Cureus
Part of SPRINGER NATURE

Open Access Review Article

Clinical Predictors of Successful Pregnancy After In Vitro Fertilization (IVF): A Comprehensive Systematic Review of Evidence

Review began 12/20/2025
Review ended 12/31/2025
Published 01/04/2026

© **Copyright** 2026
Mohamed Salih et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.100734

Egbal Lutfi Mohamed Salih [1], Abduraheem Farah [2], Hania Mohammed Saeed Mohammed [3], Hind Suliman Badre Adam [4], Reem Babkir Altayeb Abdullah [5]

1. Obstetrics and Gynecology, Sabt Alaya Hospital, Alkhaldiyah, SAU 2. Anatomical Sciences, St. George's University, St. George, GRD 3. Obstetrics and Gynecology, International University of Africa, Khartoum, SDN 4. Obstetrics and Gynaecology, Bidyia Hospital, Bidiya, OMN 5. Obstetrics and Gynecology, Ministry of National Gaurd Health Affairs, Dammam, SAU

Corresponding author: Abduraheem Farah, heemfarah93@gmail.com

Abstract

In vitro fertilization (IVF) success is influenced by a complex interplay of patient- and treatment-related factors. Identifying reliable clinical predictors is crucial for patient counseling and individualized treatment planning. This systematic review aimed to synthesize the most recent evidence on clinical predictors of successful pregnancy after IVF. A comprehensive search of five electronic databases (PubMed, Scopus, Excerpta Medica database (Embase), Web of Science, and ClinicalTrials.gov) was conducted for studies published between 2020 and 2025. Following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, studies evaluating clinical predictors of clinical pregnancy or live birth in women undergoing IVF were included. Study selection, data extraction, and risk of bias assessment (using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool) were performed independently by two reviewers. A narrative synthesis was conducted due to significant methodological heterogeneity. Eight retrospective observational studies, comprising data from 40,490 IVF/intracytoplasmic sperm injection (ICSI) cycles, were included. Female age was the most consistent and powerful predictor, with a non-linear negative impact, particularly beyond 40 years. Ovarian reserve markers, anti-Müllerian hormone (AMH) and antral follicle count (AFC), were significant, with evidence suggesting AMH better predicts oocyte yield while AFC may better forecast embryo availability. Embryo quality parameters (number of high-quality embryos) were strongly associated with success. Male factors, including total progressive motile sperm count (TPMC) and sperm DNA fragmentation index (DFI), added incremental predictive value. Several studies developed predictive models using both traditional logistic regression and machine learning (ML) algorithms (e.g., eXtreme Gradient Boosting (XGBoost) and Random Forest), which demonstrated high accuracy but raised concerns regarding interpretability and temporal validity. Successful IVF pregnancy is multifactorial, with female age, ovarian reserve, embryo quality, and sperm DNA integrity being key prognostic determinants. While ML-based models show promise, their clinical integration requires rigorous external validation and transparency. Future research should prioritize prospective, multi-center designs and the integration of novel dynamic parameters to advance toward truly personalized prognostic tools in reproductive medicine.

Categories: Obstetrics/Gynecology
Keywords: clinical predictors, in vitro fertilization, live birth rate, machine learning, ovarian reserve, pregnancy outcome, systematic review

Introduction And Background

Infertility is a significant global health concern, affecting approximately 10%-15% of couples of reproductive age worldwide [1]. Assisted reproductive technologies (ART), a group of medical procedures used to achieve pregnancy, have transformed the management of infertility, with in vitro fertilization (IVF) being the most widely used approach [2]. IVF involves the controlled stimulation of the ovaries, retrieval of oocytes, fertilization outside the body, and subsequent transfer of embryos into the uterus. Since its introduction in 1978, IVF has undergone substantial refinement, with advances in ovarian stimulation protocols, embryo culture systems, and embryo transfer techniques contributing to improved clinical outcomes [3]. Nevertheless, IVF success rates remain variable, and many individuals and couples require multiple treatment cycles without achieving a live birth, emphasizing the clinical importance of identifying factors associated with treatment success.

Successful pregnancy following IVF is determined by a complex interaction of biological, clinical, and procedural factors [4]. Among the most consistently reported predictors are maternal age and indicators of ovarian reserve, such as anti-Müllerian hormone (AMH), a serum marker reflecting the remaining follicular pool, and antral follicle count (AFC), an ultrasound-based measure of recruitable follicles [5]. Embryo quality, commonly assessed using morphological or developmental criteria, and endometrial receptivity, which refers to the uterus's capacity to allow embryo implantation, also play critical roles in IVF outcomes.

In addition, paternal factors, including sperm quality, as well as maternal lifestyle characteristics and comorbid conditions such as obesity, polycystic ovary syndrome (PCOS), and endometriosis, may further influence the probability of conception [6]. Identifying reliable clinical predictors is therefore essential for individualizing treatment strategies, counseling patients regarding prognosis, and optimizing decision-making throughout the IVF process [7].

Over the past decade, a growing body of research has examined predictors of IVF success, encompassing traditional demographic and clinical variables as well as emerging biomarkers and genetic factors [8,9]. However, substantial heterogeneity exists among studies with respect to patient characteristics, IVF protocols, outcome definitions (such as biochemical pregnancy, clinical pregnancy, or live birth), and methodological rigor. This variability limits the comparability of findings and challenges the translation of evidence into routine clinical practice, highlighting the need for a systematic and critically appraised synthesis of the literature.

