

Systematic Review

Safety of Pharmacological Agents Used in Assisted Reproductive Technology: A Comprehensive Systematic Review of Fetal Malformation Risk

Zeev Shoham¹, Ariel Weissman², Yuval Yaron³

¹ The Hebrew University of Jerusalem, Hadassah Medical School, Jerusalem 9112102, Israel, ² IVF Unit, Department of Obstetrics and Gynecology, Edith Wolfson Medical Center, Holon, Affiliated with the Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel, ³ Prenatal Genetic Diagnosis Unit, Tel Aviv Sourasky Medical Center, Israel

Keywords: assisted reproductive technology, in vitro fertilization, congenital malformations, teratogenicity, drug safety, systematic review

<https://doi.org/10.46989/001c.147524>

Journal of IVF-Worldwide

Vol. 4, Issue 1, 2026

Background

Assisted reproductive technology (ART) involves pharmacological agents that may raise fetal malformation concerns, yet their teratogenic potential remains uncertain despite widespread use, complicating clinical decisions and patient counseling.

Objective

To systematically assess fetal malformation risks from ART medications using evidence hierarchy principles for evidence-based safety evaluations.

Methods

PRISMA-compliant systematic review (PROSPERO CRD 420251118713) using narrative synthesis and selective random-effects meta-analysis for homogeneous studies, searching PubMed, Embase, Cochrane, and others (1990–2025). Included randomized controlled trials (RCTs), cohorts, case-controls, and reviews on ART-conceived pregnancies reporting major/system-specific malformations. Quality assessed via Cochrane Risk of Bias (RoB) 2.0/Newcastle-Ottawa Scale (NOS); certainty via Grading of Recommendations Assessment, Development and Evaluation (GRADE).

Results

32 studies (~1.2 million pregnancies) provided high-quality evidence for ART medication safety (pooled OR 1.01 [95% CI 0.92–1.11]; I²=20% vs. natural/alternatives). Gonadotropins (rFSH, hMG, rLH, hCG) showed no increased risk versus natural conception or alternative protocols (OR 1.01 [95% CI 0.92–1.11]; I²=18%; Level I-II from RCTs/meta-analyses). Gonadotropin-releasing hormone (GnRH) agonists/antagonists were comparable to each other (OR 1.03 [95% CI 0.89–1.19]; I²=12%). Progesterone routes (intramuscular (IM)/subcutaneous (SC)/vaginal) demonstrated equivalent safety across administration routes (OR 0.97 [95% CI 0.88–1.07]; I²=25%). Dydrogesterone demonstrated comparable safety to progesterone in RCTs (LOTUS I/II; OR 0.72 [95% CI 0.49–1.05]; I²=15%, no statistically significant difference), overriding lower-level pharmacovigilance/case-control signals. Adjuvants (metformin, letrozole, clomiphene) lacked teratogenicity versus natural conception or alternative treatments (OR 1.04 [95%

a Corresponding author

Zeev Shoham
The Hebrew University of Jerusalem, Hadassah Medical School
Ein Kerem Campus
Jerusalem 9112102, Israel
+972 52-253-0370
zeev1@mail.huji.ac.il

CI 0.90–1.20]; I²=35%), with varying certainty. Absolute major malformation risk: 2–6%, aligned with adjusted natural conception.

Conclusions

Strong evidence (high-quality GRADE) supports ART medication safety. For counseling, prioritize RCTs over observational data. Absolute risks are low, but continue surveillance to monitor long-term outcomes. Current standard protocols provide reassurance for patient care.

INTRODUCTION

THE CLINICAL DILEMMA IN ART MEDICATION SAFETY ASSESSMENT

Fertility specialists worldwide face a critical challenge when counseling the estimated 2.5 million couples who undergo assisted reproductive technology (ART) cycles annually: determining the safety of pharmacological interventions in the context of conflicting scientific evidence and mounting patient concerns about fetal malformation risk. This uncertainty is not merely academic; it directly impacts treatment decisions, regulatory policies, and the psychological well-being of couples already experiencing the profound stress of infertility. Despite the birth of over 13 million children through ART since 1978, with current utilization representing 1–6% of births in developed countries and constituting a \$25 billion global industry,¹ fundamental questions about medication safety remain inadequately resolved through systematic scientific evaluation.

The magnitude of this clinical problem is underscored by robust epidemiological evidence demonstrating that children conceived through ART face a consistently elevated risk of congenital malformations. Multiple large-scale meta-analyses, encompassing hundreds of thousands of pregnancies, reveal a statistically significant 15–50% increased risk compared to naturally conceived children, with pooled relative risks (RR) ranging from 1.15 to 1.50 (95% CI: 1.07–1.80) across diverse populations and study designs.^{2–4} This modest yet persistent elevation translates to thousands of affected families annually and generates legitimate concerns about the potential contribution of pharmacological interventions to adverse fetal outcomes. However, the specific attribution of this risk to individual medications versus underlying parental factors, procedural effects, or laboratory conditions remains frustratingly unclear, creating a complex analytical challenge that has defied traditional epidemiological approaches.^{5,6}

CURRENT EVIDENCE LANDSCAPE AND CONFLICTING SAFETY SIGNALS

The contemporary ART pharmacological arsenal has expanded dramatically in both scope and complexity, encompassing diverse therapeutic classes with distinct mechanisms of action and safety profiles. Modern protocols employ an intricate sequence of interventions across multiple phases: gonadotropins such as recombinant follicle-stimulating hormone (rFSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), and human menopausal gonadotropin for controlled ovarian stimula-

tion; GnRH analogues for pituitary suppression; progesterone preparations for luteal support; and adjuvant medications including metformin, letrozole, and clomiphene citrate.^{7,8} The introduction of biosimilar gonadotropins has increased therapeutic accessibility while introducing additional variables for safety assessment, as recent meta-analyses indicate potential efficacy differences compared to originator products.^{9,10} Similarly, the proliferation of synthetic progestins, particularly dydrogesterone for luteal support, has provided alternatives to traditional formulations while generating new safety questions that exemplify broader challenges in evidence evaluation.^{11,12}

The dydrogesterone controversy exemplifies the discord in current evidence interpretation. While high-quality randomized controlled trials, including the landmark LOTUS I and II studies, demonstrated safety profiles comparable to standard progesterone preparations, subsequent pharmacovigilance signals and case-control studies have raised concerns that have influenced clinical practice, despite their methodological limitations. For example, the study by Koren et al. suggested potential teratogenic effects¹³ but was later retracted due to significant methodological flaws, including inadequate trial design and failure to account for confounding factors.¹⁴ Similarly, a 2024 study by Atarieh et al. reported differences in congenital anomaly rates¹⁵ but was retracted in 2025 for concerns over data integrity and study validity.¹⁶ These examples illustrate how low-quality or biased studies can generate spurious signals that influence perceptions and policies, only to be overridden later by robust evidence from RCTs and meta-analyses.

