relevance to these five angles rather than through an exhaustive database search. The rationale for this fixed source set, and its limits, are set out in the Methodology section below. Consistent with the authors' clinical background in midwifery and nursing, the analysis throughout is oriented toward workflow integration, clinician trust, and bedside implementation; the underlying sensor, signal-processing, and machine learning methods are described using the framing the cited sources themselves provide, rather than through independent technical re-evaluation.

The observation that technology can outpace its clinical translation is not new; versions of this bench-to-bedside gap have been described across medicine, artificial intelligence, and digital health more generally for years. What this review contributes is not that observation itself but its application to a specific, bounded comparison: reading wearable sensing, AI-based prediction, remote monitoring, integrated IoT and digital twin systems, and equity-focused interventions in high-risk pregnancy alongside one another, on the same five analytical angles, rather than treating each technology cluster as the subject of a separate review. That cross-cluster placement, developed in Sections 3 through 8 below, is where this review's contribution lies.

2. Methodology

This review follows a narrative, rather than systematic, methodology. Sources were drawn primarily from PubMed-indexed journals, with supplementary sources identified through targeted searches of Scopus and preprint or protocol registries, covering the period 2024 to 2026. One earlier source (2022) is retained as the closest available match for a low-resource, Latin American telehealth implementation,

a combination not found in the primary window. Search terms combined “wearable”, “artificial intelligence”, “machine learning”, “patient-specific simulation”, “remote monitoring”, and “telehealth” with “pregnancy”, “obstetric”, “foetal”, and named complications (preterm birth, PROM, gestational diabetes, hypertensive disorders).

Inclusion criteria were: peer-reviewed journal articles or registered study protocols; direct relevance to wearable sensing, AI-based prediction, remote monitoring, integrated digital ecosystems, or equity-focused implementation in pregnancy care; and publication within the 2024–2026 window, to keep the review anchored to the current state of the field. Sources addressing AI or wearables outside obstetric or maternal-foetal contexts were excluded, as were non-peer-reviewed commentary pieces.

The resulting set of 21 sources is not exhaustive and was not assembled through a PRISMA-style systematic process; it was selected to illustrate recurring patterns across five technology clusters rather than to comprehensively enumerate every study published in this space. This is a scope appropriate to a narrative review within a conference-paper page limit, not to a dedicated systematic review. This scope decision, and the selection bias it may introduce, is addressed directly in the Limitations subsection of the Discussion.

3. Wearable Sensing Technologies for Fetal and Maternal Monitoring

Foetal movement counting has long relied on maternal perception, a method prone to under- and over-reporting. Yap et al. (2025) describe a compact wearable combining pressure and strain sensing to track foetal movement continuously and non-invasively, positioning it as an objective alternative to subjective kick counts. The device's contribution is not a new physiological insight; it is a measurement method that removes a known source of noise from an established clinical signal.

Foetal heart rate monitoring shows a comparable trajectory. Likitalo et al. (2026) report a low-complexity foetal heart rate monitoring approach built on carbon-based, single-channel dry electrodes, designed to reduce the hardware burden of conventional systems while remaining usable outside a monitored bed. Complementing this, Xu X et al. (2026) present a deep learning-guided ensemble empirical mode decomposition method for extracting a clean foetal ECG signal from a single maternal-abdominal channel — a long-standing signal-processing challenge, since the foetal signal is weak relative to maternal cardiac and muscular interference. Their approach targets exactly this separation problem computationally rather than through additional hardware.

Uterine electrical activity offers a third monitoring route. Barnova et al. (2025) review electrohysterography across signal processing and machine learning applications in obstetrics, framing it as a technique capable of detecting labour-related uterine activity non-invasively, with potential relevance to preterm labour risk assessment.

That the field still needs stronger evidence synthesis is signalled by Ranaei-Zamani et al. (2025), who set out a protocol for a systematic review of wearable devices for antenatal foetal monitoring rather than reporting review findings themselves. A published protocol, rather than a completed synthesis, says something on its own: even the evidence-summarising work in this area is, at the time of writing, still in preparation.

Read together, these five sources describe a coherent engineering trajectory: each addresses a specific signal-acquisition, signal-processing, or evidence-synthesis limitation in maternal-foetal monitoring, and each demonstrates feasibility or intent within its own scope. What none of them addresses directly is deployment beyond the study setting — how a device performs across body habitus, skin tone, gestational age range, or ambulatory daily life, and whether the signal quality achieved in a laboratory or short clinical session survives continuous home use. Sensing accuracy and monitoring readiness are not the same claim, and the literature at this stage supports the former considerably more firmly than the latter.

4. AI and Machine Learning for Obstetric Risk Prediction

Machine learning is the latest tool in the longstanding quest to predict preterm birth before symptoms appear. Kang et al. (2025) develop an explainable machine learning model for preterm birth prediction specifically in pregnancies complicated by gestational diabetes or hypertensive disorders, using explainability methods so that clinicians can inspect which variables drive a given prediction rather than accept an opaque score. That design choice matters. A model that cannot explain itself is difficult for a midwife or obstetrician to act on with confidence, whatever its statistical performance.