A comprehensive understanding of clinical predictors of IVF success can support clinicians in stratifying patients according to their likelihood of achieving pregnancy, facilitating personalized treatment planning, and potentially improving the efficiency and cost-effectiveness of IVF programs. Moreover, synthesizing existing evidence can help identify gaps in knowledge, methodological shortcomings, and areas requiring further investigation, thereby informing future research directions in reproductive medicine.

Accordingly, this systematic review aims to synthesize recent evidence on clinical predictors of successful pregnancy following IVF. By focusing on studies published within the last five years (2020-2025), this review seeks to capture contemporary advances in the field, assess the quality and applicability of the available evidence, and provide insights relevant to clinicians, researchers, and interdisciplinary audiences involved in reproductive health and infertility care.

Review

Methodology

Protocol and Registration

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure methodological transparency and reproducibility [10]. A predefined protocol outlining the objectives, eligibility criteria, and methodological approach was developed prior to the literature search. While formal registration with the International Prospective Register of Systematic Reviews (PROSPERO) is ideal, this review was not registered as it was conducted as a rapid synthesis of the most recent evidence (2020-2025) to inform timely clinical guidance. All subsequent review stages were rigorously documented to uphold transparency.

Eligibility Criteria

Eligibility criteria were defined based on the Population, Intervention, Comparison, Outcomes, and Study design (PICOS) framework. Studies were included if they involved adult women undergoing in vitro fertilization (Population), evaluated clinical predictors associated with pregnancy outcomes (Intervention/Exposure), compared groups with successful versus unsuccessful pregnancy outcomes (Comparator), and reported clinical pregnancy or live birth rates as primary outcomes (Outcome). Both prospective and retrospective original studies, including observational cohorts and clinical trials, were considered. Only studies published in English between 2020 and 2025 were included to ensure the incorporation of the most recent and relevant evidence. Review articles, editorials, conference abstracts, and case reports were excluded.

Information Sources and Search Strategy

A comprehensive literature search was performed using multiple electronic databases, including PubMed, Scopus, Excerpta Medica database (Embase), Web of Science, and ClinicalTrials.gov. The search strategy combined keywords and MeSH terms related to “IVF,” “in vitro fertilization,” “clinical predictors,” and “pregnancy outcomes,” with Boolean operators applied to maximize retrieval. The last search was conducted on 15 November 2025. Reference lists of relevant articles were manually screened to identify additional eligible studies. The full search strings for all databases are provided in Appendix A.

Study Selection

All retrieved records were imported into EndNote X21 (Clarivate, London, United Kingdom) to manage references and remove duplicates. Two independent reviewers screened titles and abstracts to identify potentially eligible studies. Full-text articles were subsequently assessed for eligibility based on the predefined inclusion and exclusion criteria. Any discrepancies between reviewers were resolved through discussion or consultation with a third reviewer to ensure unbiased study selection.

Data Collection Process and Data Items

Data extraction was performed independently by two reviewers using a structured data collection form. Extracted data included country, study design, sample size, patient characteristics (such as age and infertility type), clinical predictors evaluated, methods of predictor assessment, and primary pregnancy outcomes. Key findings and statistical measures were also recorded to facilitate qualitative synthesis.

Risk of Bias Assessment

The quality of included studies and risk of bias were assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [11], which evaluates patient selection, index test, reference standard, and flow and timing. Applicability concerns were also considered. Discrepancies in risk of bias assessment were resolved through consensus between reviewers.

Data Synthesis

Due to substantial heterogeneity among the included studies-including variability in patient populations, predictor definitions, study designs, outcome measures, and follow-up durations, a formal meta-analysis was not feasible. Instead, a narrative synthesis was conducted to summarize and compare clinical predictors of successful pregnancy after IVF. Findings were organized to highlight consistent associations and identify gaps in the current evidence base.

Results

Studies Selection Process

The systematic search of five electronic databases (PubMed, Scopus, Embase, Web of Science, and ClinicalTrials.gov) initially identified 179 potential records. After the removal of 108 duplicate records, a total of 71 unique citations were screened based on their titles and abstracts. This initial screening led to the exclusion of 47 records that did not meet the broad inclusion criteria. The remaining 24 reports were sought for full-text retrieval, of which two could not be retrieved due to paywall restrictions. Therefore, 22 full-text articles were assessed for eligibility. Upon detailed evaluation, a further 14 articles were excluded for the following reasons: the study design was observational, a case report, or a review without original predictive data (n=4); the study did not involve women undergoing IVF (n=6); or the study lacked sufficient details on the clinical predictors of interest (n=4). Consequently, eight studies met all pre-defined eligibility criteria and were included in the final systematic review [12-19]. The selection process is summarized in the PRISMA flow diagram (Figure *1*).

