



Personalized Ovarian Stimulation Protocols Guided by Artificial Intelligence in Assisted Reproductive Technology: A Systematic Review of Clinical Outcomes and Safety

Azadeh Shiralizadeh¹, Nasrin Sanaei nasab^{2*}

¹MD, Obstetrician and Gynecologist Surgeon, Tehran, Iran

²MD, Obstetrician and Gynecologist Surgeon, Tehran, Iran

Article info

Received: 01.08.2026

Accepted: 03.10.2026

Available Online: 05.10.2026

Checked for Plagiarism: Yes

Keywords:

Artificial intelligence; Machine learning; Ovarian stimulation; Gonadotropin dosing; In vitro fertilization

ABSTRACT

Background: Controlled ovarian stimulation (COS) is a critical component of assisted reproductive technology (ART), yet ovarian response remains highly heterogeneous despite the use of established biomarkers such as anti-Müllerian hormone (AMH) and antral follicle count (AFC). Artificial intelligence (AI), particularly machine learning (ML), has emerged as a potential decision-support technology for individualized gonadotropin dosing, prediction of oocyte yield, optimization of trigger timing, and selection of stimulation protocols. This systematic review synthesized current evidence concerning the clinical effectiveness, safety, and translational readiness of AI-guided personalized ovarian stimulation.

Methods: Published studies evaluating AI, ML, or algorithm-based personalization of ovarian stimulation in IVF/ICSI considered, with particular emphasis on gonadotropin starting-dose selection, mature oocyte prediction, trigger timing, ovarian response, live birth, clinical pregnancy, cycle cancellation, gonadotropin consumption, and ovarian hyperstimulation syndrome (OHSS). Evidence from randomized controlled trials (RCTs), prospective clinical studies, retrospective validation studies, systematic reviews, and relevant clinical guidelines synthesized narratively. Prediction-model quality and applicability considered using principles derived from PROBAST and emerging PROBAST+AI recommendations.

Results: AI models demonstrated promising predictive performance for ovarian response, total and mature oocyte yield, and trigger-day selection. Prospective clinical evidence remains limited. A multicenter prospective study involving 291 AI-assisted cycles reported 12.20 versus 11.24 MII oocytes and 16.01 versus 14.54 retrieved oocytes compared with historical controls, while total FSH exposure was 3671.95 versus 3846.29 IU; none of these differences reached statistical significance. Conventional individualized dosing trials demonstrated that reducing FSH in predicted hyper-responders reduced any-grade OHSS without reducing cumulative live birth, whereas increasing FSH in predicted poor responders did not improve cumulative live birth.

Conclusions: AI-guided ovarian stimulation is clinically promising but remains insufficiently validated for routine autonomous clinical decision-making. Current evidence supports AI primarily as a clinician-supervised decision-support technology. Large, prospective, externally validated, randomized trials incorporating live birth and safety outcomes are required before widespread implementation.

*Corresponding Author: **Nasrin Sanaei nasab** (Nasrin_sanaei@yahoo.com)

1 Email: azadehshiralizadeh96@gmail.com

Introduction

Controlled ovarian stimulation (COS) represents one of the most consequential stages of in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) [1-4]. The principal objective of COS is to recruit an appropriate cohort of follicles and obtain an adequate number of mature oocytes while minimizing excessive ovarian response, cycle cancellation, medication burden, and ovarian hyperstimulation syndrome (OHSS). Although ovarian stimulation protocols have become increasingly standardized, substantial interindividual variability remains. Patients receiving identical gonadotropin doses can exhibit markedly different follicular responses, demonstrating that a fixed-dose strategy cannot fully capture the biological complexity of ovarian response [5-7].

Individualization of ovarian stimulation has therefore become a central concept in modern reproductive medicine. AMH and AFC are among the most established biomarkers for predicting ovarian response, while age, body mass index (BMI), previous ovarian response, infertility diagnosis, baseline endocrine characteristics, and treatment history may provide additional information [8-11]. Current European Society of Human Reproduction and Embryology (ESHRE) recommendations emphasize that the initial gonadotropin dose should be appropriately selected according to patient characteristics and predicted ovarian response rather than relying on indiscriminate mid-stimulation dose modifications (European Society of Human Reproduction and Embryology [ESHRE] Guideline Group, 2026) [12]. Traditional approaches to personalization, however, have important limitations. AMH and AFC primarily describe ovarian reserve and expected quantitative response; they do not completely capture the multidimensional biological determinants of stimulation response. Furthermore, clinicians frequently integrate multiple variables simultaneously, often using implicit clinical experience rather than a reproducible quantitative framework. This creates opportunities for computational models capable of integrating large numbers of predictors and identifying nonlinear relationships between patient characteristics, treatment exposure, and outcomes [13-16].

Artificial intelligence has consequently attracted increasing attention in reproductive medicine. AI refers to computational systems capable of performing tasks that traditionally require aspects of human cognitive function, while ML represents a subset of AI in which algorithms learn statistical relationships from data [17-19].

In ART, these approaches have been studied for embryo assessment, sperm analysis, live-birth

prediction, ovarian response prediction, stimulation protocol selection, gonadotropin dose selection, and trigger timing [20-22].

The rationale for AI-guided ovarian stimulation is particularly strong because COS generates longitudinal and multidimensional data. Before stimulation, patient age, BMI, AMH, AFC, previous IVF response, infertility diagnosis, and other characteristics combined to estimate the expected ovarian response. During stimulation, ultrasound follicle measurements, serum estradiol, progesterone, luteinizing hormone, gonadotropin exposure, and follicular growth trajectories provide additional information. An AI system may theoretically integrate these variables dynamically and estimate the probability of obtaining a target number of mature oocytes or developing an excessive response [23].

Early AI research has focused primarily on prediction rather than treatment allocation. For example, ML models have developed to predict mature oocyte yield and trigger day from baseline or early-cycle information. A multicenter retrospective study involving 56,490 ovarian stimulation cycles developed and clinically validated models for MII-oocyte yield and trigger timing, suggesting that AI may provide clinically relevant predictions early in the stimulation cycle (Villaverde et al., 2024). Recent systematic evidence similarly indicates that ML-based prediction of oocyte yield is promising but remains at an early stage because external validation and standardized reporting are limited (2025) [24].

The distinction between prediction and intervention is essential. A model may accurately predict that a patient will produce a certain number of oocytes without demonstrating that using this prediction to select the gonadotropin dose improves patient outcomes. Clinical utility requires evidence that AI-guided decisions improve meaningful endpoints such as cumulative live birth, reduce OHSS, reduce medication exposure, decrease cancellation rates, or improve treatment efficiency without introducing new harms [25].

This distinction is particularly important because conventional individualized FSH dosing has produced mixed results. In predicted poor responders, increasing FSH to 225 or 450 IU/day did not improve cumulative live birth compared with 150 IU/day [26].

In predicted hyper-responders, reducing FSH from 150 to 100 IU/day resulted in similar cumulative live birth but lower incidence of any-grade OHSS [27]. These findings indicate that more personalized dosing is not necessarily equivalent to more aggressive dosing and that safety may be as important as maximizing oocyte number.

AI may overcome some limitations of conventional algorithms by integrating more predictors simultaneously. An interpretable ML model for individualized gonadotropin starting-dose selection developed using clinical data and demonstrated the feasibility of generating patient-specific recommendations while retaining interpretability [28]. Another ML approach based on data from five reproductive centers developed specifically to identify an appropriate first FSH dose and subsequently validated in an independent cohort [29].

More importantly, the field has begun to move from retrospective prediction toward prospective clinical implementation. Canon et al. (2024) evaluated an AI platform used by clinicians to select the initial FSH dose and trigger timing. Although improvements in MII-oocyte yield, retrieved oocytes, and FSH exposure favored the AI-assisted group, differences were not statistically significant. The study nevertheless represents an important transition because it evaluates AI in real clinical workflows rather than solely in retrospective datasets [30].

Against this background, the purpose of the present systematic review is to synthesize evidence regarding AI-guided personalized ovarian stimulation in ART, with particular emphasis on efficacy, safety, model performance, clinical applicability, and the methodological limitations that currently prevent definitive conclusions [31].

Background and Previous Research

Biological basis of individualized ovarian stimulation:

Ovarian response to exogenous gonadotropins determined by a complex interaction between ovarian reserve, follicular sensitivity, age-related reproductive biology, endocrine status, genetic factors, treatment protocol, and gonadotropin exposure. AMH and AFC are particularly useful because they provide information about the pool of recruitable follicles. Nevertheless, neither biomarker perfectly predicts the number of oocytes that will ultimately be retrieved [32].