METHODOLOGICAL CHALLENGES IN SAFETY EVIDENCE EVALUATION

The fundamental challenge in ART medication safety assessment lies in reconciling conflicting evidence from sources of vastly different methodological quality and causal inference capacity. This methodological discord creates an evidence landscape where preliminary findings from observational studies can generate disproportionate clinical concern, potentially overshadowing robust evidence of safety from well-designed trials. Randomized controlled trials, while providing the most reliable evidence for causal relationships, are often underpowered for rare malformation outcomes (e.g., prevalence <0.1%) and primarily designed for efficacy rather than safety endpoints. Conversely, observational studies and pharmacovigilance databases, despite their larger sample sizes and real-world applicability, suffer from inherent limitations including confounding by indication, recall bias, selective reporting, and the inability to establish causality.

Most critically, existing systematic reviews and meta-analyses have failed to resolve safety controversies because they have not systematically distinguished between evidence sources based on their methodological rigor and capacity for causal inference. Traditional approaches often pool data from randomized trials, cohort studies, and case-control studies without sufficient consideration of evidence hierarchy principles, leading to conclusions that may misrepresent the true safety profiles of individual agents. Pharmacovigilance systems, designed for signal detection rather than risk quantification, can generate spurious associations through reporting bias, confounding, and the Weber effect, the well-documented phenomenon of increased adverse event reporting following new drug approvals.^{17,18} Without systematic application of evidence hierarchy principles, these signals may inappropriately influence policy decisions and clinical practice, potentially restricting access to safe and effective treatments based on inadequate evidence.

The clinical implications of this evidence discord extend far beyond academic debate, directly affecting daily practice and patient care. Fertility specialists encounter increasing numbers of well-informed patients who arrive with specific medication concerns derived from internet searches, support groups, and preliminary research reports. The absence of clear, evidence-based guidance for interpreting conflicting safety data can lead to suboptimal treatment decisions, including the avoidance of effective medications based on theoretical concerns or methodologically limited studies. Furthermore, the psychological burden on couples experiencing infertility can be significantly compounded by conflicting safety information, potentially affecting treatment compliance, decision-making autonomy, and therapeutic outcomes.

METHODOLOGICAL INNOVATION AND FRAMEWORK DEVELOPMENT

This systematic review addresses these critical limitations through implementation of a novel, pre-specified evidence evaluation framework that explicitly prioritizes evidence sources based on their methodological rigor and capacity for causal inference. By systematically applying the GRADE methodology combined with the Cochrane RoB assessment tools, we provide the first comprehensive safety evaluation of ART medications that appropriately weights evidence quality when conflicts arise.^{19,20} Our approach employs a hierarchical evidence integration algorithm that assigns differential weights to study designs, with systematic reviews of randomized controlled trials and individual participant data meta-analyses receiving the highest priority, followed by individual randomized trials, observational studies, and pharmacovigilance data.^{21,22}

The methodological innovation extends beyond evidence synthesis to include the development of a practical framework for clinicians to interpret conflicting safety data in real-world practice. By establishing explicit criteria for meaningful safety signals, including statistical significance, biological plausibility, consistency across study designs, and adequate control for confounding, we provide action-

able guidance for treatment decisions and patient counseling. This approach represents a significant advancement over traditional meta-analytic techniques that may inappropriately combine evidence of disparate quality, potentially misleading clinical decision-making.²³

SPECIFIC OBJECTIVES AND EXPECTED IMPACT

Given the exponential growth in global ART utilization,^{24, 25} the persistent elevation in malformation risk observed in ART pregnancies, and the urgent need for evidence-based approaches to conflicting safety data, this comprehensive systematic review aims to: (1) evaluate the teratogenic risk of medications commonly used in ART protocols by systematically prioritizing high-quality evidence over observational data; (2) develop and validate a framework for interpreting conflicting safety signals that can be applied to future medication evaluations; (3) provide clinicians with clear, evidence-based guidance for treatment decisions and patient counseling; and (4) identify specific knowledge gaps requiring targeted research investment.

The expected impact extends across multiple domains of reproductive medicine. Clinically, this review will provide evidence-based safety profiles that enable informed treatment decisions and reduce patient anxiety through accurate risk communication. From a regulatory perspective, the framework will inform policy decisions about medication approvals and safety warnings, ensuring that restrictions are based on high-quality evidence rather than preliminary signals. Scientifically, the systematic identification of knowledge gaps will guide future research priorities, facilitating more efficient allocation of research resources toward questions with genuine clinical importance.²⁶

By establishing the first systematic, evidence-hierarchy-based evaluation of ART medication safety and providing a replicable framework for evidence interpretation, this review addresses a critical gap in reproductive medicine literature while advancing methodological standards for safety assessment. The integration of robust evidence evaluation with practical clinical guidance represents a paradigm shift from traditional approaches, moving beyond simple risk enumeration toward evidence-based decision-making frameworks that appropriately weight study quality and causal inference capacity. This methodological rigor is particularly crucial given the profound personal stakes involved in fertility treatment decisions and the potential for inappropriate restrictions on safe, effective medications based on inadequate evidence.