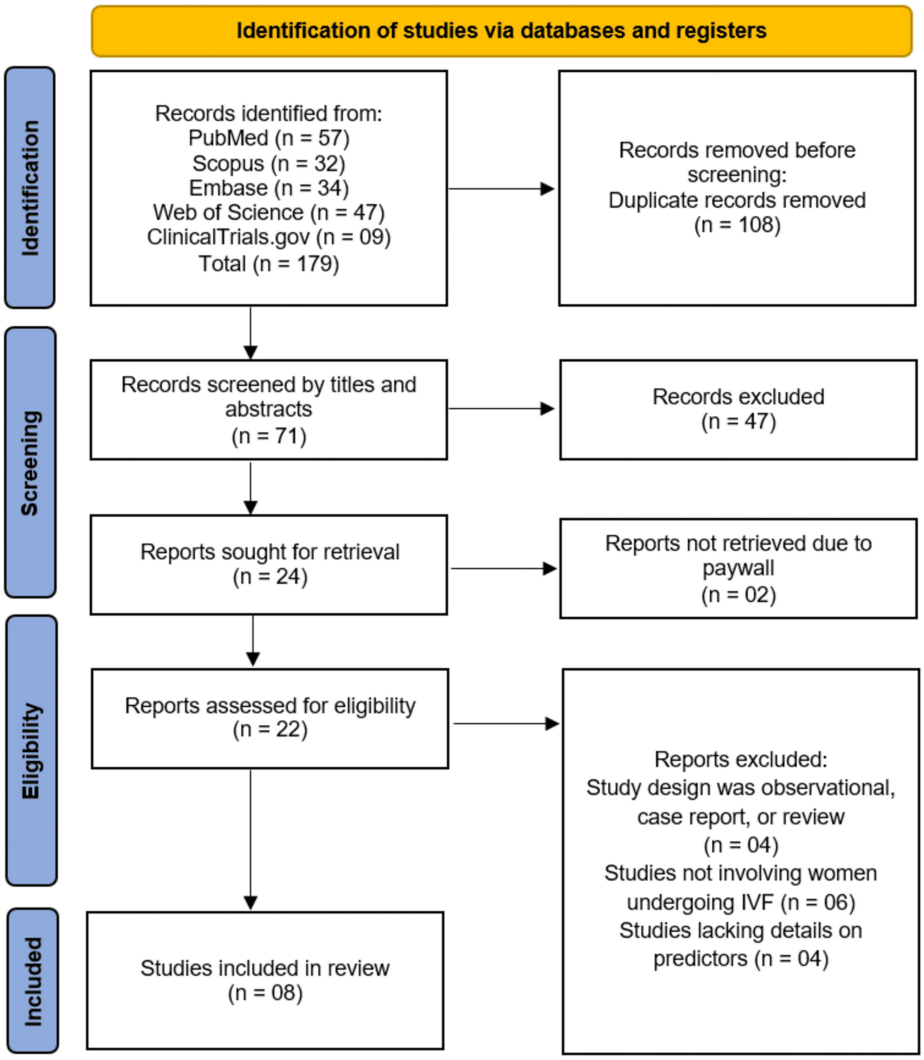


FIGURE 1: A PRISMA flowchart outlining the study selection process

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Embase: Excerpta Medica database; IVF: In Vitro Fertilization

Study Characteristics and Populations

A total of eight studies [12-19] published between 2022 and 2025 were included in this systematic review, providing data from 40,490 IVF or intracytoplasmic sperm injection (ICSI) cycles. The characteristics of these studies are summarised in Table *1*. All studies employed a retrospective observational design, with sample sizes ranging from 369 to 32,306 cycles. The primary outcomes assessed across the studies included clinical pregnancy, live birth, cumulative live birth (CLB), and fertilization failure. The predictors evaluated encompassed a wide range of female, male, and treatment-cycle factors, with analytical methods spanning traditional logistic regression to advanced machine learning (ML) techniques.

Study (Author, Year)	Country	Study Design	Sample Size (n)	Patient Population (e.g., Age, Infertility Type)	Predictors Evaluated	Predictor Measurement Method	Primary Outcome	Key Findings
Bai et al. (2025) [12]	China (Inner Mongolia)	Retrospective observational study with predictive model development	2,625	Women undergoing fresh IVF-ET cycles (2016–2022)	Multiple clinical features related to IVF-ET treatment	Clinical data preprocessing and machine-learning–based feature analysis using XGBoost and LightGBM models	Clinical pregnancy and live birth	XGBoost achieved the highest accuracy for predicting clinical pregnancy; LightGBM best predicted live birth
					Age; concurrent			Concurrent adenomyosis was associated