Kim et al. (2025) take a different entry point, examining multidimensional predictors of preterm birth risk in Black and White primiparous women and generating insights through machine learning about how risk factors differ across these groups. The comparison itself is the contribution: it treats population-specific risk structure as a variable worth modelling, rather than assuming a single risk profile generalises across ethnic groups. Few of the other sources in this review attempt that.

PROM risk stratification receives similar computational attention. Zheng et al. (2025) apply an interpretable machine learning model, built on SHAP-based feature selection and a nomogram, to identify PROM risk factors and their association with a nutritional inflammatory index in a retrospective single-hospital cohort, again prioritising interpretability alongside predictive performance. Gestational diabetes, meanwhile, has attracted a wider body of AI work. Taheri et al. (2026) survey the transformative role of AI across gestational diabetes mellitus care broadly, while Lu et al. (2024) focus specifically on digital health and machine learning technologies for blood glucose monitoring in this population, covering continuous glucose monitoring integration and algorithmic glucose pattern analysis.

A further protocol from Iran illustrates how this same ambition is being pursued outside the health systems that dominate the rest of this cluster. Shabani et al. (2025) set out a multi-phase study protocol, developed jointly by midwifery, obstetric, and electrical engineering researchers at Tabriz University of Medical Sciences, to build and validate an AI-based application predicting low birth weight, pre-eclampsia, gestational diabetes, preterm birth, and intrauterine growth restriction from a single platform. The authors explicitly position their work against an earlier Iranian model constrained by small sample size and uncertain applicability beyond the region in which it was built. That an improved version of the same tool is still at the protocol stage, years after the original attempt, is itself a small but concrete illustration of how slowly regional validation work proceeds relative to the pace of model development.

What unites these six sources is a shared methodological commitment to explainability or predictive breadth, and a shared silence on prospective clinical deployment. Each model is developed and validated on a defined dataset, or, in the Iranian case, has not yet reached that stage. None is reported as having been trialled within a live obstetric decision pathway, where a clinician would act on the prediction in real time and where downstream outcomes, not just retrospective discriminative accuracy, could be measured. Explainable does not yet mean adopted.

5. Remote Monitoring and Telehealth: Evidence from Systematic Reviews

Unlike the wearable-sensing and AI-prediction literature, remote monitoring for high-risk pregnancy already has meta-analytic evidence behind it, which makes this section the strongest empirical anchor for the review's central argument. Wang et al. (2025) synthesise randomised trials of remote foetal health monitoring in high-risk pregnancies and report that it reduced adverse maternal-foetal outcomes and improved detection of abnormal non-stress tests relative to routine prenatal care. Zhang et al. (2025), working from a similar but not identical body of trials, report comparable benefit, including reductions in adverse outcomes such as neonatal asphyxia and preterm birth alongside increased rates of natural delivery.

That two independent meta-analyses converge on similar conclusions strengthens confidence in the underlying signal, though it does not eliminate an important caveat: both note that the included trials cluster heavily within a small number of countries and healthcare systems. A meta-analysis assembled largely from one national context tells a reviewer a great deal about that context and rather less about generalisability elsewhere.

Remote monitoring extends beyond foetal surveillance into maternal blood pressure. Rajkumar et al. (2024) review remote blood pressure monitoring in hypertensive disorders of pregnancy and report evidence supporting its use for the earlier detection of hypertensive complications compared with clinic-only measurement.

This body of evidence stands apart from the two sections preceding it. Remote foetal and blood pressure monitoring have been tested through randomised evaluation and meta-analysis, an evidentiary step the wearable-sensing and AI-prediction literature reviewed above has largely not yet reached. Even so, adoption at scale remains constrained by the concentration of trial evidence in specific health systems, a limitation this review returns to directly in Section 8.

6. Integrated IoT and Digital Ecosystem Approaches

A single sensor or algorithm solves a narrow problem. A different strand of the literature asks what happens when several such tools are combined into one system, and it is here that the ambitions become considerably larger. Rani et al. (2026) present NurtureNest, an IoT-wearable predictive analytics framework designed to assess maternal risk by combining continuous sensor data with predictive modelling in a single platform, aimed at moving beyond isolated point measurements toward ongoing risk profiling.

A comparable logic drives telehealth deployment in resource-constrained settings. Escobar et al. (2022) report a telehealth and education programme connecting a tertiary referral hospital with a secondary-care centre in a low-resource region of Colombia, aimed at extending obstetric-emergency response capacity into a context where specialist review is not always readily accessible. The value of this work lies less in novel sensing hardware than in showing that a telehealth-linked referral model can function within that one setting; whether the same design would hold elsewhere is not something a single regional implementation can answer.

Patient-specific computational modelling approaches this same integration problem from an earlier stage. Scott and Oyen (2024) highlight examples of finite-element models built from patient-specific ultrasound measurements of the uterus and cervix, used to simulate biomechanical deformation during pregnancy and to flag markers associated with preterm birth risk; a related example models how a caesarean-section scar from a previous pregnancy affects tissue integrity and uterine rupture risk in a subsequent one. These remain proof-of-concept simulations built from research-grade imaging data rather than deployed clinical tools, and the authors themselves note that extracting the detailed, patient-specific geometry these models require from routine, low-resolution ultrasound, rather than research-grade MRI, is still an open technical barrier.