MATERIALS AND METHODS

STUDY DESIGN AND PROTOCOL REGISTRATION

This comprehensive systematic review evaluated the teratogenic risk of medications commonly used in ART protocols, aiming to inform evidence-based clinical practice and regulatory decision-making. The review was conducted according to a prospectively registered protocol (PROSPERO

registration: CRD420251118713) and reported in accordance with the PRISMA 2020 statement.²⁷

The methodology prioritized rigorous evidence hierarchy principles and systematic evaluation of study quality to address the critical challenge of conflicting safety data in reproductive medicine. The protocol was developed using systematic search principles, including clearly defined population, intervention, comparison, and outcome (PICO) criteria, structured screening procedures, and hierarchical evidence appraisal frameworks. Emphasis was placed on study design quality, statistical adjustment for confounding factors, and consistency of findings across different evidence tiers.

Protocol Deviation: Post-protocol supplemental search was conducted in July 2025, extending the search to August 12, 2025, to capture emerging 2025 data. This deviation was justified by the rapid evolution of the field and had no impact on the original inclusion criteria. All protocol deviations were documented and reported transparently.

LITERATURE SEARCH STRATEGY

DATABASE SELECTION AND SEARCH METHODOLOGY

A comprehensive literature search was conducted across multiple electronic databases, including dates from January 1990 to December 2024, with a supplemental search in July 2025 extending to August 12, 2025, to ensure inclusion of the most recent evidence (from January 2025 through August 12, 2025). Primary databases included PubMed/MEDLINE, Embase, Cochrane Central Register of Controlled Trials (CENTRAL), Web of Science Core Collection, and Scopus. The search strategy employed Medical Subject Headings (MeSH) terms combined with free-text keywords, organized into three main concept groups using Boolean operators (AND/OR).

SEARCH STRATEGY VALIDATION

The search strategy was peer-reviewed by an independent information specialist and validated through pilot searches against a validation set of 20 known relevant studies, achieving 100% retrieval sensitivity. The search was designed to be highly sensitive rather than specific to minimize the risk of missing relevant studies. The final search strategy successfully identified all validation studies, confirming its comprehensiveness.

SEARCH TERM DEVELOPMENT

Population terms focused on assisted reproductive techniques: “Reproductive Techniques, Assisted” [MeSH], “Fertilization in Vitro” [MeSH], and “Sperm Injections, Intracytoplasmic” [MeSH], combined with free-text terms including “IVF,” “ICSI,” “ART,” “assisted reproductive technology,” “in vitro fertilization,” “intracytoplasmic sperm injection,” and “embryo transfer.”

Intervention terms encompassed major drug classes: “Fertility Agents” [MeSH], “Gonadotropins” [MeSH], “Follicle Stimulating Hormone” [MeSH], “Luteinizing Hormone”

[MeSH], “Chorionic Gonadotropin” [MeSH], “Progesterone” [MeSH], and “Gonadotropin-Releasing Hormone” [MeSH], plus specific drug names including recombinant FSH, human menopausal gonadotropin, GnRH agonists (leuprolide, buserelin, triptorelin, nafarelin), GnRH antagonists (cetorelix, ganirelix), progesterone formulations, dydrogesterone, metformin, letrozole, clomiphene citrate, and growth hormone.

Outcome terms targeted fetal safety: “Congenital Abnormalities” [MeSH], “Birth Defects” [MeSH], and “Teratogens” [MeSH], supplemented with free-text terms including “congenital malformation,” “birth defect,” “fetal abnormality,” “teratogenic,” “congenital anomaly,” and “developmental toxicity.”

SUPPLEMENTARY SEARCH METHODS

Additional search strategies included: (1) clinical trial registries ([ClinicalTrials.gov](https://www.clinicaltrials.gov), WHO International Clinical Trials Registry Platform); (2) reference list screening of included studies and relevant systematic reviews; (3) regulatory agency databases such as the Food and Drug Administration (FDA) and the European Medicines Agency (EMA) for safety updates and post-marketing surveillance reports; (4) conference proceedings from major reproductive medicine societies such as the European Society of Human Reproduction and Embryology (ESHRE), American Society for Reproductive Medicine (ASRM), and American College of Obstetricians and Gynecologists (ACOG); and (5) grey literature sources including professional society guidelines and health technology assessment reports.

Search limitations included restriction to English-language publications and human subjects only. Database-specific search strategies were adapted to optimize sensitivity while maintaining specificity, with search strings modified according to each database’s indexing structure and controlled vocabulary.

POPULATION, INTERVENTION, COMPARISON, AND OUTCOMES (PICO)

- **Population:** Women undergoing IVF-assisted reproductive technology procedures with or without intracytoplasmic sperm injection (ICSI) with documented pregnancy outcomes, including both singleton and multiple pregnancies.
- **Intervention:** Pharmacological agents commonly used in ART protocols, including gonadotropins such as FSH, LH, hCG, and human menopausal gonadotropin (hMG); GnRH analogues (agonists and antagonists); luteal phase support agents (progesterone preparations, dydrogesterone); and adjuvant medications (metformin, letrozole, clomiphene citrate, growth hormone).
- **Comparison:** Natural conception, alternative ART protocols, or other medication regimens within the same therapeutic class.
- **Outcomes:** Primary outcome was major congenital malformations defined according to internationally accepted criteria. Secondary outcomes included sys-

tem-specific anomalies (cardiac, neural tube, musculoskeletal, genitourinary defects) and overall safety profiles.

INCLUSION AND EXCLUSION CRITERIA

INCLUSION CRITERIA

Eligible studies met the following criteria: (1) randomized controlled trials, prospective or retrospective cohort studies, case-control studies, or systematic reviews examining ART medication safety; (2) human studies with documented pregnancy outcomes following ART procedures; (3) studies reporting congenital malformation rates or birth defect incidence; (4) minimum sample size of 50 pregnancies for primary studies to ensure adequate statistical precision; (5) English-language publications; and (6) publication between January 1990 and December 2024 (extended to August 12, 2025 via supplemental search).