Zhu et al. (2025) [13]	China	Retrospective cohort study	1457	Women with endometriosis undergoing IVF/ICSI	adenomyosis; duration of infertility; number of fertilized oocytes; number of high-quality embryos	Clinical record review; univariate analysis and multivariable logistic regression	CLB rate	with reduced CLB; longer infertility duration decreased CLB; higher numbers of fertilized oocytes and high-quality embryos increased CLB; the predictive nomogram showed good discrimination
Wang et al. (2025) [14]	China	Retrospective single-center observational study	691 IVF cycles (594 c-IVF, 97 rescue ICSI)	Infertile couples undergoing conventional IVF; variables included female age, female BMI, male age, semen quality parameters, and infertility duration	Female age, female BMI, male age, sperm concentration, TPMC, DFI, and infertility duration	Clinical records and laboratory semen analysis; machine learning–based modeling	Fertilization failure in conventional IVF cycles	Logistic regression showed the best predictive performance; protective predictors were male age and TPMC, while female BMI and sperm DNA fragmentation index were significant risk factors for fertilization failure
Zeng et al. (2025) [15]	China	Retrospective observational study	1,403 IVF/ICSI cycles (1,039 patients); 441 fresh embryo transfer cycles	Patients with DOR; subgroup analyses by age (<40 vs ≥40 years)	AMH, AFC, basal FSH, age, D3 available cleavage-stage embryos, D3 top-quality embryos	Serum hormone assays (AMH, basal FSH), transvaginal ultrasound (AFC), embryo morphological grading on Day 3, logistic regression with ROC analysis	Oocyte retrieval, D3 available cleavage-stage embryos, clinical pregnancy (fresh cycle), viable blastocyst formation	AMH was a stronger predictor of oocyte retrieval than AFC and basal FSH; AFC showed superior predictive accuracy for obtaining D3 available embryos; D3 top-quality embryos predicted clinical pregnancy better than age in patients <40 years, whereas age was more predictive in patients ≥40 years
Bereczki et al. (2025) [16]	Hungary	Retrospective cohort study with model development and validation	1243 IVF/ICSI cycles (development cohort); 92 cycles	Women and couples undergoing IVF/ICSI; baseline characteristics collected at first consultation, including female age, BMI, hormonal profile, infertility duration, and male semen parameters	Female age, AMH, BMI, FSH, LH, infertility duration, sperm concentration, sperm motility, and other baseline clinical variables	Routine preprocedural clinical and laboratory measurements were collected at the first IVF consultation	IVF success per cycle	A nine-variable model showed strong predictive performance, with consistent discrimination in external validation. Female age was the strongest predictor, while AMH and BMI were key contributors; male factors added incremental predictive value.
Xia et al. (2024) [17]	China	Retrospective cohort study	32,306 complete IVF cycles (29,023 couples)	Women/couples undergoing IVF (2014–2020); varying female age; primary and secondary infertility; tubal factor, male factor, scarred uterus; multiple prior IVF attempts	Female age, AFC, BMI, number of previous IVF attempts, previous embryo transfer failures, infertility type, tubal factor, male factor, scarred uterus, number of oocytes retrieved	Clinical and laboratory measurements during pre-treatment, post-stimulation, and post-treatment phases; non-linear effects modeled using restricted cubic splines	CLBR	Three stage-specific predictive models showed good discrimination. Female age and oocyte number had non-linear associations with CLBR. Internal validation showed good calibration; temporal validation showed reduced accuracy, possibly due to improvements in IVF practices over time.
Liu et al. (2024) [18]	China	Retrospective cohort	1405	Patients undergoing IVF; age not specified, infertility type not specified	Age, OSI, COS regimen, Gn starting dose, endometrial thickness on HCG day, Progesterone value on HCG day, embryo transfer strategy	Clinical records; IVF cycle data; laboratory measurements	LBO	Seven predictors significantly associated with LBO; ANN and SVM models built, SVM performed best; models can predict LBO and assist in customizing embryo transfer strategy
Yang et al., (2022) [19]	China	Nested Case-Control	369	Women receiving IVF-ET (age and infertility type not specified)	Age, BMI, Number of cycles, HCT, LH, P, Endometrial thickness, FSH	Clinical/laboratory measurement of hormones, hematocrit, BMI, endometrial thickness; IVF treatment records	Clinical pregnancy after IVF-ET	Age, BMI, P, FSH were positively associated with clinical pregnancy; three cycles, HCT, LH, and endometrial thickness were negatively associated; the random forest model validated to predict clinical pregnancy outcomes.

TABLE 1: Summary of Included Studies Evaluating Clinical Predictors of Pregnancy Outcomes Following IVF

IVF-ET: In Vitro Fertilization and Embryo Transfer; ICSI: Intracytoplasmic Sperm Injection; CLB: Cumulative Live Birth; TPMC: Total Progressive Motile Count; DFI: DNA Fragmentation Index; DOR: Diminished Ovarian Reserve; AMH: Anti-Müllerian Hormone; AFC: Antral Follicle Count; FSH: Follicle-

Stimulating Hormone; LH: Luteinizing Hormone; OSI: Ovarian Sensitivity Index; COS: Controlled Ovarian Stimulation; hCG: Human Chorionic Gonadotropin; LBO: Live Birth Outcome; ANN: Artificial Neural Network; AVM: Adaptive Vector Machine; HCT: Hematocrit; P: Progesterone; D3: Day 3 embryos.

The included studies were predominantly conducted in China (n=7), with one study from Hungary [16]. Patient populations varied, including general IVF populations [12, 17, 18], women with specific conditions such as endometriosis [13] or diminished ovarian reserve (DOR) [15], and couples undergoing conventional IVF [14]. Study designs were exclusively retrospective, comprising cohort studies [13, 16-18], model development studies [12, 14, 16], and one nested case-control study [19]. The large variation in sample size is notable, with Xia et al. [17] analyzing a substantial cohort of over 32,000 cycles, providing considerable statistical power.

Predictors of Pregnancy and Live Birth Outcomes

Female age was consistently identified as a critical predictor across multiple studies. It was reported as the strongest individual predictor of IVF success in several analyses [16, 17]. The relationship was often non-linear, with advanced maternal age (typically ⩾40 years) being associated with significantly reduced probabilities of success [15, 17]. Ovarian reserve markers, including AMH and AFC, were also significant predictors. Zeng et al. [15] found AMH to be a stronger predictor of oocyte retrieval than AFC and basal FSH in women with DOR, while AFC was superior for predicting the availability of Day-3 embryos. Bereczki et al. [16] confirmed AMH as a key contributor in predictive models.

Other female factors played important roles. BMI was identified as a significant risk factor, with higher BMI associated with reduced success in some studies [14, 16, 19]. For patients with endometriosis, Zhu et al. [13] found that concurrent adenomyosis was independently associated with a reduced CLB rate. Endometrial thickness was another recurring factor, though its association was inconsistent; Yang et al. [19] reported a negative association with clinical pregnancy, whereas Liu et al. [18] included it in their predictive model for live birth.

Male factors contributed incremental predictive value. Wang et al. [14] identified total progressive motile sperm count (TPMC) as a protective factor against fertilization failure, while sperm DNA fragmentation index (DFI) was a significant risk factor. Bereczki et al. [16] also noted that incorporating male semen parameters improved model performance.