Population-scale ambitions appear again in Jayaratne et al. (2026), who set out the study protocol for an AI-powered, population-based birth cohort in Sri Lanka. A protocol paper, by definition, describes intended methodology rather than reported outcomes; its inclusion here signals where the field is heading rather than what it has already achieved. These four sources share a pattern: each integrates existing sensing or predictive components into a more ambitious system architecture, and each sits at an earlier stage of real-world validation than the individual components reviewed in Sections 3 and 4. Integration multiplies capability on paper. It also multiplies the number of points at which clinical validation has yet to occur.

7. Equity, Access, and Behavioural Engagement

Technology that works in a well-resourced trial setting does not automatically reach the populations who might benefit most. This section reviews three sources that engage directly with that problem, each from a different angle.

Saleh et al. (2025) report an AI-based gamified intervention for maternal health delivered in Lebanon, tested through a community interventional trial. Gamification here is not a superficial add-on; it is a behavioural engagement strategy intended to sustain participation in a population where conventional digital health tools have struggled to hold attention over time. The community-based trial design is itself notable, since it places the intervention within an existing social structure rather than testing it in isolation.

A wider pattern emerges when this trial is set alongside the review by Correia et al. (2025), which scopes AI-driven approaches to personalised maternal and foetal health across the literature. Their review surfaces a recurring observation: much of the innovation surveyed elsewhere in this paper has been developed and validated in high-income health systems, with comparatively little attention paid to how such tools would need to be adapted, or whether they would work at all, in lower-resource contexts.

Foo et al. (2026) address a related but distinct problem: moving care to the patient rather than the patient to care, using AI, IoT, and point-of-care testing devices. Their framing inverts the usual assumption that patients travel to fixed monitoring infrastructure, proposing instead that monitoring infrastructure travel to the patient. This is a conceptual shift worth taking seriously, though it depends on exactly the device reliability and connectivity questions raised in Sections 3 and 6.

Read against the rest of this review, these three sources are the exception that proves the rule. Most of the literature surveyed here treats implementation context as an afterthought. These three treat it as the starting design constraint, and in doing so, they illustrate what the rest of the field would need to incorporate for the innovation-clinical gap to close in anything other than high-income settings.

8. The Innovation-Clinical Gap: A Structural Analysis

The preceding five sections describe five clusters of technology at five different stages of maturity. This section asks a harder question: what, specifically, separates a working device from an adopted one?

Technology readiness and clinical validation are frequently conflated, and the sources reviewed above illustrate why the two must be kept separate. A sensor achieving reliable signal extraction in a controlled session (Xu X et al., 2026; Likitalo et al., 2026) has cleared an engineering hurdle. It has not thereby cleared a regulatory one. None of the wearable-sensing or AI-prediction sources in this review report a completed regulatory submission or approval pathway. The remote-monitoring meta-analyses (Wang et al., 2025; Zhang et al., 2025; Rajkumar et al., 2024) come closer to clinical validation because they report randomised outcome data, yet even these draw on trials concentrated in a small number of health systems, which limits how far “validated” can honestly be said to extend. It is also worth being precise about what is being compared across these five clusters, since they are not equivalent in kind. A wearable sensor (Yap et al., 2025; Likitalo et al., 2026) is a physical medical device with its own hardware certification route; an AI prediction model (Kang et al., 2025; Zheng et al., 2025) is closer to software, which regulators increasingly treat as a distinct category; and a patient-specific biomechanical simulation (Scott and Oyen, 2024) is, at this stage, a research modelling technique with no settled regulatory category at all. Placing these on a single innovation-to-adoption spectrum, as Figure 1 does, is a useful simplification for comparison, not a claim that the three face the same approval pathway or the same timeline.

Nor does this pattern apply uniformly across all five clusters. Remote foetal and blood-pressure monitoring, discussed in Section 5, already rest on randomised trial evidence in a way the other four

clusters do not, which is precisely why it is treated as the comparative benchmark here rather than folded into the same story of unvalidated promise. The claim advanced here is not that every reviewed technology is equally far from clinical use, but that, cluster by cluster, validation and adoption consistently lag behind the pace of engineering progress within that cluster's own trajectory.

Workflow integration is a separate barrier again, and one the literature addresses only indirectly. An integrated IoT platform (Rani et al., 2026) presumes a clinical team willing and able to act on continuous data streams, yet the source reports nothing on how such data would be triaged, who would review it out of hours, or what happens when the system generates an alert a midwife judges to be a false positive. The patient-specific simulations discussed above (Scott and Oyen, 2024) raise a related but earlier-stage workflow question: who would order, interpret, and act on a biomechanical simulation built from a patient's own imaging, and at what point in antenatal care this would happen, are questions a two-page comment does not have space to resolve. Foo et al. (2026) engage with the broader workflow question more directly than most, proposing care delivered to the patient rather than at a fixed site, but even this proposal does not resolve who bears clinical responsibility when monitoring is continuous and staffing is not.