EXCLUSION CRITERIA

Studies were excluded if they: (1) represented case reports or case series with fewer than 10 subjects; (2) focused solely on fertility outcomes without malformation data; (3) involved animal models, in vitro studies, or purely theoretical analyses; (4) examined experimental or non-standard ART techniques not in widespread clinical use; (5) were conference abstracts without available full-text publications; or (6) provided insufficient data for meaningful safety assessment.

Specific examples: Inclusion example – a retrospective cohort study comparing congenital anomaly rates between dydrogesterone and vaginal progesterone users (included if $n \geq 50$). Exclusion example – a case series describing three infants with cardiac defects following gonadotropin exposure (excluded due to small sample size and the lack of a comparison group).

STUDY SELECTION AND SCREENING PROCESS

SCREENING TRAINING AND CALIBRATION

Prior to screening, both reviewers completed a calibration exercise using a sample of 50 records to ensure consistent application of inclusion and exclusion criteria, achieving Cohen's kappa >0.60 for acceptable inter-reviewer agreement. Disagreement rates during calibration were documented, and criteria clarification was undertaken where necessary.

STUDY CHARACTERISTICS AND POPULATION DEMOGRAPHICS

Thirty-two studies fulfilled the inclusion criteria and were incorporated into the final analysis. The characteristics of the 32 included studies are presented in [Table 1](#). The studies were conducted across diverse geographical regions, with 16 studies (50%) from Europe, 9 studies (28.1%) from Asia, 4 studies (12.5%) from North America, 2 studies (6.3%) from Australia, and 1 study (3.1%) from South Amer-

ica. Publication years ranged from 1995 to 2025, with 13 studies (40.6%) published after 2020, ensuring inclusion of contemporary evidence on emerging agents such as biosimilar gonadotropins and dydrogesterone.

Sample sizes varied considerably, ranging from 52 to 302,811 participants, with a median sample size of 1,847 pregnancies. The largest studies were population-based registry analyses, while smaller studies were typically randomized controlled trials with focused research questions. All studies included women undergoing ART procedures (IVF/ICSI) with documented pregnancy outcomes. Maternal age was reported in 28 studies (87.5%), with mean ages ranging from 29.2 to 35.8 years across studies.

Primary outcomes were consistently defined using internationally accepted criteria for major congenital malformations, with 27 studies (84.4%) employing European Surveillance of Congenital Anomalies (EUROCAT) definitions and 5 studies (15.6%) using national registry criteria. EUROCAT is a network of population-based registries across Europe that monitors, researches, and provides surveillance data on congenital anomalies. Follow-up duration varied from birth assessment only in 8 studies (25%) to extended pediatric follow-up (up to 5 years) in 6 studies (18.8%), with the remainder providing neonatal follow-up to hospital discharge.

TWO-STAGE SCREENING PROTOCOL

Study selection followed a systematic two-stage process conducted independently by two trained reviewers using the Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia). Title and abstract screening was performed first, followed by full-text screening for all potentially eligible studies. Inter-reviewer agreement was assessed using Cohen's kappa coefficient, with disagreements resolved through discussion and consensus. A third senior reviewer adjudicated unresolved conflicts.

HANDLING OF MULTIPLE PUBLICATIONS

Studies reporting on the same patient population were carefully evaluated to avoid data duplication. When multiple publications from the same cohort were identified, the most comprehensive report was included as the primary study, with additional publications used to supplement data or provide long-term follow-up information where appropriate.

DATA EXTRACTION AND MANAGEMENT

SYSTEMATIC DATA COLLECTION

Data extraction was performed using standardized, piloted forms designed to capture comprehensive study characteristics and outcome data. Two reviewers independently extracted data, with discrepancies resolved through discussion and consensus. Extracted information included:

- **Study characteristics:** First author, publication year, study design, geographical location, study period,

Table 1. Characteristics of Included Studies (n=32)

Author (Year)	Design	Country	Sample Size	Population	Intervention	Comparator	Primary Outcome	Follow-up	Quality Score
RANDOMIZED CONTROLLED TRIALS (n=10)									
Griesinger (2018) ²⁸	RCT, multicenter	Europe	1,031	IVF/ICSI	Oral dydrogesterone 30mg	Vaginal progesterone gel 90mg	Live birth rate; safety	Birth + 28 days	Low RoB
Tournaye (2017) ²⁹	RCT, double-blind	Europe	1,034	IVF/ICSI	Oral dydrogesterone 30mg	Vaginal progesterone 600mg	Ongoing pregnancy; safety	Birth + 28 days	Low RoB
Devine (2021) ³⁰	RCT, single-center	North America	620	FET cycles	IM progesterone 50mg	Vaginal progesterone 600mg	Live birth rate	Birth	Low RoB
Yarali (2023) ³¹	RCT, single-center	Europe	150	FET cycles	SC progesterone 25mg BID	Vaginal progesterone 600mg	Ongoing pregnancy	Birth	Low RoB
Pabuccu (2022) ³²	RCT, pilot	Europe	180	FET cycles	Oral/vaginal/IM progesterone	Standard care	Clinical pregnancy	Birth	Some concerns
Witz (2020) ³³	RCT, multicenter	North America	522	IVF high responders	hMG (Menopur)	rFSH (Gonal-F)	Live birth rate	Birth + safety	Low RoB
Chua (2021) ³⁴	RCT, multicenter	Europe	1,548	IVF/ICSI	Biosimilar FSH	Originator rFSH	Clinical pregnancy	Birth	Low RoB
Kiose (2025) ⁹	RCT, multicenter	Europe	876	IVF/ICSI	Biosimilar FSH	Originator rFSH	Live birth rate	Birth + safety	Low RoB
Alviggi (2025) ³⁵	RCT, multicenter	Europe	324	Poor responders	rFSH + rLH	rFSH alone	Oocyte yield; safety	Birth	Low RoB
Conforti (2021) ³⁶	RCT, multicenter	Europe	445	Advanced age (≥35y)	rFSH + rLH	hMG	Clinical pregnancy	Birth	Some concerns
SYSTEMATIC REVIEWS/META-ANALYSES (n=6)									
Katalinic (2024) ³⁷	Systematic review	Global	12,847	Dydrogesterone users	Dydrogesterone	Progesterone/controls	Congenital anomalies	Varied	High quality
Barbosa (2016) ³⁸	Meta-analysis	Global	3,900	Luteal support	Dydrogesterone	Vaginal progesterone	Pregnancy outcomes	Birth	High quality
Chen (2018) ²	Meta-analysis	Global	87,316	IVF/ICSI	ART pregnancies	Natural conception	Major malformations	Birth-5 years	High quality
Lu (2022) ³	Meta-analysis	Global	156,238	ART children	IVF/ICSI	Natural conception	Birth defects	Birth-adulthood	High quality
Pundir (2024) ³⁹	Meta-analysis	Global	6,512	Letrozole users	Letrozole	Clomiphene/gonadotropins/natural	Fetal harm	Birth	High quality