Embryological parameters were strongly predictive. The number of fertilized oocytes and the number of high-quality or top-quality embryos were positively associated with CLB and clinical pregnancy rates [13, 15]. Zeng et al. [15] highlighted that in patients under 40 years, the quality of Day-3 embryos was a better predictor of clinical pregnancy than age itself.

Performance and Validation of Predictive Models

Several studies developed and tested multivariable predictive models. Traditional statistical methods like logistic regression were used effectively [13-15]. For instance, Zhu et al. [13] developed a nomogram based on logistic regression that showed good discrimination for predicting CLB in endometriosis patients.

A prominent trend was the application of ML algorithms, which generally demonstrated high predictive accuracy. Bai et al. [12] compared multiple ML techniques, reporting that the XGBoost model achieved the highest accuracy for predicting clinical pregnancy, while LightGBM was best for predicting live birth. Similarly, Liu et al. [18] found that a support vector machine (SVM) model outperformed an artificial neural network (ANN) for predicting live birth outcomes. Yang et al. [19] validated a random forest model for predicting clinical pregnancy. Bereczki et al. [16] developed a nine-variable model, likely using ML methods, which showed strong and consistent predictive performance upon external validation.

However, model performance over time can be variable. Xia et al. [17] developed stage-specific predictive models with good initial discrimination and calibration but noted that temporal validation showed reduced accuracy, potentially due to evolving IVF laboratory and clinical practices.

The evidence synthesized from these eight studies indicates that successful pregnancy outcomes following IVF are multifactorial. The most robust predictors include female age, ovarian reserve markers (AMH, AFC), embryo quality parameters, and specific male factor semen analyses. Predictive modelling using both traditional regression and advanced ML techniques shows promise for individualized prognosis, though the generalizability and temporal stability of these models require further validation.

Risk of Bias Assessment

The methodological quality of the eight included studies was assessed using the QUADAS-2 tool, adapted for

predictive model studies. Regarding risk of bias, most studies demonstrated low risk in the domains of the index test (predictors and model development), reference standard (pregnancy outcomes), and flow and timing [13, 15-17]. However, concerns were identified in patient selection for three studies due to either unclear inclusion criteria [12] or insufficient detail on key population characteristics such as age and infertility type [18, 19], resulting in a high risk of selection bias. Furthermore, one study was rated high risk in the flow and timing domain because it analyzed IVF cycles rather than independent patients, potentially violating statistical independence [14]. In terms of applicability to the review question, all studies were deemed to have low concern regarding the predictors and outcomes evaluated. Nonetheless, applicability regarding patient selection was of great concern for two studies [18, 19] due to their poorly described populations, limiting generalizability. Overall, while the core predictive methodologies in most studies [13-17] appear robust, the findings from Bai et al. [12], Liu et al. [18], Yang et al. [19], and Wang et al. [14] should be interpreted with caution due to the identified biases (Table *2*).

Study (Author, Year)	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow & Timing	Patient Selection	Index Test	Reference Standard
Bai et al. (2025) [12]	□ High	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Zhu et al. (2025) [13]	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Wang et al. (2025) [14]	□ Low	□ Low	□ Low	□ High	□ Low	□ Low	□ Low
Zeng et al. (2025) [15]	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Bereczki et al. (2025) [16]	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Xia et al. (2024) [17]	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Liu et al. (2024) [18]	□ High	□ Low	□ Low	□ Low	□ High	□ Low	□ Low
Yang et al. (2022) [19]	□ High	□ Low	□ Low	□ Low	□ High	□ Low	□ Low

TABLE 2: Risk of Bias Assessment for the Included Studies Using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) Tool

Discussion

This systematic review synthesises evidence from eight recent studies exploring the clinical predictors of successful pregnancy following IVF. The findings collectively underscore the multifactorial nature of IVF success, identifying a core set of robust prognostic factors while highlighting the burgeoning role of ML in crafting personalised predictive models. The most consistent and potent predictors emerging from this synthesis are female age, markers of ovarian reserve (AMH and AFC), embryo quality parameters, and specific semen analysis metrics. These factors align with the fundamental biological pillars of reproduction: ovarian capacity, oocyte and embryo competence, and spermatozoal contribution. The pronounced and often non-linear negative impact of advanced female age, particularly beyond 40 years, is a recurrent theme, corroborated by multiple studies in this review [15-17]. This finding is biologically intuitive, reflecting the well-documented decline in oocyte quantity and quality, increased aneuploidy rates, and altered endometrial receptivity associated with aging. Our review reinforces that age is not merely a demographic variable but the single strongest clinical predictor, a conclusion that echoes decades of epidemiological and clinical research in reproductive medicine.

The assessment of ovarian reserve through AMH and AFC has become a cornerstone of pre-IVF evaluation, and our findings provide nuanced insights into their predictive utility. Zeng et al. [15] offered a valuable distinction, demonstrating that in patients with DOR, AMH was superior for predicting oocyte yield, while AFC was more accurate for forecasting the availability of viable Day-3 embryos. This suggests these markers, while correlated, may inform different stages of the IVF process, with AMH reflecting the follicular pool and AFC offering a more direct, cycle-specific snapshot. This is consistent with existing literature that positions AMH as a robust marker of ovarian aging and quantitative response, while AFC may have additional value in predicting oocyte maturity. The confirmation by Bereczki et al. [16] that AMH is a key contributor in a multi-factorial model further solidifies its clinical relevance. However, it is crucial to contextualize these markers; they are excellent predictors of ovarian response but are imperfect surrogates for oocyte quality and ultimate live birth, a distinction highlighted by the independent predictive value of embryo quality parameters in our reviewed studies [13, 15].