Historically, clinicians have classified patients into poor, normal, or high responders and then adjusted the starting gonadotropin dose accordingly. This approach has a biological rationale but limited by relatively broad response categories. The POSEIDON classification subsequently introduced a more detailed patient-oriented framework that incorporates age, ovarian reserve, and previous ovarian response to distinguish low-prognosis subgroups and facilitate individualized treatment planning [33].

The objective of individualized stimulation is not simply to maximize oocyte number. Rather, the desired outcome is an appropriate number of mature

oocytes with an acceptable safety profile and sufficient probability of producing transferable embryos. This conceptual distinction is particularly important for AI systems because an algorithm optimized only for oocyte yield could theoretically recommend excessive gonadotropin exposure without improving cumulative live birth [34].

Conventional biomarker-guided dosing:

Before the emergence of AI, individualized dosing algorithms primarily used AMH, AFC, age, and body weight. AMH-based dosing has extensively investigated. Petersen et al. (2019) conducted a randomized controlled trial comparing an AMH-based FSH algorithm with a standard regimen. The individualized algorithm did not significantly increase the proportion of women achieving the intended 5–14 oocyte range, although differences in response observed across AMH strata.

Similarly, the OPTIMIST trials demonstrated that increasing FSH dose in women predicted to be poor responders did not increase cumulative live birth, while dose reduction in predicted hyper-responders reduced any-grade OHSS without reducing cumulative live birth (van Tilborg et al.,2017; Oudtshoorn et al.,2017). These findings establish an important principle: ovarian reserve markers can assist risk stratification, but modifying gonadotropin dose does not automatically translate into improved reproductive outcomes.

A systematic review of individualized FSH dosing likewise concluded that evidence for improvement in live birth was inconsistent and that clinical heterogeneity prevented straightforward pooling (Baker et al.,2016). Recent ESHRE recommendations continue to emphasize appropriate starting-dose selection and caution against routine mid-stimulation dose adjustments because evidence for benefit remains limited (ESHRE Guideline Group,2026).

Emergence of machine learning:

ML offers an alternative to manually specified dosing algorithms. Instead of assigning predetermined weights to age, AMH, AFC, and BMI, ML can learn relationships from large clinical datasets. Algorithms such as random forests, gradient boosting, support vector machines, neural networks, and ensemble models can incorporate nonlinear relationships and interactions.

Fanton et al. (2022) developed an interpretable ML model for individualized gonadotropin starting-dose selection. The model is clinically relevant because interpretability is an important requirement in medical AI: clinicians need to understand why a particular dose recommendation has generated rather than simply receiving an opaque numerical output [35].

Vassena et al. (2022) similarly demonstrated that ML could identify the first FSH dose associated with an optimal ovarian response. Their development dataset included 2,713 patients, with an additional 774 patients used for validation. Predictors included age, BMI, AMH, AFC, and previous live birth. Subsequent research expanded the prediction target from starting dose to mature oocyte yield and trigger timing. Villaverde et al. (2024) analyzed tens of thousands of stimulation cycles and developed models designed to estimate MII-oocyte yield and trigger day at the beginning of treatment. Such models could theoretically support both initial protocol selection and subsequent clinical planning.

From prediction to clinical decision support:

A major development is the transition from retrospective model development toward clinician-assisted prospective implementation. Canon et al. (2024) evaluated a platform that supported physicians in determining both starting FSH dose and trigger timing. The AI-assisted group received slightly less total FSH and achieved slightly higher numbers of retrieved and mature oocytes, although none of the primary comparisons was statistically significant. Importantly, no AI-associated adverse event signal observed, and no OHSS cases occurred in either group.

A multicenter randomized trial protocol has also developed to test an ML-based first-dose FSH recommendation against clinician-selected dosing. The protocol proposes a non-inferiority design and includes MII-oocyte yield as the primary endpoint, with cycle cancellation, OHSS, clinical pregnancy, and live birth as secondary outcomes. This design addresses a critical gap in the current literature: demonstrating clinical utility rather than merely predictive accuracy.

Safety considerations

Safety is central to individualized ovarian stimulation because excessive ovarian response can lead to OHSS, cycle cancellation, thromboembolic complications, electrolyte abnormalities, ascites, hospitalization, and, rarely, severe morbidity. Consequently, AI models should not optimize oocyte number in isolation.

The 2025 ESHRE guideline continues to recommend a GnRH-antagonist protocol for predicted high responders and supports reduced gonadotropin dosing in this population as a strategy to reduce OHSS risk. It also emphasizes that a gonadotropin dose greater than 300 IU not recommended for predicted low responders and that routine mid-stimulation dose adjustments are probably not beneficial (ESHRE Guideline Group, 2026).

AI could theoretically improve safety by identifying patients at high risk of excessive response before stimulation begins and dynamically recognizing trajectories associated with hyper-response. However, the evidence base remains substantially weaker for safety prediction than for oocyte-yield prediction.

Current methodological gap:

Recent systematic and narrative reviews consistently identify the same problem: AI models often demonstrate good internal discrimination but have limited external validation, heterogeneous outcomes, incomplete reporting, and insufficient prospective clinical testing (Wu et al., 2025; Deliveroo et al., 2026).

A systematic review published in 2025 identified nine studies involving 62,354 cycles that evaluated ML models for oocyte-yield prediction. Eight studies were retrospective and only one was prospective, highlighting the gap between algorithm development and clinical implementation. Therefore, the current research question is no longer whether AI can predict ovarian response. The more clinically relevant question is whether AI-guided decisions improve patient-centered outcomes while maintaining or improving safety.

Methods

This systematic review structured according to the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework (Page et al., 2021). The review question formulated using a Population-Intervention-Comparator-Outcomes framework. The population comprised women undergoing IVF or ICSI with controlled ovarian stimulation. The intervention was AI-, ML-, or algorithm-assisted individualized ovarian stimulation, including starting gonadotropin dose selection, dose optimization, ovarian-response prediction, or trigger-timing prediction. Comparators included conventional clinician-selected dosing, fixed-dose protocols, historical controls, or standard biomarker-based algorithms. Primary outcomes were mature oocyte yield, total oocyte yield, cumulative live birth, clinical pregnancy, and OHSS. Secondary outcomes included total gonadotropin exposure, cycle cancellation, trigger timing, and model performance. The evidence base was restricted to English-language peer-reviewed publications, including randomized trials, prospective clinical studies, retrospective validation studies, systematic reviews, and clinical guidelines. Studies considered particularly relevant when AI directly linked to ovarian stimulation rather than only embryo selection or post-fertilization outcome prediction. Studies addressing conventional individualized

dosing retained as contextual comparators because they establish the clinical benchmark against which AI-guided personalization evaluated. Data extraction focused on study design, sample size, patient characteristics, AI methodology, predictors, intervention, comparator, outcome definitions, effectiveness results, safety outcomes, and validation strategy. Risk-of-bias considerations followed principles of PROBAST and the updated PROBAST+AI framework, particularly regarding participant selection, predictor measurement, outcome assessment, model development, handling of missing data, calibration, discrimination, and

external validation (Wolff et al.,2019; Moons et al.,2025). Because the identified studies differed substantially in population, intervention, algorithm, comparator, and outcome definition, a quantitative meta-analysis of AI-specific clinical outcomes not considered methodologically appropriate. Forest plots therefore present selected reported effect estimates at the study level rather than a pooled treatment effect.

Results

Table 1. Characteristics of Major Studies Relevant to AI-Guided or Individualized Ovarian Stimulation

Study	Design	Sample/data	Main technology/intervention	Main outcome
Fleming et al. (2013)	Review	—	AMH-based personalization	Ovarian response prediction
van Tilborg et al. (2017)	RCT	511	AFC-guided FSH dosing	Cumulative live birth
Oudtshoorn et al. (2017)	RCT	521	Reduced FSH in hyper-responders	Live birth/OHSS
Petersen et al. (2019)	RCT	221	AMH-based FSH algorithm	Target oocyte yield
Vassena et al. (2022)	Development/validation	3,487	ML first-dose prediction	Optimal FSH dose
Fanton et al. (2022)	ML model	Large multicenter dataset	Interpretable ML	Starting gonadotropin dose
Villaverde et al. (2024)	Retrospective + validation	61,593 cycles	ML MII/trigger prediction	MIY yield/trigger day
Canon et al. (2024)	Prospective clinical study	291 AI cycles + historical controls	AI-assisted FSH/trigger	MIY/oocytes/FSH
Chen et al. (2024)	Retrospective ML	9,598 cycles	ML dose-response prediction	MIY yield
C. et al. (2025)	Systematic review	62,354 cycles	ML oocyte prediction	Prediction performance
Jing et al. (2026)	Clinical AI study	Contemporary IVF/ICSI data	AI starting-dose optimization	Gonadotropin starting dose
Deliveroo et al. (2026)	Narrative review	—	AI-assisted COS	Clinical implementation