Evidence quality compounds these barriers. Much of the wearable-sensing and AI-prediction literature reviewed in Sections 3 and 4 is single-centre, and several sources are explicit about small or narrowly defined study populations (Kim et al., 2025). Kang et al. (2025) and Zheng et al. (2025) both build explainable models, a methodological strength, but explainability addresses trust in a prediction, not the separate question of whether that prediction has been tested prospectively on an independent population. Generalisability across geography and demography is addressed by only a minority of sources in this review, principally Kim et al. (2025) and Correia et al. (2025); most others are silent on it.

A further evidence gap sits underneath all of this: accuracy is not the same claim as outcome. Wearable sensors and AI-prediction models across this review are consistently evaluated against how well they detect a signal or classify a risk, not against whether using them changed a maternal or neonatal outcome such as mortality, severe morbidity, or emergency intervention. The remote-monitoring meta-analyses (Wang et al., 2025; Zhang et al., 2025; Rajkumar et al., 2024) are the exception, since randomised outcome data is precisely what a meta-analysis of trials provides. For the wearable-sensing, AI-prediction, and integrated-systems clusters, this review found no comparable outcome evidence among the 21 sources; a device or model can be accurate and still never be shown to make a mother or baby measurably safer. This distinction between accuracy and outcome is, in the authors' assessment, one of the more consequential gaps this review identifies, precisely because it is easy to overlook when a paper's own framing emphasises performance metrics.

Cost and reimbursement receive almost no direct treatment across the 21 sources, and that silence is telling in its own right. A technology cannot become standard obstetric care without a payment pathway, whether through public health-system funding or insurance reimbursement, and the absence of this discussion across a body of otherwise technically sophisticated work suggests the field is further from clinical translation than its engineering maturity implies.

Clinician trust follows a similar pattern. Explainable AI design (Zheng et al., 2025; Kang et al., 2025) is a partial answer, since a clinician can at least inspect why a model reached a given output. It is not a complete one. Trust also requires training, familiarity, and a track record of the tool performing correctly under real clinical pressure, none of which a validation study alone can establish.

Each barrier identified here, considered alone, might appear surmountable. Set together across the 21 sources, they describe something closer to a systemic lag. This synthesis is the authors' own interpretation rather than a conclusion stated directly by any single reviewed source: none of the 21 sources studies regulatory clearance or reimbursement policy as its primary subject, so the claim that engineering capability has outpaced these slower processes is an inference drawn across the set, not a reported finding. Regulatory clearance, reimbursement, interoperability, and workflow integration

emerge repeatedly as plausible barriers across the reviewed literature, but they cannot be established here as the definitive causes of limited clinical adoption to the exclusion of other explanations. Commercial decisions not to pursue certification, unresolved investment or market considerations, incomplete product development, and local health-system capacity may each also contribute, and the 21 sources reviewed were not designed to weigh these possibilities against one another. Closing the lag described above will likely require deliberate investment in these slower processes, but establishing their relative importance is a task for dedicated implementation research, not for a narrative synthesis of this kind.

9. Tables and Figures

Table 1. Summary of the 21 reviewed sources by technology type, evidence stage, and clinical stage

Study	Technology	Evidence Stage	Clinical Stage
Yap et al. (2025)	Wearable pressure-strain foetal movement sensor	Device validation	Pre-clinical
Likitalo et al. (2026)	Dry-electrode foetal heart rate monitor	Device validation	Pre-clinical
Xu X et al. (2026)	Deep learning foetal ECG extraction	Algorithm validation	Pre-clinical
Barnova et al. (2025)	Electrohysterography + ML review	Narrative synthesis	Mixed / early
Zheng et al. (2025)	Interpretable ML for PROM risk (SHAP/nomogram)	Retrospective single-centre model	Pre-clinical
Kang et al. (2025)	Explainable ML for preterm birth (GDM/HDP)	Model development	Pre-clinical
Kim et al. (2025)	ML predictors of preterm birth by population	Model development	Pre-clinical
Taheri et al. (2026)	AI review across GDM care	Narrative synthesis	Mixed / early
Lu et al. (2024)	ML/digital health glucose monitoring in GDM	Narrative synthesis	Early clinical use
Shabani et al. (2025)	AI application for multi-outcome pregnancy prediction	Study protocol	Pre-clinical
Wang et al. (2025)	Remote foetal monitoring	Meta-analysis of RCTs	Clinical evidence
Zhang et al. (2025)	Remote foetal monitoring	Meta-analysis of RCTs	Clinical evidence
Rajkumar et al. (2024)	Remote blood pressure monitoring	Meta-analysis	Clinical evidence
Rani et al. (2026)	NurtureNest IoT-wearable platform	Framework/platform development	Pre-clinical
Escobar et al. (2022)	Telehealth referral programme (Colombia)	Retrospective ecological evaluation	Single-region clinical use
Scott and Oyen (2024)	Patient-specific biomechanical simulation (uterus/cervix)	Proof-of-concept modelling	Pre-clinical
Jayarathne et al. (2026)	Population-scale AI birth cohort (Sri Lanka)	Study protocol	Pre-clinical
Saleh et al. (2025)	Gamified AI maternal health intervention (Lebanon)	Community interventional trial	Community clinical use
Correia et al. (2025)	AI for personalised maternal/foetal health	Scoping review	Mixed / early
Foo et al. (2026)	AI/IoT/POCT care-to-patient model	Conceptual/technical proposal	Pre-clinical
Ranaei-Zamani et al. (2025)	Wearable devices for antenatal monitoring	Systematic review protocol	Pre-clinical