The significance of embryo morphology, specifically the number of high-quality or top-quality embryos, as a powerful predictor of CLB and clinical pregnancy [13, 15] underscores a critical transition in the prognostic timeline: from pre-treatment potential to actual laboratory outcome. Zeng et al.'s [15] finding that in younger patients (<40 years), day-3 embryo quality surpassed female age as a predictor of clinical pregnancy is particularly instructive. It implies that once a cohort of good-quality embryos is obtained from a young ovary, the age-related decline in endometrial receptivity may be a lesser barrier to success than the inherent viability of the embryos themselves. This aligns with the broader paradigm shift in IVF towards embryo-centric selection, championed by the widespread adoption of comprehensive chromosome screening (PGS/PGT-A), which seeks to directly assess embryonic competence beyond morphology. Our reviewed studies, focused on clinical and embryological predictors, complement the genetic landscape, suggesting that even without genetic testing, traditional embryo grading retains significant prognostic power.

The inclusion of male factors in several studies [14, 16] marks an important evolution from a predominantly female-centric prediction model to a more holistic couple-focused approach. Wang et al. [14] identified TPMC as protective and sperm DFI as a significant risk factor for fertilization failure in conventional IVF cycles. This reinforces a growing body of evidence that sperm contribution extends beyond simple motility and morphology to include genomic integrity. Bereczki et al. [16] further noted that male parameters added incremental predictive value, challenging the historical underestimation of the male role in IVF outcomes beyond severe male factor infertility. These findings are supported by external literature, such as the work of Osman et al. [20], which demonstrated that high sperm DNA fragmentation negatively impacts blastocyst development and pregnancy rates, and the systematic review by Cissen et al. [21], which confirmed the clinical value of advanced sperm function tests. Our review consolidates the view that comprehensive male assessment, including DFI, should be integrated into pre-IVF prognostic evaluations.

A striking trend across the included studies is the comparative application of ML algorithms alongside traditional logistic regression. Studies by Bai et al. [12], Liu et al. [18], and Yang et al. [19] demonstrated that models like XGBoost, SVM, and Random Forest could achieve high predictive accuracy, often outperforming or complementing conventional statistics. This aligns with a broader movement in medicine towards leveraging complex, non-linear algorithms to handle high-dimensional clinical data. For instance, the work of Salih et al. [22] in Fertility and Sterility successfully used deep learning to predict live birth from time-lapse embryo imaging data, surpassing embryologist assessment. Similarly, a study by Liu et al. [23] developed an ML model integrating clinical and genetic data to predict IVF success with high discrimination. The advantage of ML lies in its ability to model intricate interactions and non-linear relationships without pre-specified assumptions, as seen in Xia et al.'s [17] use of restricted cubic splines to capture the non-linear effect of age and oocyte number. However, the "black box" nature of some ML models can limit clinical interpretability compared to the transparent odds ratios of a logistic regression nomogram, as produced by Zhu et al. [13]. The critical challenge, noted by Xia et al. [17], is model decay over time due to evolving clinical practices, underscoring the need for continuous external validation and updating, a requirement even more pressing for complex ML models to ensure their generalizability beyond the development cohort.

When contextualized within the wider existing literature, our synthesized findings both confirm and refine established knowledge. The primacy of female age is universally acknowledged, mirroring conclusions from large registry analyses such as those of the Society for Assisted Reproductive Technology (SART) database in the United States and the Human Fertilisation and Embryology Authority (HFEA) data in the United Kingdom. Our observation on the differential predictive strengths of AMH and AFC finds resonance in the meta-analysis by Tal & Seifer [24], which also debated their respective roles in predicting poor response versus live birth. The strong predictive value of embryo quality is a cornerstone of IVF practice, supported by landmark studies such as the one by Qiu et al. [25], which established the correlation between blastocyst morphology and implantation potential. Furthermore, the incremental value of male factors corroborates the conclusions of the systematic review by Goyal et al. [26] on sperm DFI and IVF outcomes. The promising results from ML applications in our review are part of a rapidly expanding field, evidenced by the pioneering work of Zhang et al. [27], who developed an AI platform to predict implantation potential from static embryo images. However, our review also highlights gaps often present in the existing literature, such as the under-reporting of CLB rates, addressed by Zhu et al. [13] and Xia et al. [17], which provide a more patient-centered outcome than singleton cycle-based pregnancy rates.

Despite these coherent findings, the risk of bias assessment necessitates a cautious interpretation. The high risk of selection bias in studies by Bai et al. [12], Liu et al. [18], and Yang et al. [19], primarily due to unclear inclusion criteria or poorly described populations, means their reported predictor strengths or model accuracies may not be fully generalizable. Similarly, the analysis of cycles rather than unique patients by Wang et al. [14] introduces a unit-of-analysis error that could inflate the statistical significance of their findings. These methodological limitations are not uncommon in retrospective, single-center model development studies, which form the bulk of the current evidence. They highlight a critical need for future research to adhere to stricter reporting guidelines, such as the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement for predictive model studies, and to prioritize prospective, multicenter designs with explicit protocols for patient selection and clear documentation of population characteristics. The high applicability concerns for two studies [18, 19] further remind us that predictive models are only as good as the population they were derived from; a model built on an obscure population has limited utility in broader clinical practice.