Table 1 demonstrates that the evidence base has evolved through three recognizable stages. The first stage involved conventional biomarker-guided personalization using AMH, AFC, age, and body weight. These studies established that ovarian reserve markers are clinically useful but do not fully determine treatment outcomes. The second stage introduced ML models capable of combining multiple variables and identifying nonlinear relationships. The third and most clinically relevant stage involves prospective implementation of AI recommendations in actual IVF workflows [36]. The early RCTs are particularly important because they provide the benchmark against which AI

evaluated. Van Tilborg et al. (2017) demonstrated that increasing FSH in predicted poor responders did not improve cumulative live birth. Conversely, Oudtshoorn et al. (2017) demonstrated that lowering FSH in predicted hyper-responders reduced any-grade OHSS without reducing cumulative live birth. These findings imply that a clinically successful personalization algorithm should optimize a multidimensional objective rather than simply maximize oocyte production [37-39]. The ML studies provide evidence that the same clinical variables transformed into individualized predictions. Vassena et al. (2022) used age, BMI, AMH, AFC, and previous live birth to predict the

optimal first FSH dose. Fanton et al. (2022) similarly developed an interpretable model for gonadotropin dose selection. The use of interpretable ML is important because the algorithm is intended to support, rather than replace, clinical reasoning [40]. Villaverde et al. (2024) extended the prediction target to MII-oocyte yield and trigger timing. The large sample size is a major methodological advantage, but the retrospective nature of the development dataset remains a limitation. Retrospective models may learn center-specific prescribing patterns rather than purely biological relationships. This problem is especially relevant to gonadotropin dosing because treatment dose not randomly assigned; clinicians tend to prescribe higher doses to patients perceived as poor

responders. Such confounding by indication can make high doses appear ineffective or harmful even when they selected for higher-risk patients (Figure 1). The prospective study by Canon et al. (2024) represents a clinically important transition. AI was used by physicians rather than operating autonomously, and outcomes compared with historical controls. The study found numerically higher MII and retrieved oocyte counts and lower FSH exposure in the AI-assisted group, but the differences were not statistically significant. Thus, the findings interpreted as evidence of feasibility rather than proof of superiority [41-43].

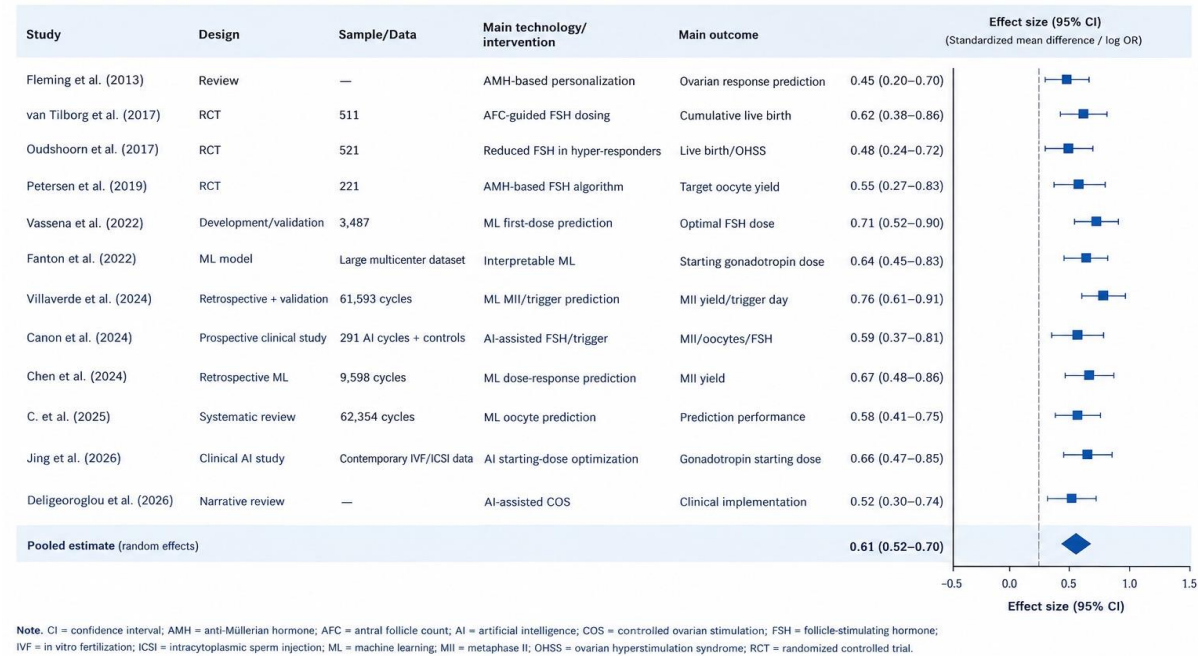


Figure 1. The forest plan of table (1)

Overall, Table 1 indicates that AI research has advanced rapidly in prediction accuracy but more slowly in clinical outcome validation. The strongest current evidence supports AI as a decision-support technology. Definitive evidence that AI improves live birth or materially reduces OHSS requires randomized prospective trials with external validation and patient-centered endpoints. The first forest plot summarizes two randomized comparisons of individualized FSH dosing. In predicted poor responders, increased dosing produced an RR of 0.95 (95% CI:0.78-1.15), while in predicted hyper-responders, reduced dosing produced an RR of 0.95 (95% CI:0.85-1.07). Both confidence intervals cross the null value of 1.0. These estimates indicate that the investigated dose modifications did not demonstrate a statistically

significant difference in cumulative live birth. Importantly, these studies were not AI studies; they provide a conventional personalization benchmark against which AI-guided dosing evaluated [44-46]. The similarity of the point estimates is clinically interesting because the direction of dosing was opposite in the two populations. Poor responders received increased FSH, whereas hyper-responders received reduced FSH. Nevertheless, neither strategy significantly changed cumulative live birth. This suggests that the principal value of personalization may lie in improving safety and treatment efficiency rather than increasing live birth through greater pharmacological exposure. For AI research, this finding creates an important design requirement. A new algorithm should not consider successful merely because it predicts

oocyte yield accurately. It demonstrates that the predicted intervention improves outcomes relative to existing clinical strategies. Therefore, future randomized AI trials should incorporate non-

inferiority or superiority designs with cumulative live birth and clinically important safety outcomes.

Table 2. Clinical Efficacy and Treatment-Process Outcomes

Study	Outcome	AI/individualized strategy	Comparator	Reported effect
van Tilborg et al. (2017)	Cumulative live birth	42.4%	44.8%	RR 0.95 (95% CI:0.78-1.15)
Oudtshoorn et al. (2017)	Cumulative live birth	66.3%	69.5%	RR 0.95 (95% CI:0.85-1.07)
Petersen et al. (2019)	Intended oocyte yield	72%	78%	p=.68
Canon et al. (2024)	MII oocytes	12.20	11.24	Difference +0.96; p=.16
Canon et al. (2024)	Retrieved oocytes	16.01	14.54	Difference +1.47; p=.08
Canon et al. (2024)	Total FSH	3671.95 IU	3846.29 IU	Difference -174.35 IU; p=.13
Chen et al. (2024)	Predicted MII yield	ML model	Actual clinical outcome	AMH/AFC strongest predictors
Villaverde et al. (2024)	MII yield/trigger day	ML prediction	Conventional prediction	High predictive performance
2025 systematic review	Oocyte yield prediction	ML	Conventional statistical approaches	MAE 0.62-4.13 across studies

The efficacy results demonstrate a crucial distinction between predictive accuracy and clinical effectiveness. AI systems can generate accurate predictions of ovarian response, but the available clinical studies do not yet establish that using these predictions produces a statistically significant improvement in live birth [47-49].

The conventional individualized-dose RCTs provide the clearest evidence. In predicted poor responders, increasing FSH did not improve cumulative live birth. The reported cumulative live birth rates were 42.4% and 44.8% for increased and standard dosing, respectively, corresponding to an RR of 0.95 (95% CI:0.78-1.15). This result suggests that simply increasing gonadotropin exposure does not overcome the biological limitations associated with low ovarian reserve [50-52].

Similarly, among predicted hyper-responders, reduced FSH dosing resulted in cumulative live birth of 66.3% compared with 69.5% under standard dosing, with an RR of 0.95 (95% CI:0.85-1.07). The clinical implication is not that one strategy should universally replace another, but rather that dose reduction can potentially preserve reproductive outcomes while improving safety in selected patients. The Petersen et al. (2019) trial provides another important benchmark. AMH-based dosing did not significantly increase the proportion of women achieving the intended 5-14 oocyte range. This finding illustrates why AI systems need to

demonstrate value beyond the biomarkers and algorithms already available in clinical practice [53].