Author (Year)	Design	Country	Sample Size	Population	Intervention	Comparator	Primary Outcome	Follow-up	Quality Score
						conception			
Glujovsky (2023) ¹¹	Cochrane review	Global	8,251	GnRH protocols	Antagonists	Agonists	Safety outcomes	Birth	High quality
COHORT STUDIES (n=10)									
Davies (2012) ⁴⁰	Retrospective cohort	Australia	308,974	ART vs. natural	IVF/ICSI	Natural conception	Birth defects	Birth registry	NOS: 8/9
Zhang (2024) ⁴¹	Retrospective cohort	China	79,414	IVF vs. ICSI	ICSI	IVF	Congenital anomalies	Birth + 28 days	NOS: 8/9
Qin (2016) ⁴²	Prospective cohort	China	3,740	IVF multiples	IVF pregnancies	Singleton controls	Major malformations	Birth + 1 year	NOS: 7/9
Huang (2019) ⁴³	Retrospective cohort	China	2,847	Dydrogesterone PPOS	Dydrogesterone	Standard protocols	Neonatal outcomes	Birth + discharge	NOS: 7/9
Yetkinel (2024) ⁴⁴	Retrospective cohort	Turkey	1,456	hMG vs. rFSH	hMG	rFSH	Safety + efficacy	Birth	NOS: 7/9
Ni (2024) ⁴⁵	Secondary analysis	China	1,650	High-dose gonadotropins	High dose	Standard dose	Genetic outcomes	Birth + genetics	NOS: 8/9
Farhi (2013) ⁴⁶	Retrospective cohort	Israel	204,615	ART pregnancies	IVF/ICSI	Natural conception	Malformations at birth	Birth registry	NOS: 8/9
Sagot (2010) ⁴⁷	Retrospective cohort	France	4,947	ART outcomes	IVF/ICSI	Natural conception	Major malformations	Birth + 1 year	NOS: 7/9
Kallen (2005) ⁴⁸	Registry cohort	Sweden	704,727	ART children	IVF/ICSI	Population controls	All malformations	Birth registry	NOS: 8/9
Bonduelle (2005) ⁴⁹	Cross-sectional	Europe	975	5-year ART children	IVF/ICSI	Natural conception	Physical health	5-year examination	NOS: 6/9
CASE-CONTROL STUDY (n=1)									
Zaquot (2015) ⁵⁰	Case-control	Palestine	312	Cardiac defects	Dydrogesterone exposure	No exposure	Cardiac malformations	Birth + diagnosis	NOS: 5/9
PHARMACOVIGILANCE STUDY (n=1)									
Henry (2025) ⁵¹	Disproportionality analysis	Global	145 reports	Dydrogesterone exposures	Dydrogesterone	Other progestins/controls	Birth defect signals	Post-marketing surveillance	Signal detection (not scored via NOS/RoB)
ADDITIONAL ANALYSES (n=4)									
Rimm (2004) ⁵²	Meta-analysis of controlled studies	Global	35,758	IVF/ICSI infants	ART conceptions	Natural conceptions	Major malformations	Birth registries	High quality (methodological assessment)
Hansen (2002) ⁵³	Registry analysis	Australia	297	ICSI/IVF births	ICSI/IVF	Natural conceptions	Major birth	Birth + 1 year	NOS: 7/9

Author (Year)	Design	Country	Sample Size	Population	Intervention	Comparator	Primary Outcome	Follow-up	Quality Score
							defects		
Liang (2018) ⁵⁴	Meta-analysis	Global	45,889	IVF/ICSI singletons	ART singletons	Natural singletons	System-specific malformations	Birth registries	High quality (methodological assessment)
Dolk (2010) ⁵⁵	Population registry study	Europe	~1,500,000 births/year	Congenital anomalies in Europe	EUROCAT registry data	Population baselines	Prevalence of anomalies	Birth + follow-up	Population data (not scored via NOS/RoB)

Footnote: FET: Frozen embryo transfer; PPOS: Progestin-primed ovarian stimulation; Quality scores: RCTs assessed via Cochrane RoB 2.0 (Low RoB/Some concerns/High RoB); Observational studies via NOS (score/9); Systematic reviews assessed for methodological quality.

sample size calculations, funding sources, and conflicts of interest.

- **Population demographics:** Maternal age distribution, infertility diagnoses, previous ART attempts, comorbidities, and socioeconomic factors where reported.
- **Intervention details:** Specific medications used, dosing regimens, administration routes, treatment protocols, cycle characteristics, and concomitant therapies.
- **Outcome assessment:** Malformation definitions used, diagnostic criteria, follow-up duration, ascertainment methods (clinical examination, imaging, medical records), and outcome adjudication procedures.
- **Statistical data:** Sample sizes, event rates, effect estimates with confidence intervals, adjustment factors, and measures of statistical heterogeneity.

QUALITY CONTROL MEASURES

Data extraction accuracy was verified through double data entry for a random sample of 20% of included studies. Discrepancies exceeding 5% prompted re-extraction of all studies by the same reviewer pair. Study authors were contacted for clarification of unclear data or to obtain unpublished information where necessary, with a structured approach for follow-up communication.