Evidence stage reflects the type of study reported: device or algorithm validation and model development indicate performance testing on a defined dataset; narrative synthesis indicates a review of existing literature rather than new primary data; meta-analysis of RCTs and meta-analysis indicate pooled trial evidence; study protocol indicates a planned but not yet completed investigation. Clinical stage reflects how close the work is to deployed care: pre-clinical indicates no reported use with patients

outside a study setting; single-region or community clinical use indicates deployment within one defined setting; clinical evidence indicates outcome data from randomised or comparative trials.

Figure 1. Conceptual Position of Reviewed Technology Clusters on the Innovation-to-Clinical-Adoption Spectrum

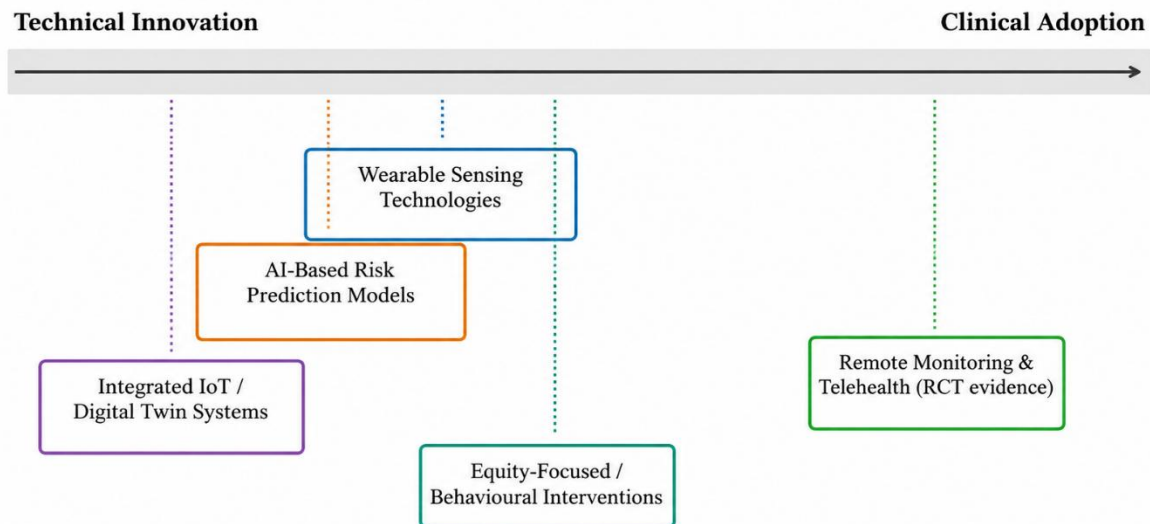


Figure 1. Conceptual position of the five reviewed technology clusters on a spectrum from technical innovation to clinical adoption. Placement is qualitative and derived from the authors' reading of the 21 sources, not from a quantitative technology-readiness scale reported in any single source. The figure is included as a heuristic summary of the comparative argument developed across Sections 3 to 8, not as a validated measurement instrument, and should be read alongside Table 1 rather than in place of it.

Table 2. How this review's scope relates to common categories of prior review work in this space

Typical scope in prior reviews	Technologies covered	Primary lens	This review, by contrast
Wearable-sensing reviews	Fetal/maternal sensors, signal processing	Device capability and signal accuracy	Treats wearables as one of five clusters, compared on adoption readiness rather than sensing accuracy alone
AI-prediction reviews	Risk-prediction algorithms for specific complications	Model performance and explainability	Reads AI prediction alongside remote monitoring and IoT clusters to compare relative evidence maturity
Low-resource/equity-focused reviews	mHealth and digital health broadly	Access and implementation barriers in under-resourced settings	Uses equity-focused sources as one point of comparison against better-resourced clusters, not as the sole lens
This review	Wearables, AI prediction, remote monitoring, IoT/digital twin, equity interventions	Cross-cluster comparison of innovation-to-adoption readiness	—

Table 2 characterises categories of prior review work in general terms rather than naming specific publications, since a named comparison would require introducing sources beyond the 21 that anchor this review's citation discipline. The categories reflect patterns commonly seen in this literature: reviews organised around a single technology type (wearables, or AI prediction, in isolation) or around a single implementation setting (equity or low-resource contexts), rather than a comparison across technology clusters on a shared set of adoption-readiness criteria.