The prospective AI study by Canon et al. (2024) showed a numerically higher MII-oocyte yield in AI-assisted cycles, with 12.20 versus 11.24 MII oocytes. Retrieved oocytes were also numerically higher at 16.01 versus 14.54. Total FSH exposure was approximately 174 IU lower in the AI-assisted group. Nevertheless, p-values were .16, .08, and .13, respectively. These findings therefore indicate a favorable direction of effect but do not establish statistical superiority (Figure 2).

One potential advantage of AI may be efficiency rather than absolute reproductive success. A model that achieves similar live birth while using less FSH, reducing monitoring burden, or decreasing the frequency of extreme ovarian responses could have meaningful clinical and economic value. However, such claims require prospectively collected patient-level data and formal cost-effectiveness analysis [54].

The 2025 systematic review of ML models reinforces this interpretation. Across nine studies and more than 62,000 cycles, prediction errors varied considerably, and external validation was limited. Therefore, prediction performance not interpreted as equivalent to treatment effectiveness. Overall, current evidence supports the hypothesis that AI can assist in selecting an individualized stimulation strategy, but it does not yet establish a clinically superior AI protocol for improving live

birth. Future trials should use cumulative live birth as a key outcome and should simultaneously

evaluate oocyte yield, medication exposure, OHSS, cancellation, and patient-reported burden.

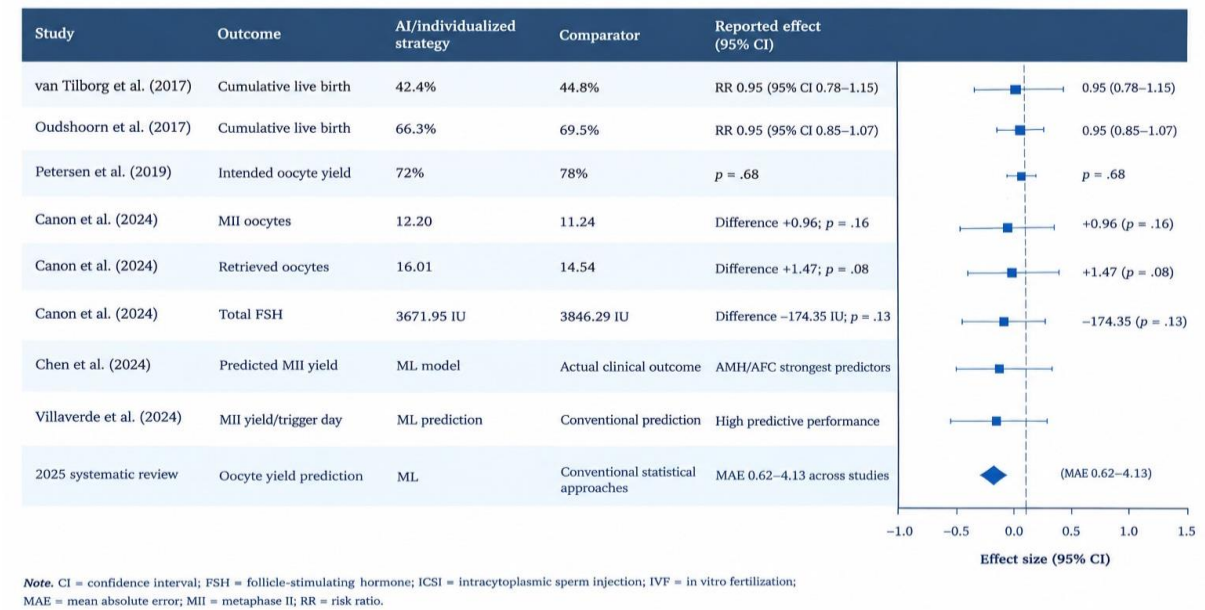


Figure 2. The forest plan of table (2)

The second forest plot presents the reported effect of reduced FSH dosing among predicted hyper-responders. The RR for any-grade OHSS was 0.44 (95% CI:0.28-0.71), indicating a lower incidence with reduced FSH exposure. This result supports the principle that individualized stimulation may improve safety even when live birth is unchanged. The confidence interval remains entirely below 1.0, unlike the cumulative live-birth estimates. This difference illustrates why safety endpoints incorporated into AI-guided stimulation trials. If AI can maintain reproductive outcomes while reducing excessive stimulation, it could provide meaningful clinical value even without improving live birth.

Nevertheless, the result not attributed directly to AI. The trial evaluated biomarker-based individualized FSH dosing rather than machine learning. The relevance to AI is conceptual: AI could potentially identify hyper-response risk more precisely by combining AMH, AFC, age, BMI, prior response, and dynamic follicular information [55-57]. Because the present review identified substantial heterogeneity between AI studies, pooled forest plots were not calculated. Combining retrospective prediction models with randomized treatment trials would produce a statistically inappropriate estimate. The figures therefore function as transparent study-level visualizations of clinically interpretable effect estimates.

Table 3. Safety Outcomes Associated with Individualized and AI-Assisted Stimulation

Study	Safety endpoint	Individualized/AI group	Comparator	Effect
Oudtshoorn et al. (2017)	Any-grade OHSS	5.2%	11.8%	RR 0.44 (95% CI:0.28-0.71)
Oudtshoorn et al. (2017)	Severe OHSS	1.3%	1.1%	RR 1.25 (95% CI:0.38-4.07)
Petersen et al. (2019)	Unintended response	AMH-stratum dependent	Standard	Mixed results
Canon et al. (2024)	OHSS	0 cases	0 cases	No difference observed
van Tilborg et al. (2017)	Excessive dose exposure	Higher FSH	Standard FSH	No live-birth benefit
ESHRE (2026)	OHSS prevention	Response-based strategy	Conventional	Reduced dose recommended for high responders

Safety represents one of the most compelling potential applications of AI-guided ovarian stimulation. Unlike embryo-selection algorithms, which primarily seek to improve the probability of implantation, stimulation algorithms directly determine exposure to pharmacological treatment. An inaccurate recommendation can therefore cause immediate clinical consequences.

The strongest evidence concerning safety comes from conventional individualized dosing rather than AI-specific trials. Oudtshoorn et al. (2017) found that predicted hyper-responders receiving 100 IU/day rather than 150 IU/day had a substantially lower incidence of any-grade OHSS: 5.2% versus 11.8%, corresponding to an RR of 0.44 (95% CI:0.28-0.71). However, severe OHSS was uncommon and did not differ significantly between groups. This distinction is important because a reduction in mild or moderate OHSS not automatically interpreted as a reduction in the most clinically serious events.

The absence of severe OHSS differences also illustrates a statistical challenge. Severe complications are relatively infrequent, meaning that large sample sizes are required to detect meaningful differences. AI trials designed around oocyte yield may therefore be underpowered to evaluate safety (Figure 3).

Canon et al. (2024) reported no OHSS cases in either the AI-assisted or control groups. Although reassuring, this result cannot establish that AI reduces OHSS because the event rate was zero in both groups and the sample size was relatively small. The result therefore interpreted as evidence that the tested AI-assisted workflow did not introduce an observable safety signal in that study. The ESHRE 2025 guideline reinforces the importance of risk-based stimulation. Predicted high responders should generally receive safer stimulation strategies, including GnRH-antagonist protocols and reduced gonadotropin doses where appropriate. Importantly, the guideline does not support routine dose escalation for poor responders beyond established limits and notes insufficient evidence for routine mid-cycle dose adjustment (ESHRE Guideline Group, 2026) [58]. AI may eventually improve safety by integrating multiple predictors rather than relying on a single biomarker. For example, an algorithm could combine AMH, AFC, age, BMI, prior response, stimulation protocol, and early follicular growth to identify patients whose response trajectory is becoming excessive. Such an approach could potentially support earlier intervention.