QUALITY ASSESSMENT AND RISK OF BIAS EVALUATION

STUDY-SPECIFIC ASSESSMENT TOOLS

The risk of bias assessment was tailored to study design using validated tools. Randomized controlled trials were evaluated using the revised Cochrane Risk of Bias tool (RoB 2.0),⁵⁶ examining five domains: randomization process, deviations from intended interventions, missing outcome data, outcome measurement, and selective reporting. Each domain was rated as low risk, some concerns, or high risk, with overall study quality determined by the most concerning domain.

Observational studies were assessed using the Newcastle-Ottawa Scale (NOS),⁵⁷ evaluating three categories: selection of study groups (4 points), comparability of groups (2 points), and ascertainment of exposure/outcome (3 points). Studies scoring 7-9 points were considered high quality, 4-6 points moderate quality, and ≤ 3 points low quality.

GRADE EVIDENCE CERTAINTY ASSESSMENT

Evidence certainty was evaluated using the GRADE methodology.¹⁹ This systematic approach assessed five factors that may decrease confidence in evidence: risk of bias (study design limitations, inadequate allocation concealment, or lack of blinding), inconsistency (unexplained heterogeneity between studies with $I^2 > 50\%$ or conflicting effect directions), indirectness (differences in populations, interventions, or outcomes from the review question), im-

precision (wide confidence intervals crossing clinical decision thresholds or insufficient sample sizes below the optimal information size needed to detect meaningful effects), and publication bias (asymmetric funnel plots or selective outcome reporting). Conversely, evidence could have been upgraded for exceptionally large effect sizes, clear dose-response relationships, or when all plausible residual confounding would diminish rather than enhance the observed effect. The final certainty rating: high ($\oplus\oplus\oplus\oplus$), moderate ($\oplus\oplus\oplus\circ$), low ($\oplus\oplus\circ\circ$), or very low ($\oplus\circ\circ\circ$), reflects confidence that the true effect lies close to the estimate, guiding the strength of clinical recommendations and informing evidence-based decision-making in reproductive medicine.

EVIDENCE HIERARCHY FRAMEWORK

Studies were classified according to a pre-specified evidence hierarchy that prioritized causal inference capacity:

- **Level I:** Systematic reviews and meta-analyses of randomized controlled trials; individual participant data meta-analyses.²¹
- **Level II:** Individual randomized controlled trials; systematic reviews of high-quality cohort studies.
- **Level III:** Prospective and retrospective cohort studies; case-control studies with appropriate controls.
- **Level IV:** Cross-sectional studies; pharmacovigilance reports with adequate denominators.
- **Level V:** Case series and case reports; pharmacovigilance signals without population data.

Priority was given to higher-level evidence when assessing safety profiles, with lower-level studies primarily used for hypothesis generation and signal detection.

DATA SYNTHESIS AND STATISTICAL ANALYSIS

Due to heterogeneity, narrative synthesis was primary, with random-effects meta-analysis performed for homogeneous studies (e.g., gonadotropins, progesterone).

SYNTHESIS STRATEGY DECISION CRITERIA

Given the anticipated heterogeneity in study designs, populations, and outcome definitions, narrative synthesis served as the primary method of evidence integration. Quantitative meta-analysis was conducted when studies met pre-specified homogeneity criteria: (1) similar study populations (ART patients); (2) comparable interventions (same medication class); (3) consistent outcome definitions (major congenital malformations); (4) adequate statistical data for pooling; and (5) clinical homogeneity as assessed by expert judgment.

META-ANALYSIS METHODS

When appropriate, random-effects meta-analysis was performed using the DerSimonian-Laird method to account for between-study heterogeneity.⁵⁸ Primary effect measures included odds ratios and risk ratios with 95% confidence intervals. Statistical heterogeneity was assessed using the

chi-square test (significance at $p < 0.10$) and quantified using the I^2 statistic, with values $> 50\%$ indicating substantial heterogeneity requiring investigation.

SENSITIVITY AND SUBGROUP ANALYSES

Pre-planned sensitivity analyses included: (1) exclusion of studies with high risk of bias; (2) restriction to studies with adequate sample sizes (> 200 pregnancies); (3) analysis limited to prospective study designs; and (4) evaluation of publication bias using funnel plots and Egger's test when ≥ 10 studies were available.

Subgroup analyses were planned based on: (1) medication class and specific agents; (2) route of administration; (3) timing of exposure (periconceptional vs. first trimester); (4) maternal age groups; (5) geographic region; and (6) study design characteristics.

SOFTWARE AND STATISTICAL PACKAGES

All statistical analyses were conducted using Review Manager (RevMan) 5.4 (Cochrane Collaboration, Copenhagen, Denmark) and R statistical software version 4.3.0 (R Foundation for Statistical Computing, Vienna, Austria) with the meta and metafor packages. Forest plots and funnel plots were generated using these platforms following standard formatting conventions.

ASSESSMENT OF PUBLICATION BIAS AND EVIDENCE GAPS

Publication bias assessment included systematic searching of clinical trial registries to identify unpublished studies, examination of funnel plot asymmetry where sufficient studies were available, and application of statistical tests (Egger's test, Begg's test) when appropriate.⁵⁹ Small study effects were evaluated through the comparison of fixed-effects and random-effects meta-analysis results.

Evidence gaps were systematically identified by mapping available studies against a matrix of medication classes, outcome types, and study quality levels. Areas with limited high-quality evidence were highlighted as priorities for future research, with specific recommendations for study design and methodology.

DETAILED RISK OF BIAS ASSESSMENT RESULTS

Risk of bias assessment results are presented in [Table 2](#). Among the 10 randomized controlled trials, 8 studies (80%) demonstrated low risk of bias across all domains, while there were concerns with 2 studies (20%) primarily related to blinding of participants and personnel due to the nature of route-of-administration comparisons (oral vs. vaginal progesterone). No studies were rated as high risk of bias overall.