10 .Discussion

Placed side by side, the six preceding sections resist a single tidy summary, and that resistance is itself informative. Remote foetal and blood-pressure monitoring stand out as the most clinically mature strand of this literature, supported by randomised evidence and meta-analysis (Wang et al., 2025; Zhang et al., 2025; Rajkumar et al., 2024). Wearable sensing (Yap et al., 2025; Likitalo et al., 2026; Xu X et al., 2026; Barnova et al., 2025) and AI-based risk prediction (Zheng et al., 2025; Kang et al., 2025; Kim et al., 2025; Taheri et al., 2026; Lu et al., 2024; Shabani et al., 2025) sit considerably earlier on that same trajectory, technically accomplished but largely unvalidated beyond the studies that produced them. Integrated IoT systems and patient-specific simulation approaches (Rani et al., 2026; Escobar et al., 2022; Scott and Oyen, 2024; Jayaratne et al., 2026) are earlier still, often at protocol, single-site, or proof-of-concept stage. Equity-focused work (Saleh et al., 2025; Correia et al., 2025; Foo et al., 2026) is the smallest cluster by volume, and arguably the most clinically honest about implementation context.

That remote monitoring carries the strongest evidence base in this review, yet still falls short of routine, system-wide adoption, is not a contradiction; it separates two different bottlenecks. Wang et al. (2025) and Zhang et al. (2025) demonstrate clinical benefit within the trials they analyse, which answers the evidence question. Adoption beyond those trial contexts depends on a different set of barriers neither meta-analysis was designed to test: whether remote monitoring is reimbursed as a billable service, whether its data stream integrates with the electronic record a unit already uses, and whether staffing models allow someone to review continuous data outside daytime hours. Rajkumar et al. (2024) hint at this by noting that included trials cluster in a small number of health systems, which are precisely the systems most likely to already have the reimbursement and staffing arrangements needed. Strong evidence, in other words, is necessary for adoption but not sufficient for it, and this review's five-cluster comparison makes that distinction visible in a way that no single meta-analysis, read alone, would.

A lag between engineering capability and clinical adoption is not unique to obstetrics; digital health more broadly shows the same pattern. What is specific to the high-risk pregnancy context reviewed here is the stakes attached to it. A wearable or algorithm in this setting is never assessed against a single patient's outcome alone; a false negative or a missed alert carries consequences for two interdependent physiologies at once, on a timescale that can shift from stable to emergent within hours. That dual-patient, time-critical character is part of why the cautious, slower pace of adoption documented across this review is not simply institutional inertia. It reflects a legitimately higher bar for evidence before a technology is trusted with two lives rather than one.

What would closing this gap actually require? Five things stand out from the analysis above, each tied to specific gaps identified in the sources reviewed.

Prospective validation studies, designed around patient-centred outcomes rather than algorithmic accuracy alone, are the most pressing need for the AI-prediction cluster specifically. Kang et al. (2025) and Zheng et al. (2025) report discriminative performance on retrospective data; both models now need testing in a prospective, multi-site trial that measures whether acting on the output changes a maternal or neonatal outcome, rather than only whether the prediction correlates with a retrospective label. Shabani et al. (2025) are, at the time of writing, still at the protocol stage of exactly this kind of validation; their trial, once completed, would be a natural test case for the broader cluster.

Regulatory pathway clarity follows closely behind, and applies most directly to the wearable-sensing cluster. None of Yap et al. (2025), Likitalo et al. (2026), or Xu X et al. (2026) describe a completed regulatory clearance process for the devices or algorithms they present. Device developers in this space would do well to engage a regulatory pathway, such as a software-as-a-medical-device or connected-sensor route, before rather than after publishing performance data, so that clearance status becomes something future papers can actually report.

Interoperability standards matter most for the integrated-systems cluster. An IoT framework such as NurtureNest (Rani et al., 2026) is only as useful as its ability to receive data from the electronic health record and existing monitoring equipment already in clinical use; the source does not specify which data standards, such as HL7 FHIR, the platform would use to exchange information with hospital systems, and naming one would be a low-cost, high-value addition for any follow-up work on it. The patient-specific simulation models discussed by Scott and Oyen (2024) face an earlier-stage version of the same problem: before interoperability is even a relevant question, these models need a pipeline capable of extracting their required geometry from routine, low-resolution ultrasound rather than research-grade imaging.

Clinician-in-the-loop design needs to move earlier in the development process than it currently sits, particularly for the explainable AI models in this review. Zheng et al. (2025) and Kang et al. (2025) build interpretability into the model's output, which helps a clinician audit a prediction after it is made. What is still missing is earlier involvement: midwives and obstetricians specifying which variables and thresholds a model should use before training begins, rather than only explaining its choices after the fact.

A related silence runs through nearly all 21 sources: none discusses what happens when the technology itself fails or misleads. A wearable sensor can detach or lose contact; a signal-processing algorithm can mistake maternal motion artefact for a foetal event; a prediction model trained on one population can generate a confident but wrong output on another. None of the sources reviewed here reports how clinicians should distinguish a genuine alert from a false one at the bedside, who is accountable when an algorithm's error contributes to a missed or delayed diagnosis, or how frontline staff should be trained to interpret a system's output rather than defer to it automatically. Reliability and error handling are not peripheral concerns; a technology that clinicians cannot trust to fail safely is unlikely to be adopted regardless of how accurate it is on average, and this gap deserves the same attention this review has given to validation and regulation.