Limitations

This systematic review has several limitations. First, the exclusive inclusion of retrospective studies inherently carries risks of confounding and bias that cannot be fully mitigated, despite our quality assessment. Second, all but one study were conducted in China, potentially limiting the generalizability of the findings to other ethnic and clinical practice populations, as ovarian reserve markers and treatment protocols can vary globally. The Hungarian study [16] offers a valuable but singular point of external validation. Third, the heterogeneity in predictor selection, outcome definitions (e.g., clinical pregnancy vs. CLB), and modelling techniques precluded a formal meta-analysis, limiting our synthesis to a narrative summary. Fourth, the search was restricted to published literature, introducing the possibility of publication bias, where studies with negative or null findings may be underrepresented. Finally, the rapidly evolving field of IVF, with continuous improvements in laboratory techniques (e.g., vitrification, PGT-A, time-lapse imaging) and stimulation protocols, means that predictors identified from cycles performed several years ago may have changing relevance, as hinted at by the temporal validation results of Xia et al. [17].

Conclusions

This systematic review affirms that successful pregnancy after IVF is predicated on a constellation of clinical factors, with female age, ovarian reserve markers, embryo quality, and sperm DNA integrity standing out as key prognostic determinants. The evidence demonstrates a promising adjunct role for ML models in synthesizing these variables into personalized predictions, though they have yet to surpass the clinical interpretability and established utility of traditional regression-based tools like nomograms. The path forward requires a concerted effort to strengthen the methodological rigor of predictive research through prospective, multi-center collaborations and adherence to robust reporting standards. Future studies should strive to integrate novel, dynamically assessed parameters, such as time-lapse morphokinetics, proteomic profiles of endometrial fluid, or advanced sperm functional assays, into these predictive frameworks. Ultimately, the goal is to move beyond population-level statistics to provide couples with truly individualized prognostic estimates, thereby guiding clinical decision-making, managing expectations, and optimizing the allocation of resources in the emotionally and financially demanding journey of IVF.

Appendices

Appendix A

Database	Search String
PubMed (MEDLINE)	("In Vitro Fertilization"[Mesh] OR IVF OR "Assisted Reproductive Technology" OR ART) AND ("Pregnancy Outcome"[Mesh] OR pregnancy OR "clinical pregnancy" OR "live birth") AND (predict* OR determinant* OR factor* OR prognostic* OR predictor*)
Scopus	TITLE-ABS-KEY ("in vitro fertilization" OR IVF OR "assisted reproductive technology" OR ART) AND TITLE-ABS-KEY (pregnancy OR "clinical pregnancy" OR "live birth" OR outcome*) AND TITLE-ABS-KEY (predict* OR determinant* OR factor* OR prognostic*)
Embase	('in vitro fertilization'/exp OR IVF OR 'assisted reproductive technology') AND ('pregnancy outcome'/exp OR pregnancy OR 'clinical pregnancy' OR 'live birth') AND (predict* OR determinant* OR factor* OR prognostic*)
Web of Science (Core Collection)	TS=("in vitro fertilization" OR IVF OR "assisted reproductive technology" OR ART) AND TS=(pregnancy OR "clinical pregnancy" OR "live birth" OR outcome*) AND TS=(predict* OR determinant* OR factor* OR prognostic*)
ClinicalTrials.gov	Condition or disease: Infertility OR IVF AND Other terms: pregnancy outcome OR live birth OR predictor OR prognostic

TABLE 3: Search Strategy for Databases

MEDLINE: Medical Literature Analysis and Retrieval System Online; Embase: Excerpta Medica database

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Aburaheem Farah, Egbal Lutfi Mohamed Salih, Hania Mohammed Saeed Mohammed, Hind Suliman Badre Adam

Critical review of the manuscript for important intellectual content: Aburaheem Farah, Hania

Mohammed Saeed Mohammed, Reem Babkir Altayeb Abdullah

Supervision: Abduraheem Farah

Acquisition, analysis, or interpretation of data: Egbal Lutfi Mohamed Salih, Hind Suliman Badre Adam, Reem Babkir Altayeb Abdullah

Drafting of the manuscript: Egbal Lutfi Mohamed Salih, Hania Mohammed Saeed Mohammed, Hind Suliman Badre Adam, Reem Babkir Altayeb Abdullah

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

The authors gratefully acknowledge Egbal Lutfi Mohamed Salih and Abduraheem Farah for their equal and substantial contributions to this work. They are to be considered joint first authors. Though affiliated with different institutions, all authors collaborated equally on this research. We share a common interest in advancing reproductive medicine and continue to collaborate on joint research projects.