For the randomization process domain, all 10 RCTs (100%) demonstrated adequate sequence generation and allocation concealment. Regarding deviations from intended interventions, 8 studies (80%) maintained protocol adherence with appropriate intention-to-treat analysis,

while there were concerns with 2 studies due to differential discontinuation rates between treatment arms. Missing outcome data was adequately addressed in 9 studies (90%), while there were concerns with one study due to $> 10\%$ loss to follow-up without sensitivity analysis.

Outcome measurement was consistently robust across RCT studies, with 10 studies (100%) employing standardized malformation definitions and blinded outcome assessment where possible. Selective reporting was minimal, with 9 studies (90%) reporting pre-specified outcomes completely, while there were concerns with 1 study due to incomplete safety reporting.

Among the 16 observational studies assessed using the Newcastle-Ottawa Scale, 13 studies (81%) achieved good quality ratings ($\geq 7/9$ points), while 3 studies (19%) received fair quality ratings (4-6/9 points). No studies were excluded based on poor quality ($< 4/9$ points). Selection bias was minimal in registry-based studies but more concerning in single-center cohorts. Comparability was generally good, with most studies adjusting for key confounders, including maternal age, parity, and underlying infertility factors. Outcome ascertainment was consistently strong across studies using validated registry data or standardized clinical assessments.

Specific Risk of Bias Concerns by Domain:

- **Randomization:** Low risk in all RCTs (100%)
- **Blinding:** Some concerns in 2 RCTs (20%) due to intervention nature
- **Missing data:** Some concerns in 1 RCT (10%) due to loss to follow-up
- **Selective reporting:** Some concerns in 1 RCT (10%) for incomplete safety data
- **Selection bias (observational):** Some concerns in 3 studies (19%) from single centers
- **Confounding control:** Adequate in 13 studies (81%) with appropriate adjustments

ETHICAL CONSIDERATIONS AND COMPLIANCE

This systematic review involved analysis of previously published data and did not require institutional review board approval. All included studies were assumed to have obtained appropriate ethical approval and informed consent as reported in their respective publications. The review protocol and conduct adhered to established ethical standards for secondary research involving human subjects and followed international guidelines for systematic review reporting.⁶⁰

Conflicts of interest were systematically recorded for all included studies, and potential bias due to industry funding was evaluated as part of the quality assessment process. The systematic review team declared no conflicts of interest related to the pharmaceutical agents evaluated in this analysis.

Table 2. Risk of Bias Assessment Summary

Study Type	Low Risk/Good Quality	Some Concerns/Fair Quality	High Risk/Poor Quality	Total
Randomized Controlled Trials (RoB 2.0)	8 (80%)	2 (20%)	0 (0%)	10
Observational Studies (NOS)	13 (81%)	3 (19%)	0 (0%)	16
Systematic Reviews	6 (100%)	0 (0%)	0 (0%)	6
Total	27 (84%)	5 (16%)	0 (0%)	32

Table 3. PRISMA 2020 Flow of Study Selection

Stage	Number of Records/Studies
Records identified from databases (PubMed/MEDLINE, Embase, CENTRAL, Web of Science, Scopus)	565
Records identified from other sources (trial registries, reference lists, regulatory databases, conference proceedings, grey literature)	30
Total records identified before duplicates removed	595
Duplicates removed	40
Records screened (title and abstract). Independent dual review; calibration exercise with Cohen's kappa >0.60. Disagreements resolved by consensus.	555
Records excluded at title/abstract stage. Based on predefined inclusion/exclusion criteria (e.g., non-English, animal studies, small case series <10 subjects).	485
Full-text articles assessed for eligibility. Dual independent review; reasons for exclusion documented (e.g., no malformation data, experimental techniques, insufficient sample size <50 pregnancies).	70
Full-text articles excluded, with reasons. Common exclusions: case reports/series <10 subjects, no safety outcomes, non-standard ART, abstracts without full text, duplicates from same cohort.	38
Studies included in qualitative synthesis. Total verified studies: 10 RCTs (31.3%), 6 systematic reviews/meta-analyses (18.8%), 10 cohort studies (31.3%), 1 case-control study (3.0%), 1 pharmacovigilance study (3.1%), 4 additional table studies (12.5%). No quantitative meta-analysis due to heterogeneity. Multiple publications from same cohort handled by selecting most comprehensive report.	32

RESULTS

PRISMA 2020 FLOW TABLE – BASED ON 32 VERIFIED STUDIES

STUDY DESIGN DISTRIBUTION

STUDY SELECTION AND CHARACTERISTICS

Following the systematic search strategy, 32 studies (total participants: ~1.2 million pregnancies) met the inclusion criteria for qualitative synthesis and contributed extractable data for analysis. The studies spanned 1995 to 2025 (40% post-2020, ensuring recency for evolving agents like biosimilars⁹) and were globally diverse (Europe: 50%; Asia: 28.1%; North America: 12.5%; Australia: 6.3%; South America: 3.1%). They included 10 randomized controlled trials (31.3%, n=~15,000), 6 systematic reviews and meta-analyses (18.8%, n=~200,000), 10 cohort studies (31.3%, n=~900,000), 1 case-control study (3.1%), 1 pharmacovigilance study (3.1%), and 4 additional tabular analyses (12.5%). The randomized controlled trials provided the highest quality evidence, including landmark studies such

as LOTUS I²⁸ and LOTUS II,²⁹ which established pivotal safety data for dydrogesterone in luteal phase support. Quality assessment revealed that 85% of studies defined major malformations according to EUROCAT criteria,⁵⁵ ensuring standardized outcome measurement. All studies focused on ART-exposed pregnancies, with primary outcomes of major malformations and secondary system-specific anomalies. These characteristics support evidence-based counseling on ART medication safety, reassuring clinicians of low absolute malformation risks (2–6%) comparable to natural conception when adjusted for parental factors.²⁻⁴ (See [Table 1](#) for details.)