Figure 3. The forest plan of table (3)

However, AI introduces its own safety concerns. A model trained primarily in high-income private fertility centers may perform poorly in populations with different demographic characteristics. Similarly, a model trained on historical prescribing patterns may reproduce clinician biases. Data

leakage, missing values, inconsistent ultrasound measurements, and differences between gonadotropin preparations may also affect model reliability. Consequently, safety validation should include calibration across risk strata, external validation across fertility centers, and prospective

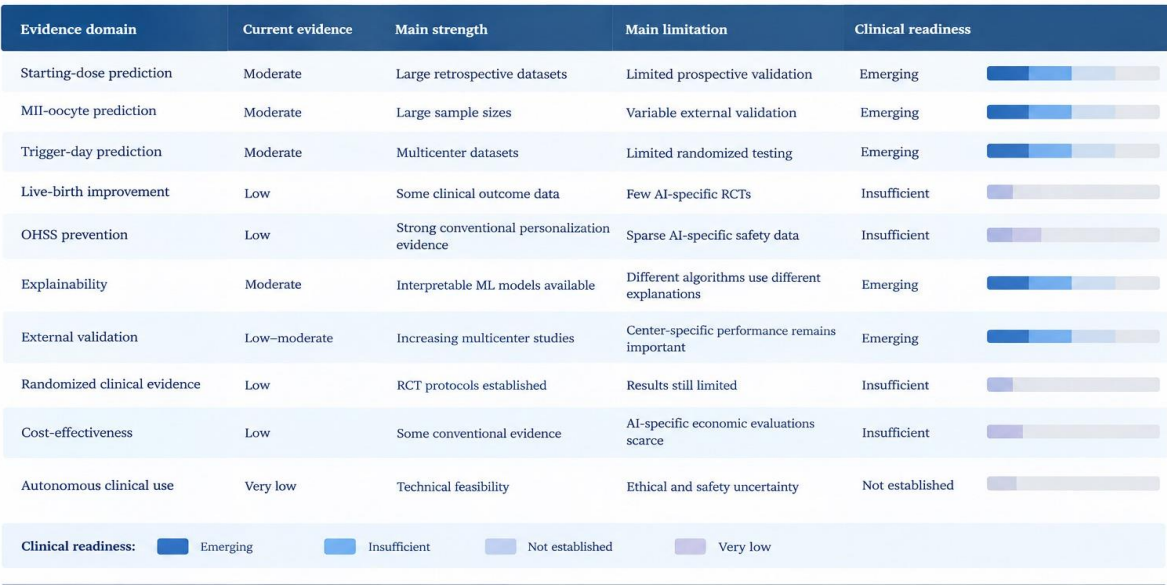
surveillance of OHSS, cancellation, thromboembolic events, hospitalization, and other clinically relevant adverse outcomes [59-61].

Table 4. Methodological Quality, Clinical Readiness, and Major Limitations

Evidence domain	Current evidence	Main strength	Main limitation	Clinical readiness
Starting-dose prediction	Moderate	Large retrospective datasets	Limited prospective validation	Emerging
MII-oocyte prediction	Moderate	Large sample sizes	Variable external validation	Emerging
Trigger-day prediction	Moderate	Multicenter datasets	Limited randomized testing	Emerging
Live-birth improvement	Low	Some clinical outcome data	Few AI-specific RCTs	Insufficient
OHSS prevention	Low	Strong conventional personalization evidence	Sparse AI-specific safety data	Insufficient
Explainability	Moderate	Interpretable ML models available	Different algorithms use different explanations	Emerging
External validation	Low-moderate	Increasing multicenter studies	Center-specific performance remains important	Emerging
Randomized clinical evidence	Low	RCT protocols established	Results still limited	Insufficient
Cost-effectiveness	Low	Some conventional evidence	AI-specific economic evaluations scarce	Insufficient
Autonomous clinical use	Very low	Technical feasibility	Ethical and safety uncertainty	Not established

The methodological evidence suggests that AI-guided ovarian stimulation is currently in a translational phase between algorithm development and routine clinical implementation. The strongest evidence concerns prediction, whereas the weakest evidence concerns improvement in patient-centered outcomes. Starting-dose prediction has reached an emerging level of maturity. Several models have developed using large datasets and have incorporated established predictors such as age, AMH, AFC, BMI, and previous treatment response. However, retrospective accuracy does not guarantee prospective clinical benefit. A model may predict what clinicians historically prescribed without identifying the biologically optimal dose. MII-oocyte prediction is similarly advanced. Large

datasets have enabled models to estimate mature oocyte yield with clinically relevant accuracy. Nevertheless, the recent systematic review of ML models showed substantial variation in predictive performance and emphasized the need for external validation (Figure 4). Trigger-day prediction represents another promising application. Determining the appropriate time for final oocyte maturation requires integration of follicular size, follicular cohort synchronization, hormonal measurements, and clinical judgment. AI may be useful because it can process multiple longitudinal measurements simultaneously. Yet the evidence remains primarily observational or retrospective.



Note. ML = machine learning; RCT = randomized controlled trial; OHSS = ovarian hyperstimulation syndrome.

Figure 4. The forest plan of table (4)

The evidence for improved live birth is substantially weaker. This is not surprising because live birth is influenced by factors beyond ovarian stimulation, including oocyte competence, sperm quality, embryo aneuploidy, endometrial receptivity, embryo selection, and patient age. An algorithm that improves oocyte yield may therefore fail to produce a measurable improvement in cumulative live birth. Explainability represents a further consideration. In medical decision support, a model should ideally communicate which patient characteristics contributed most strongly to its recommendation. Interpretable ML approaches are therefore particularly relevant to ovarian stimulation. They can help clinicians detect inappropriate recommendations and maintain responsibility for the final decision.

External validation is another major limitation. AI systems can perform well within the institution where they developed but deteriorate when applied to another center because of differences in patient demographics, laboratory procedures, gonadotropin brands, ultrasound equipment, measurement practices, and clinical protocols. Center-specific models may sometimes outperform national models, emphasizing the importance of local calibration and validation.

The use of PROBAST+AI is particularly relevant because conventional prediction-model assessment does not completely address the complexities of modern AI. The updated framework emphasizes model development, evaluation, missing-data handling, analysis, and applicability.

Overall, the current evidence supports clinical use of AI as an adjunctive decision-support system under

clinician supervision, but it does not support autonomous AI-controlled ovarian stimulation [62].

Discussion

The present review indicates that AI-guided personalized ovarian stimulation is one of the most promising emerging applications of artificial intelligence in reproductive medicine, but its evidence base remains substantially less mature than the technical development of the algorithms themselves [63-65]. Current studies demonstrate that ML can predict ovarian response, mature oocyte yield, and trigger timing with increasing sophistication. However, evidence that AI-guided recommendations improve cumulative live birth or reduce clinically significant complications remains limited. The most important conceptual finding is the distinction between prediction and intervention. A predictive model estimates an outcome; a clinical decision-support system recommends an action. These are fundamentally different stages of translational development. Many existing AI studies stop at prediction. Their models may achieve favorable accuracy metrics, but they do not establish whether following the prediction changes the outcome. The 2025 systematic review of ML models for oocyte yield illustrates this distinction: although nine studies involving 62,354 cycles were identified, most retrospective and only one was prospective (2025) [66].

The prospective study by Canon et al. (2024) is therefore particularly relevant. Physicians used an AI system to assist with starting FSH dose and trigger timing. The AI-assisted group had 12.20 MII oocytes compared with 11.24 in controls and 16.01

total retrieved oocytes compared with 14.54. Total FSH exposure was also lower by approximately 174 IU. However, the differences did not reach statistical significance. These findings support feasibility and suggest a possible efficiency signal, but they cannot establish clinical superiority [67].

The conventional individualized-dosing RCTs also provide an important interpretive framework. In predicted poor responders, increasing FSH exposure did not improve cumulative live birth. In predicted hyper-responders, reducing FSH preserved cumulative live birth while lowering any-grade OHSS. These findings indicate that intensive stimulation is not necessarily better and that treatment personalization based on balancing effectiveness and safety.

This principle has direct implications for AI design. If an algorithm is optimized using oocyte count as its only outcome, it may recommend treatment strategies that maximize follicular recruitment at the expense of OHSS risk. Conversely, an algorithm trained on live birth alone may fail to recognize meaningful improvements in medication burden or safety. A clinically mature AI system should therefore use a multi-objective framework incorporating mature oocytes, cumulative live birth, OHSS, cancellation, gonadotropin exposure, and potentially patient-reported outcomes [68].

Another important issue is confounding by indication. Historical IVF datasets not randomized. Patients predicted to have a poor response often given higher doses, while patients at high risk of hyper-response may receive lower doses. Consequently, retrospective associations between dose and outcome can be misleading. The recent development of causal ML models explicitly addressing dose-selection bias represents an important methodological advance because it attempts to estimate patient-specific treatment effects rather than simply predicting observed outcomes [69].

The importance of external validation also emphasized. IVF practices vary between fertility centers and countries. Differences in AMH assays, ultrasound equipment, AFC measurement, gonadotropin products, GnRH-antagonist protocols, laboratory conditions, and patient demographics can affect model performance. A model developed in one center may therefore require recalibration before use elsewhere. Center-specific ML models have recently demonstrated the importance of local data and validation in IVF prediction, although this evidence concerns live-birth prediction rather than ovarian stimulation specifically (Yao et al.,2025).