Equity-centred development frameworks address a gap the sections above only partly close. Saleh et al. (2025), Correia et al. (2025), and Foo et al. (2026) show what happens when implementation context is designed for from the start rather than assumed away, yet they remain a minority within this literature. Extending their approach in practice would mean testing the sensors and models reviewed in Sections 3 and 4 within the same kind of community or resource-constrained settings that Saleh et al. (2025) and the Iranian protocol of Shabani et al. (2025) already occupy, instead of validating once in a well-resourced centre and presuming the result travels.

These five priorities are not equally urgent, in the authors' judgement. Prospective validation comes first, because none of the other four steps is worth taking for a model or device that has not yet shown it changes an outcome. Regulatory clarity and interoperability follow closely and can largely proceed in parallel with validation work already underway. Clinician-in-the-loop design and equity-centred testing matter most once a technology is closer to deployment, since both concern how, not whether, a validated tool is introduced into practice. This ordering reflects the authors' reading of the reviewed literature rather than a finding reported by any of the 21 sources themselves.

10. Limitations

This review has the limitations inherent to a narrative, rather than systematic, methodology. The 21 sources were selected for relevance to five predefined analytical angles rather than identified through an exhaustive, PRISMA-style search, which introduces a risk of selection bias: sources that fit the innovation-clinical gap argument may have been easier to notice than sources that complicate it. A related risk is that this selection process could under-represent cases where a wearable, algorithm, or platform has in fact reached full clinical adoption; such cases, being less novel from a research-publication standpoint, are less likely to appear as recent, highly cited journal articles and so are structurally less likely to have been captured by the search strategy described in the Methodology. No

meta-analysis or other quantitative synthesis was attempted, and claims about the field as a whole rest on qualitative pattern-matching across a bounded set of studies rather than pooled effect estimates. The review is also weighted toward sources published in English and indexed in mainstream biomedical databases, which may under-represent regionally significant work published elsewhere. A further, related limitation is that all 21 sources report technologies in broadly favourable terms; none documents an algorithm that underperformed after deployment, a wearable that was discontinued, or a project abandoned after piloting. Published literature is known to favour positive results generally, and this review has no way to correct for that tendency with the source set it draws on; the pattern described here should therefore be read as characteristic of what gets published about these technologies, which is not necessarily identical to everything that has been attempted with them. These constraints do not undermine the recurring pattern identified across sections, but they do mean the pattern should be read as an interpretation of a defined literature sample rather than a definitive census of the field.

This should not be mistaken for dismissing the technical achievements catalogued across this review. The engineering is frequently sound, occasionally elegant, and consistently well-documented. What it lacks, almost uniformly, is the slower institutional work that turns a working prototype into standard care, and that work deserves to be treated as the central research task ahead, not an afterthought once the engineering is finished.

Several of the 21 sources acknowledge their own limitations individually, a small sample here, a single health system there. What this review adds is not new primary data but a comparative reading placed alongside those individual admissions: set next to one another, the five technology clusters show the same limitation recurring in slightly different form at every stage, from single-sensor validation through to population-scale cohort protocols. That cross-cluster repetition, visible only once the sources are read together rather than one at a time, is this review's central contribution.

11 .Conclusion

The literature reviewed here does not describe a field waiting for a technical breakthrough. Wearable sensors already extract usable foetal signals, machine learning models stratify risk with reported discriminative accuracy, and remote monitoring shows measurable benefit in randomised trials. What this body of work still lacks is the connective tissue between demonstrated capability and routine obstetric practice: validation across populations, cleared regulatory pathways, systems that speak to existing clinical infrastructure, and clinicians involved from the point of design rather than the point of deployment.

Closing that gap is a different kind of project from building the next sensor or refining the next algorithm. It asks health systems, regulators, and technology developers to treat implementation as a research question that deserves the same rigour currently reserved for accuracy metrics, rather than as an administrative afterthought. Until that shift occurs, high-risk pregnancy monitoring will keep producing capable technologies that arrive at the bedside more slowly than the engineering that made them possible.

Declarations

Conflict of Interest: The authors declare no conflict of interest.

Funding: This research received no external funding.

Use of Artificial Intelligence: Generative AI tools were used to assist with language editing, structuring, and reference formatting during preparation of this manuscript. All source selection, interpretation, and scientific content were reviewed and approved by the authors, who take full responsibility for the accuracy of the manuscript.

Scope of Expertise: The authors' clinical background in midwifery and nursing informed the review's focus on workflow integration, clinician trust, and implementation barriers, which are the paper's central analytical contribution. Descriptions of the underlying sensor, signal-processing, and machine learning methods in Sections 3 and 4 are drawn directly from the framing given by the cited sources themselves, rather than from independent technical evaluation of the engineering.