References

1. Liang Y, Huang J, Zhao Q, et al.: Global, regional, and national prevalence and trends of infertility among individuals of reproductive age (15-49 years) from 1990 to 2021, with projections to 2040. Hum Reprod. 2025, 40:529-44. 10.1093/humrep/deae292
2. Tenchov R, Zhou QA: Assisted reproductive technology: a ray of hope for infertility . ACS Omega. 2025, 10:22347-65. 10.1021/acsomega.5c01643
3. Abdullah KA, Atazhanova T, Chavez-Badiola A, Shivhare SB: Automation in ART: paving the way for the future of infertility treatment. Reprod Sci. 2023, 30:1006-16. 10.1007/s43032-022-00941-y
4. Wang C, Li QL, Xu YS, et al.: Interplay of endocrine and psychological factors in IVF/ICSI outcomes: a prospective cohort analysis. Front Endocrinol (Lausanne). 2025, 16:1596664. 10.3389/fendo.2025.1596664
5. Ribeiro S, Sousa M: In vitro fertilisation and intracytoplasmic sperm injection predictive factors: a review of the effect of female age, ovarian reserve, male age, and male factor on IVF/ICSI treatment outcomes. JBRA Assist Reprod. 2023, 27:97-111. 10.5935/1518-0557.20220000
6. Khan HL, Bhatti S, Suhail S, et al.: Antral follicle count (AFC) and serum anti-Müllerian hormone (AMH) are the predictors of natural fecundability have similar trends irrespective of fertility status and menstrual characteristics among fertile and infertile women below the age of 40 years. Reprod Biol Endocrinol. 2019, 17:20. 10.1186/s12958-019-0464-0
7. Cedars MI: Evaluation of female fertility-AMH and ovarian reserve testing . J Clin Endocrinol Metab. 2022, 107:1510-9. 10.1210/clinem/dgac039
8. Hernández-Vargas P, Muñoz M, Domínguez F: Identifying biomarkers for predicting successful embryo implantation: applying single to multi-OMICs to improve reproductive outcomes. Hum Reprod Update. 2020, 26:264-301. 10.1093/humupd/dmz042
9. Sadegh-Zadeh SA, Khanjani S, Javanmardi S, Bayat B, Naderi Z, Hajiyavand AM: Catalyzing IVF outcome prediction: exploring advanced machine learning paradigms for enhanced success rate prognostication. Front Artif Intell. 2024, 7:1392611. 10.3389/frai.2024.1392611
10. Page MJ, McKenzie JE, Bossuyt PM, et al.: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021, 372:n71. 10.1136/bmj.n71
11. Whiting PF, Rutjes AW, Westwood ME, et al.: QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011, 155:529-36. 10.7326/0003-4819-155-8-201110180-00009
12. Bai R, Li JW, Hong X, Xuan XY, Li XH, Tuo Y: Predictive modeling of pregnancy outcomes utilizing multiple machine learning techniques for in vitro fertilization-embryo transfer. BMC Pregnancy Childbirth. 2025, 25:316. 10.1186/s12884-025-07433-2
13. Zhu L, Lin T, Yang X, Zhang D: Nomogram to predict cumulative live birth rate following in vitro fertilization/intracytoplasmic sperm injection cycles in patients with endometriosis. BMC Pregnancy Childbirth. 2025, 25:38. 10.1186/s12884-025-07147-5
14. Wang H, Pan H, Xu Z, Zheng X, Xia S, Zheng J: Development of a single-center predictive model for conventional in vitro fertilization outcomes excluding total fertilization failure: implications for protocol selection. J Ovarian Res. 2025, 18:138. 10.1186/s13048-025-01728-w
15. Zeng Y, Zhang Y, Cao Q, Gao Z, Tang T: Independent predictors and thresholds of in vitro fertilization outcomes in patients with diminished ovarian reserve. Sci Rep. 2025, 15:18875. 10.1038/s41598-025-01937-9
16. Bereczki K, Bukva M, Vedelek V, et al.: Machine learning-based prediction of IVF outcomes: the central role of female preprocedural factors. Biomedicines. 2025, 13:2768. 10.3390/biomedicines13112768
17. Xia L, Han S, Huang J, et al.: Predicting personalized cumulative live birth rate after a complete in vitro fertilization cycle: an analysis of 32,306 treatment cycles in China. Reprod Biol Endocrinol. 2024, 22:65.

10.1186/s12958-024-01237-3
18. Liu L, Liang H, Yang J, Shen F, Chen J, Ao L: Clinical data-based modeling of IVF live birth outcome and its application. Reprod Biol Endocrinol. 2024, 22:76. 10.1186/s12958-024-01253-3
19. Yang H, Liu F, Ma Y, Di M: Clinical pregnancy outcomes prediction in vitro fertilization women based on random forest prediction model: a nested case-control study. Medicine (Baltimore). 2022, 101:e32232. 10.1097/MD.0000000000032232
20. Osman A, Alsomait H, Seshadri S, El-Toukhy T, Khalaf Y: The effect of sperm DNA fragmentation on live birth rate after IVF or ICSI: a systematic review and meta-analysis. Reprod Biomed Online. 2015, 30:120-7. 10.1016/j.rbmo.2014.10.018
21. Cissen M, Wely MV, Scholten I, et al.: Measuring sperm DNA fragmentation and clinical outcomes of medically assisted reproduction: a systematic review and meta-analysis. PLoS One. 2016, 11:e0165125. 10.1371/journal.pone.0165125
22. Salih M, Austin C, Warty RR, et al.: Embryo selection through artificial intelligence versus embryologists: a systematic review. Hum Reprod Open. 2023, 2023:hoad031. 10.1093/hropen/hoad031
23. Liu X, Chen Z, Ji Y: Construction of the machine learning-based live birth prediction models for the first in vitro fertilization pregnant women. BMC Pregnancy Childbirth. 2023, 23:476. 10.1186/s12884-023-05775-3
24. Tal R, Seifer DB: Ovarian reserve testing: a user's guide . Am J Obstet Gynecol. 2017, 217:129-40. 10.1016/j.ajog.2017.02.027
25. Qiu J, Li P, Dong M, Xin X, Tan J: Personalized prediction of live birth prior to the first in vitro fertilization treatment: a machine learning method. J Transl Med. 2019, 17:317. 10.1186/s12967-019-2062-5
26. Goyal A, Kuchana M, Ayyagari KP: Machine learning predicts live-birth occurrence before in-vitro fertilization treatment. Sci Rep. 2020, 10:20925. 10.1038/s41598-020-76928-z
27. Zhang Y, Shen L, Yin X, Chen W: Live-birth prediction of natural-cycle in vitro fertilization using 57,558 linked cycle records: a machine learning perspective. Front Endocrinol (Lausanne). 2022, 13:838087. 10.3389/fendo.2022.838087