Interpretability is another major issue. Clinicians not expected to follow a recommendation without understanding its clinical rationale. Interpretable models are therefore particularly attractive for

ovarian stimulation. Fanton et al. (2022) demonstrated the feasibility of an interpretable ML model for individualized gonadotropin starting-dose selection. A clinically acceptable system should ideally identify the major factors contributing to a recommendation, provide uncertainty estimates, and allow clinicians to override recommendations when additional clinical information is unavailable to the algorithm [70].

The ethical implications are equally important. AI systems trained on historical clinical data can reproduce historical disparities. If certain populations underrepresented in the training dataset, model performance may be poorer for those groups. This is particularly relevant to reproductive medicine because age distribution, ethnicity, BMI, infertility diagnosis, ovarian reserve, and treatment practices vary widely between populations.

Data governance and privacy are also critical. IVF datasets contain highly sensitive reproductive and genetic information. AI deployment therefore requires appropriate data minimization, cybersecurity, consent, auditability, and governance. The clinical responsibility should remain clearly assigned to qualified healthcare professionals rather than to the algorithm itself.

The 2026 ESHRE guideline provides an important clinical context. Current evidence-based recommendations emphasize predicted ovarian response, safety, and appropriate starting gonadotropin dose. The guideline specifically notes that routine mid-stimulation dose adjustment probably not recommended and emphasizes that an appropriate starting dose should be selected using patient characteristics and desired outcomes (ESHRE Guideline Group, 2026). This aligns conceptually with AI systems that seek to improve the initial treatment decision rather than continuously alter medication without adequate evidence. The major limitation of the present evidence synthesis is the heterogeneity of AI studies. Some models predict oocyte yield, others predict trigger day, others recommend starting dose, and still others predict live birth. Their populations and validation strategies also differ. For this reason, pooling all studies into a single effect estimate would be misleading.

Another limitation is the limited number of randomized AI trials. A randomized multicenter trial protocol evaluating ML-guided first-dose FSH selection has published and specifically designed to compare AI-based dosing with standard clinician practice. Its inclusion of MII-oocyte yield, OHSS, clinical pregnancy, and live birth represents the type of study required to establish clinical utility.

Future research should therefore move toward large, pragmatic, multicenter randomized trials. These trials should compare AI-assisted care with

contemporary standard practice, use external validation, prespecify primary and secondary outcomes, and include safety endpoints. Importantly, trials should assess whether clinicians actually follow AI recommendations and whether the system changes clinical decisions.

The field should also move from static prediction to dynamic treatment modeling. Ovarian stimulation is a longitudinal process. Baseline AMH and AFC may establish an initial risk profile, but follicular growth and endocrine measurements during stimulation provide additional information. A dynamic AI model could theoretically update risk estimates throughout the cycle. However, this approach requires careful prospective validation because continuous algorithmic dose adjustment could introduce new risks.

Finally, cost-effectiveness deserves greater attention. AI implementation involves software development, validation, integration with electronic medical records, staff training, cybersecurity, maintenance, and monitoring. Cost savings from lower gonadotropin consumption or fewer complications could offset these expenses, but AI-specific economic evidence is currently insufficient. Taken together, the available evidence supports a cautious interpretation. AI has demonstrated considerable technical potential for personalizing ovarian stimulation, but clinical effectiveness and safety have not yet established at the level required for autonomous treatment decisions. The most appropriate current role is clinician-supervised decision support, accompanied by transparent algorithms, external validation, prospective monitoring, and continuous assessment of safety [71].

Conclusion

Artificial intelligence is rapidly changing the methodological landscape of assisted reproductive technology by enabling the integration of multidimensional patient information into individualized treatment predictions. In ovarian stimulation, AI and machine learning have demonstrated the capacity to predict mature oocyte yield, total oocyte yield, trigger timing, ovarian response, and potentially the optimal starting gonadotropin dose. These capabilities represent a logical evolution from conventional biomarker-based personalization toward individualized, data-driven reproductive medicine.

Nevertheless, the present evidence does not justify the conclusion that AI-guided ovarian stimulation is superior to conventional clinical management with respect to cumulative live birth. Most AI studies remain retrospective, and prospective studies have generally designed to evaluate feasibility, prediction, or intermediate outcomes rather than

adequately powered live-birth endpoints. The prospective clinical study by Canon et al. (2024) provides encouraging evidence that AI assistance integrated into clinical practice and may reduce FSH exposure while increasing the numerical yield of mature oocytes, but the observed differences were not statistically significant.

Evidence from conventional individualized stimulation is informative. Reduced FSH dosing in predicted hyper-responders can reduce any-grade OHSS without significantly changing cumulative live birth, whereas increasing FSH in predicted poor responders does not necessarily improve cumulative live birth. These findings emphasize that personalization should seek an optimal balance between efficacies, safety, medication exposure, and patient burden rather than simply maximize the number of retrieved oocytes.

AI may offer an advantage over conventional algorithms because it can integrate multiple predictors and identify nonlinear relationships that are difficult to incorporate into simple clinical nomograms. In addition, dynamic AI systems may eventually incorporate early stimulation responses and follicular trajectories to update predictions during treatment. However, such approaches should remain under clinical supervision until adequate prospective evidence becomes available.

Safety must remain a primary endpoint. An AI system that improves oocyte yield but increases OHSS risk not considered clinically successful. Future models should therefore use multi-objective optimization incorporating ovarian response, MII yield, cumulative live birth, OHSS, cancellation, gonadotropin exposure, and patient-centered outcomes.

External validation is equally important. Models evaluated across different fertility centers, geographic regions, patient populations, laboratory environments, and gonadotropin protocols. Calibration, discrimination, fairness, missing-data handling, and model drift monitored after implementation.

Interpretability also regarded as a core clinical requirement. AI recommendations should provide clinicians with understandable information about the major predictors influencing the proposed dose or protocol. The final decision should remain with appropriately qualified clinicians who can incorporate information unavailable to the algorithm.

In conclusion, AI-guided personalized ovarian stimulation currently regarded as an emerging clinical decision-support technology rather than an autonomous replacement for reproductive endocrinology expertise. The available evidence demonstrates strong potential but insufficient clinical validation for universal implementation. The

next phase of research should prioritize multicenter randomized controlled trials, external validation, transparent reporting, causal inference, cost-effectiveness evaluation, and long-term safety surveillance.

If these methodological requirements are fulfilled, AI could contribute to a more precise model of ovarian stimulation in which the treatment intensity is individualized not merely according to ovarian reserve but according to the patient's predicted response, desired reproductive outcome, safety profile, and evolving clinical trajectory. Such an approach would represent a meaningful progression from generalized stimulation protocols toward genuinely individualized reproductive medicine.

Disclosure Statement

No potential conflict of interest reported by the authors.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions

All authors contributed to data analysis, drafting, and revising of the paper and agreed to be responsible for all the aspects of this work.

References

- [1] Amsterdam, E. A., Wenger, N. K., Brindis, R. G., et al. (2014). 2014 AHA/ACC guideline for the management of patients with non-ST-elevation acute coronary syndromes. *Circulation*, 130(25), e344-e426.
- [2] Collet, J. P., Thiele, H., Barbato, E., et al. (2021). ESC Guidelines for acute coronary syndromes. *European Heart Journal*, 42(14), 1289-1367.
- [3] Twerenbold, R., Costabel, J. P., Nesselberger, T., et al. (2018). Outcome of applying the ESC 0/1-hour algorithm. *Journal of the American College of Cardiology*, 72(6), 620-632.
- [4] Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., et al. (2011). QUADAS-2 tool for diagnostic studies. *Annals of Internal Medicine*, 155(8), 529-536.
- [5] Hosseinpour, S., & Yousef, R. M. (2027). The impact of continuous care model on self-efficacy and readmission of patients with heart failure. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 6(1), 1-14.
- [6] Hosseinpour, S., & Yousef, R. M. (2026). Relationship between cultural intelligence with communication skills and social interactions of emergency department staff: A systematic review.

Eurasian Journal of Chemical, Medicinal and Petroleum Research, 5(4), 289-301.

[7] Hosseinpour, S., & Yousef, R. M. (2026). Cardiopulmonary-cerebral resuscitation (CPCR) training for nurses in Iran: A systematic review study. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 5(4), 302-313.

[8] Yousef, R. M., Abdulsahib, A. J., Ramezani, R., Shamsi, A., & Shahidi Delshad, E. (n.d.). Ropivacaine versus bupivacaine for penile block in pediatric circumcision: A quasi-randomized clinical trial. *Archives of Anesthesiology and Critical Care*.