References

1. Barnova, K., et al. (2025). Electrohysterography in modern obstetrics: Advances in signal processing, machine learning, and clinical applications. *Artificial Intelligence in Medicine*. Vol. 171. P. 103303.
2. Correia, V., et al. (2025). Smart Pregnancy: AI-Driven Approaches to Personalised Maternal and Foetal Health — A Scoping Review. *J Clin Med*. Vol. 14. No 19. 6974
3. Escobar, M. F., Echavarria, M. P., Vasquez, H., et al. (2022). Experience of a telehealth and education program with maternal and perinatal outcomes in a low-resource region in Colombia. *BMC Pregnancy and Childbirth*. Vol. 22. 604.
4. Foo, L., et al. (2026). Managing maternity: Moving care, not patients, using AI, IoT and POCT devices. *International Journal of Gynaecology and Obstetrics*. Vol. 173. No. 2. 602-613
5. Jayaratne, K., et al. (2026). AI-powered population-based birth cohort study, Sri Lanka: study protocol. *BMC Pregnancy and Childbirth*. Vol. 26. 616
6. Kang, L., et al. (2025). An explainable ML model for predicting preterm birth in GDM/HDP pregnancies. *Frontiers in Endocrinology*. Vol. 16. 1665935
7. Kim, S., Barandouzi, Z., Grant, S. et al. (2025). Multidimensional predictors of preterm birth risk in Black/White primiparous women — ML insights. *BMC Pregnancy and Childbirth*. Vol. 25. No. 1.
8. Likitalo, S., Anzanpour, A., Axelin, A., et al. (2026). Low-complexity fetal heart rate monitoring from carbon-based single-channel dry electrodes maternal electrocardiogram. *Physiological Measurement*. Vol. 47. No. 1.
9. Lu, H. Y., Ding, X., Hirst, J. E., et al. (2024). Digital Health and ML Technologies for Blood Glucose Monitoring and Management of Gestational Diabetes. *IEEE Reviews in Biomedical Engineering*. Vol. 17. 98-117
10. Rajkumar, T., Hennessy, A., Makris, A. (2024). Remote blood pressure monitoring in hypertensive disorders of pregnancy: SR and meta-analysis. *International Journal of Gynaecology and Obstetrics*. Vol. 169. No. 1. 89-104
11. Ranaei-Zamani, N., Kowobari, O., Siassakos, D., et al. (2025). Protocol for a systematic review of wearable devices for antenatal fetal monitoring. *Wellcome Open Res*. 10;365
12. Rani, D., Wala, T., Kumar, R., et al. (2026). NurtureNest: An IoT-wearable predictive analytics framework for maternal risk assessment. *Biomedical Physics & Engineering Express*. Vol. 12. No. 4.
13. Saleh, S., Arnaout, N. E., Sabra, N., et al. (2025). The Effectiveness of an Artificial Intelligence–Based Gamified Intervention for Improving Maternal Health Outcomes Among Refugees and Underserved Women in Lebanon: Community Interventional Trial. *JMIR mHealth and uHealth*. Vol. 13. e65599

14. Scott, A. K., Oyen, M. L. (2024). Virtual pregnancies: predicting and preventing pregnancy complications with digital twins. *The Lancet Digital Health*. Vol. 6. No. 7. e436-e437
15. Shabani, F., Jodeiri, A., Mohammad-Alizadeh-Charandabi, S., et al. (2025). Developing and validating an artificial intelligence-based application for predicting some pregnancy outcomes: a multi-phase study protocol. *Reproductive Health*. Vol. 22. No. 1. 99
16. Taheri, Z., Tajdari M., Ghadikolaie Y. H., et al. (2026). The transformative role of AI in gestational diabetes mellitus: advancements in screening, management, and long-term outcome prediction. *Journal of Diabetes & Metabolic Disorders*. Vol. 25. No. 1. 120
17. Wang, J., Dai, Y., Li, H., et al. (2025). Effectiveness of remote foetal health monitoring in improving maternal and foetal outcomes among high-risk pregnancies: A systematic review and meta-analysis. *International Journal of Nursing Studies*. Vol. 175. 105318
18. Xu, X., Zhang, Y., Rao, Y., et al. (2026). A Deep Learning-Guided Ensemble EMD Method for Single-Channel Fetal Electrocardiogram Extraction. *Sensors*. Vol. 26. No. 7. 2037
19. Yap, L. W., Levin, A., Jiang, Y., et al. (2025). An intelligent, compact wearable pressure-strain combo sensor system for continuous fetal movement monitoring. *Science Advances*. Vol. 11. No. 48. eady2661
20. Zhang, J., Li, Q., Li, M., et al. (2025). Effect of Remote Fetal Monitoring on Outcomes in High-Risk Pregnancies: A Systematic Review and Meta-analysis. *Journal of Ultrasound in Medicine*. Vol. 45. No. 2. 275-89
21. Zheng, M., Zhang, X., Wang, H., et al. (2025). Interpretable machine learning model for identification and risk factor of premature rupture of membranes (PROM) and its association with nutritional inflammatory index: a retrospective study. *Frontiers in Medicine*. Vol. 12. 1557919