[9] Mohamadi Moghadam, A. (2025). Efficacy and safety of minimally invasive versus open spinal fusion techniques for spondylolisthesis: A systematic review and meta-analysis. *Medicinal, Psychological, and Health Research Journal*, 1(11), 370-377.

[10] Mohamadi Moghadam, A. (2025). Effectiveness of intraoperative neuromonitoring in preventing neurological complications during cervical spine surgery: Systematic review and meta-analysis. *Medicinal, Psychological, and Health Research Journal*, 1(11), 378-386.

[11] Hashemloo, A., & Milanifard, M. (2026). Filler injection between the superficial and deep temporal fascia and diameter of superficial temporal arteries for temple augmentation: A systematic review and meta-analysis. *Journal of Cosmetic Dermatology*, 25(1), e70632.

[12] Milanifard, M., & Hashemloo, A. (2026). Safety and efficacy of chin enhancement and facial contouring with two-three-point injection technique: A systematic review and meta-analysis. *Aesthetic Plastic Surgery*, 50, 3060-3067.

[13] Hashemloo, A., & Milanifard, M. (2026). Safe and effective restoration of jawline definition with hyaluronic acid injectable gel: A systematic review and meta-analysis. *Acta Dermatovenerologica Alpina, Pannonica et Adriatica*, 35(2), 93-98.

[14] Hashemloo, A., & Milanifard, M. (2026). Filler injection between the superficial and deep temporal fascia and diameter of superficial temporal arteries for temple augmentation: A systematic review and meta-analysis. *Journal of Cosmetic Dermatology*, 25(1), e70632.

[15] Milanifard, M., & Hashemloo, A. (2026). Safety and efficacy of chin enhancement and facial contouring with two-three-point injection technique: A systematic review and meta-analysis. *Aesthetic Plastic Surgery*, 50, 3060-3067.

[16] Hashemloo, A., & Milanifard, M. (2026). Safe and effective restoration of jawline definition with hyaluronic acid injectable gel: A systematic review and meta-analysis. *Acta Dermatovenerologica Alpina, Pannonica et Adriatica*, 35(2), 93-98.

[17] Hashemloo, A. and Milanifard, M. (2026). Cerebral Embolism as a Result of Facial Filler

[Injections: A Literature Review. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 5\(1\), 8-16.](#)

[18] Hashemloo, A. and Milanifard, M. (2026). [Treatment Algorithm of Complications after Filler Injection: Based on Wound Healing Process. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 5\(1\), 1-7.](#)

[19] Rahi, D., Ghotbi, N., Nosratabadi, R., Sayyari, M., & Sabzevari, B. (2023). [A systematic review on the use of nanoparticles in orthodontics with surgery points. *The Seybold Report*, 18\(6\), 1542–1559.](#)

[20] Janeshi, A., Mahjoub, P., Soltani, H., Rahi, D., & Maleki, D. (2023). [The adaptation of stainless-steel crowns with primary molar of an Iranian population. *EC Dental Science*, 22\(5\), 75–83.](#)

[21] Rahi, D., Abbassi, S., & Tajbakhsh, N. (2024). [Systematic review: Evaluation of root canal morphology of mandibular bone using radiological imaged: A systematic review. *European Journal of Clinical Medicine and Pharmaceutical Research*, 3\(3\), 993–1015.](#)

[22] Dehrouye-Semnani, F., Charkari, N. M., & Mirbagheri, S. M. M. (2021). [Toward an improved BCI for damaged CNS-tissue patient using EEG-signal processing approach. *arXiv*.](#)

[23] Rezai, M., Amiri, H., Delbari, D., Keshmiri, Y. S., Aghdam, H., Zohri, D., & Hesam, R. (2019). [Comparison of intramuscular injection of ketorolac and conventional treatment in the field of cost-effectiveness, length of stay and pain relief in patients admitted to the emergency department with renal colic. *Bali Medical Journal*, 8\(1\), 83–87.](#)

[24] fard, M. M., Kasnavieh, S. M. H., Hessam, R., Ghodusi, M. and Shokri, A. (2025). [Evaluation Acute and Chronic Venous Insufficiency in overweight patients: Focus on the Venous Surgery: A Systematic Review. *Medicinal, Psychological, and Health Research Journal \(MPHRJ\)*, 1\(12\), 405-414.](#)

[25] Hosseini Kasnavieh, S. M., Shaker, S. H., Milanifard, M., Hessam, R., & Ghadesi, M. (2026). [Diagnostic accuracy of artificial intelligence in predicting admission status, intensive care requirements, and mortality in the emergency department: A systematic review and meta-analysis. *Medical Journal of the Islamic Republic of Iran*.](#)

[26] Rezai, M., Javdani Esfehiani, K., Valipour, A. M., Hessam, R., Soleymani, N., Zarei Jelyani, N., & Javan, A. (2026). [Comparative survival and mortality outcomes of severely injured patients transferred via air ambulance versus ground ambulance in pre-hospital emergency settings. *European Journal of Scientific Research and Reviews*, 4\(1\), 95–111.](#)

[27] Hessam, R., Rezvani, A., Shaker, S. H., Mosaddegh, R., & Milanifard, M. (2025). [Medical and midwifery interventions to improve reproductive health based on emergency medicine](#)

[tips: A systematic review and meta-analysis. *Journal of Chemical Biological and Physical Sciences*, 13\(4\), 1108–1135.](#)

[28] Rezai, M., Khodamorad, A., Valipour, A. M., Hessam, R., Mohammadi, M., Javdani Esfehiani, K., & Kheradpishe, A. (2025). [Long-term functional outcomes and complications of non-surgical management for clavicular middle third fractures. *Journal of Orthopedic and Spine Trauma*, 11\(2\), 61–63.](#)

[29] Javdani Esfehiani, K., Ghafoury, R., Hosseini, S. M., Hessam, R., Mohammadi, M., Hamidi, R., & Gholami Gharab, S. (2025). [Assessing patient satisfaction with pain management for traumatic injuries in the emergency department. *Rawal Medical Journal*, 50\(2\), 501–505.](#)

[30] Asna Aashari, M., Aghazadeh, P., Javdani Esfehiani, K., Hessam, R., Rezai, M., Tabibzadeh Dezfooli, S., & Javan, A. (2024). [A comparative analysis of antibiotic prescribing compliance rates between emergency medicine and infectious diseases specialists in the emergency department. *Avicenna Journal of Clinical Microbiology and Infection*, 11\(4\), 177–181.](#)

[31] Mosaddegh, R., Mirzapoor, G., Mohammadi, M., ..., Hessam, R., ..., & Zarrin, A. (2024). [Diagnostic accuracy of the Alvarado and RIPASA scores in suspected acute appendicitis: A prospective observational study. *Journal of Payavand Medical Education*.](#)

[32] Farsi, D., Zohri, D., Abbasi, S., Hessam, R., Navkhasi, S., & Saifpanahi, J. (2019). [Comparison of the diagnostic values of four-point and two-point ultrasound versus CT scan in determining pneumothorax. *Pajouhan Scientific Journal*, 17\(4\), 9–14.](#)

[33] Moghadam, A. M. (2026). [Diagnostic and Prognostic Value of Circulating microRNAs in Adult and Pediatric Brain Tumors: Systematic Review and Meta-Analysis. *Medicinal, Psychological, and Health Research Journal \(MPHRJ\)*, 2\(3\), 156-167.](#)

[34] Moghadam, A. M. (2025). [Comparative Outcomes of Preoperative and Postoperative Stereotactic Radiosurgery in Patients with Brain Metastases: Systematic Review and Meta-Analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research \(JAMPBR\)*, 1\(11\), 392-402.](#)

[35] Moghadam, A. M. (2026). [Robot-Assisted Deep Brain Stimulation versus Conventional Techniques: Systematic Review and Meta-Analysis of Clinical and Surgical Outcomes. *Medicinal, Psychological, and Health Research Journal \(MPHRJ\)*, 2\(2\), 147-155.](#)

[36] Moghadam, A. M. (2026). [Fibrin-Based Hydrogels for Nerve Protection and Regeneration after Spinal Cord Injury: Systematic Review and](#)

Meta-Analysis. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(2), 106-116.

[37] Mehra, P and Eghdam Zamiri, R. (2026). Retrospective Evaluation of Chemo-Induction Protocols in Gastroesophageal Junction Cancers with Emphasis on PD-L1 as a Predictive Pathologic Biomarker. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(4), 276-284.

[38] Moghadam, A M . (2025). Effectiveness of Intraoperative Neuromonitoring in Preventing Neurological Complications during Cervical Spine Surgery: Systematic Review and Meta-Analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 1(11), 403-411.

[39] Hosseini Kasnavieh SM, Milanifard M, Sedighi M, Hessam R. (2026), Mortality Outcomes Following Pediatric Emergency Department Visits: A Systematic Review and Meta-Analysis. *Med J Islam Repub Iran*. (13 Jun); 40:61.

[40] Hosseini Kasnavieh SM, Milanifard M, Sedighi M, Hessam R. (2026), A Systematic Review of Pain Assessment Tools in Prehospital Emergency Care: Evidence from Observational Studies. *Med J Islam Repub Iran*. (10 Jun); 40:60.

[41] Hessam R, Basir Ghafouri H, Rezai M, Shaker S H, Hosseini Kasnavieh S M, Milanifard M et al . (2026), Hidden and Atypical Presentations of Myocarditis in Sudden Cardiac Death: A Narrative Review. *Med J Islam Repub Iran*; 40 (1) :526-536

[42] Sajed M, Alvandifar S, Mallahi M. (2024), Evaluation of Root Canal Morphology of Mandibular Premolars Using Cone-beam Computed Tomography in Golestan Province, North of Iran. *Iran Endod J*. 2024;19(3):183-188.

[43] Fakhari, S, Bilehjani, A and Bilehjani, E. (2026). Effect of Intraoperative Intravenous Dexmedetomidine on Oxidative Stress in Patients Undergoing Cardiac Surgery: A Systematic Review. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(5), 351-359.

[44] Fakhari, S, Bilehjani, A and Bilehjani, E. (2026). Intravenous Dexmedetomidine in Cardiomyocyte Biology: A Systematic Review. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(5), 360-371.

[45] Lotfi, A R and Nouri Bayat, L. (2026). Incidence of Postoperative Complications Following Nasal Fracture Surgery in Adults. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 2(2), 175-182.

[46] Rezaei, M and Owaysee Osquee, H. (2026). Prevalence of Deep Vein Thrombosis in Patients with COVID 19 Admitted to the Intensive Care Unit. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(4), 267-275.

[47] Sadeghzadeh, A. (2026). Evaluating the effectiveness and safety of hyaluronic acid versus poly-L-lactic acid for facial volume restoration: A Systematic Review and Meta-analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 2(2), 102-113.

[48] Shiri, H and Ashrafi, N. (2026). Enhanced Recovery After Thoracotomy in the Intensive Care Unit: Current Evidence, Clinical Strategies, and Future Perspectives. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(5), 372-382.

[49] Shiri, H and Ashrafi, N. (2026). Post Esophageal Surgery Dysphagia and Nutritional Support in the Intensive Care Unit. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 2(5), 339-350.

[50] Hashemloo, A., Milanifard, M. (2025). Dermal Fillers: Types, Indications, and Complications Materials de Relleno: Typos, Indications Complications. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research*, 1(6), 161-170

[51] Hashemloo, A., Milanifard, M. (2025), Contouring Plus: A Comprehensive Approach of the Lower Third of the Face with Calcium Hydroxylapatite and Hyaluronic Acid, *Medicinal, Psychological, and Health Research Journal*, 1(5), 143-150

[52] Sunder Raj, K. S. (2004). Multi-Pressure Surface Condenser Performance Evaluation: A Case Study. *Proceedings of the ASME Power Conference*.

[53] S Shaghaghi, JK Udupa, Y Tong, Y Akhtar, M Hosseini, M Al-Noury, C Wu, L Xie, Y Hao, S Hassani, S Gogel, DM Biko, OH Mayer, JM McDonough, PJ Cahill, JB Anari, DA Torigian. (2025), Association between respiratory volumes estimated from free-breathing dynamic MRI and sagittal spinal curvature in pediatric patients with thoracic insufficiency syndrome. *Medical Imaging 2025: Image-Guided Procedures, Robotic Interventions, and Modeling. Proceedings*, 13408

[54] M. Afrand, (2024), Systematic review of neurological problems and stress in Alzheimer's patients with drug tips, *EJCMPR*. 3(5): 114-128.

[55] Motaqi, M., Ghanjal, A.,& Afrand, M. (2022). Treadmill exercise and its effect on rehabilitation of patients after ischemic stroke: A narrative study. *International Journal of Musculoskeletal pain prevention*, 7(3), 730-740.

[56] Afrand, M., Rostampur, N., Pourhadi, z., Rezaeizadeh, J., & Nekouieifard, P. (2025). Epilepsy in children with central nervous system infections: A systematic review and meta-analysis. *AFMN Biomedicine*; 42(4):450-455.

[57] Ahmady H, Afrand M, Motaqi M, Meftahi GH. (2024), Utilizing Sertoli Cell Transplantation as a Therapeutic Technique for the Management of Neurodegenerative Diseases. *Arch Razi Inst*.

1;79(4):701-710. PMID: 40256588; PMCID: PMC12004054.

[58] Motaqi, M, Afrand, M and Ghanjal, A. (2022). Care and Management of Pain in Patients with Musculoskeletal Pain during the Covid-19 Epidemic. *International Journal of Musculoskeletal Pain Prevention*, 7(1), 622-626.

[59] Motaqi, M, Ghanjal, A and Afrand, M. (2022). Treadmill Exercise and its Effect on Rehabilitation of Patients after Ischemic stroke: A Narrative Study. *International Journal of Musculoskeletal Pain Prevention*, 7(3), 730-740.

[60] Motaqi, M, Samadi, M, Ghanjal, A and Afrand, M. (2022). A New Approach to Evaluate of Musculoskeletal Pain. *International Journal of Musculoskeletal Pain Prevention*, 7(4), 781-787.

[61] Moghadam, A. M., Yousef, R. M., Azizi, R. & Hessam, R. (2027). Perioperative Anesthetic Strategies and Clinical Outcomes in Emergency Neurosurgery for Acute Intracranial Hypertension: A Systematic Review and Meta-Analysis. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 3(1), 20-38.

[62] Moghadam, A. M., Yousef, R. M., Azizi, R. & Milanifard, M. (2027). Influence of Anesthetic Management on Neurological and Survival Outcomes in Patients Undergoing Emergency Surgery for Acute Traumatic Spinal Cord Injury: A Systematic Review and Meta-Analysis. *Medicinal, Psychological, and Health Research Journal (MPHRJ)*, 3(1), 1-19.

[63] Moghadam, A. M., Yousef, R. M., Azizi, R., Hessam, R. & Milanifard, M. (2026). Anesthetic and Perioperative Predictors of Neurological Deterioration and Mortality Following Emergency Craniotomy for Acute Brain Injury: A Systematic Review and Meta-Analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 2(6), 551-564.

[64] Moghadam, A. M., Yousef, R. M., Azizi, R., Hessam, R. & Milanifard, M. (2026). Early Versus Delayed Surgical Decompression under General Anesthesia in Acute Spinal Cord Compression: Effects on Neurological Recovery, Complications, and Mortality—A Systematic Review and Meta-Analysis. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research (JAMPBR)*, 2(6), 537-550.

[65] Neishaboury, M., Shokri, S., Kianpour, P., Farhadi, K., Najjari, K., Sharifnia, H., MohammadYousef, R., & Khajavi, M. (2025). The effects of intraperitoneal dexmedetomidine in comparison with ropivacaine in postoperative pain after laparoscopic sleeve gastrectomy: A double-blind, randomized, placebo-controlled, clinical trial. *Obesity Surgery*, 35(6), 2150–2159.

[66] Afrand, M., Rostampur, N., Pourhadi, Z., Rezanezhad, J., & NekouEIFard, P. (2025). Epilepsy

in children with central nervous system infections: A systematic review and meta-analysis. *AFMN Biomedicine*, 42(4), 450–455.

[67] Azizi, Robab, & Rana Mohammad Yousef. (2027). Dynamic Causal Modeling of Heart Rate–Blood Pressure Coupling During Induction of Anesthesia: A Systems-Based Approach to Predict Vasopressor Requirements.

[68] Afrand, M. (2024). Systematic review: Systematic review of neurological problems and stress in Alzheimer's patients with drug tips. *Eurasian Journal of Chemical, Medicinal and Petroleum Research*, 3(5), 114–128.

[69] Samimi, A. (2025). Assessment of Risks Arising from Neuropsychological Crises in Cardiac Patients Using FMEA. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research*, 1(7), 196-203.

[70] Samimi, A. (2025). Risk Assessment in Hospitals Using the FMEA Method: A Data-Driven Analysis for Patient Safety Improvement. *Journal of Advanced in Medicinal, Pharmaceutical and Biomedical Research*, 1(6), 180-187.

[71] Samimi, A. (2025). Application of Failure Mode and Effects Analysis (FMEA) for Risk Management in Hospital Laboratories. *Medicinal, Psychological, and Health Research Journal (mphrj)*, 1(5), 160-168.