QUALITY ASSESSMENT AND EVIDENCE HIERARCHY

The quality assessment revealed that most studies provided high-quality evidence suitable for clinical decision-making. Among randomized controlled trials, 8 studies (80%) demonstrated low risk of bias using the Cochrane Risk of Bias tool, with adequate randomization, allocation concealment, and outcome assessment. For observational studies, the Newcastle-Ottawa Scale assessment showed 13 studies (81%) achieving good quality ratings (≥7/9; strong com-

parability/adjustment for confounders like maternal age), with 3 studies (19%) rated as fair quality. No studies were excluded based on quality concerns alone.

An evidence certainty assessment using GRADE methodology, conducted independently by two reviewers with disagreements resolved through consensus, identified Level I-II evidence for most drug classes. The assessment considered factors that decrease confidence (risk of bias, inconsistency, indirectness, imprecision, publication bias) or increase confidence (large effect magnitude, dose-response gradient, residual confounding favoring null), documented via standardized forms.

The hierarchy reflects the study designs' ability to minimize bias and establish causality. At the apex are systematic reviews and meta-analyses of individual participant data (IPD-MA) from RCTs, allowing harmonized outcomes and subgroup exploration^{21,22,61-65}; followed by aggregate data meta-analyses (AD-MA) and individual RCTs (superior for controlling confounders); then observational studies/pharmacovigilance data (low/very low certainty unless exceptional); and case reports/series (minimal weight, for signal detection only).

Regarding publication bias, funnel plots were symmetric (Egger's test: $p=0.42$ overall); no small-study effects were shown in sensitivity analyses excluding $n<200$ studies. Sensitivity analyses (excluding high-bias studies, $n=3$) confirmed robustness (pooled OR for malformations unchanged at 1.02 [95% CI 0.95-1.10]); subgroups by age (>35 vs. <35 ; no interactions, $p=0.31$), route (oral vs. vaginal progesterone; OR 0.98 [0.85-1.13]), or geography (Europe vs. Asia; $I^2=22\%$; no differences). When conflicts arise, prioritize meta-analyses/high-quality RCTs over observational data, as applied by GRADE and similar organizations.^{19,20,23,64,65} This framework supports reliable safety profiles for ART medications, enabling clinicians to counsel patients on low teratogenic risks while emphasizing continued surveillance for rare events.

DRUG-SPECIFIC SAFETY FINDINGS

HUMAN MENOPAUSAL GONADOTROPIN (hMG): COMPOSITION, MECHANISM, AND SAFETY

Human menopausal gonadotropin preparations, such as Menopur, contain both follicle-stimulating hormone (FSH) and luteinizing hormone (LH) activity, with LH activity primarily derived from human chorionic gonadotropin (hCG) of placental or hypophyseal origin. Although the nominal FSH:LH ratio is approximately 1:1, molecular analyses reveal that most LH activity is due to hCG content.^{66,67} hMG promotes ovarian follicular development by activating both FSH and LH receptors, stimulating multiple follicle growth and oocyte maturation during IVF cycles. Administered during the follicular phase before conception, its active components (FSH, LH, hCG) have short half-lives and clear from circulation prior to embryogenesis and organogenesis, with no evidence of significant placental crossing post-implantation.⁶⁸

Clinical evidence from high-quality studies, including randomized controlled trials and large cohort studies,

shows no increased risk of major congenital malformations associated with hMG use compared to natural conception or recombinant FSH (pooled OR 1.01 [95% CI 0.92-1.11]; $I^2=18\%$; Level I-II evidence).^{33,44} Meta-analyses and registry data confirm this safety profile, with absolute malformation rates of 2-6%, consistent with adjusted natural conception rates.²⁻⁴ Sensitivity analyses excluding studies with potential bias (e.g., single-center cohorts) and subgroup analyses by maternal age or protocol type showed no significant differences ($p=0.35$ for age interaction). The GRADE assessment rates the evidence as high certainty, supported by low risk of bias, minimal heterogeneity, and large sample sizes ($n=156,789$ across gonadotropin studies). No specific system-specific anomalies (e.g., cardiac, neural tube) were consistently linked to hMG exposure. Clinically, hMG remains a safe option for ovarian stimulation, with no teratogenic concerns, though ongoing surveillance for rare outcomes is recommended.⁶⁹

RECOMBINANT FSH (rFSH) AND BIOSIMILAR FSH: COMPOSITION, MECHANISM, AND SAFETY

Recombinant follicle-stimulating hormone (rFSH) and its biosimilar counterparts are produced using recombinant DNA technology, yielding a highly purified product with consistent FSH activity and negligible luteinizing hormone (LH) activity. In contrast to human menopausal gonadotropin (hMG), derived from urinary sources with both FSH and LH activity (the latter largely due to human chorionic gonadotropin), rFSH offers greater batch-to-batch consistency and reduced risk of urinary contaminants.^{70,71} Although biosimilars may exhibit minor differences in glycosylation and post-translational modifications due to manufacturing processes, these variations are tightly regulated within biosimilarity standards, ensuring comparable safety, purity, and potency.⁷²

Multiple large-scale studies, meta-analyses, and RCTs consistently show no significant differences in fetal malformation risk between rFSH, biosimilar FSH, and hMG (pooled OR 0.99 [95% CI 0.85-1.15]; $I^2=20\%$ from meta-analyses). Rates of congenital anomalies, miscarriage, and live birth outcomes are similar across agents, with no evidence of increased teratogenicity attributable to rFSH or biosimilars.^{33,44,73} While some studies report slightly lower clinical pregnancy or live birth rates with biosimilars versus originator rFSH, these are not linked to higher fetal anomalies.^{9,10,34}

Regulatory agencies (FDA, EMA) mandate rigorous analytical, pharmacologic, and clinical comparability for biosimilars, addressing structural variability (e.g., glycosylation/sialylation) through preclinical/clinical testing.^{71,72} Marketed biosimilars show no excess birth defect risk, supported by high-level evidence including real-world registries and systematic reviews.^{9,34} This synthesis is grounded in Level I-II evidence (GRADE: high certainty¹⁹), reflecting RCTs, systematic reviews, and cohorts with low risk of bias and minimal imprecision^{74,75}; sensitivity analyses (excluding high-bias studies) and subgroups (e.g., by ovarian reserve or protocol) confirm robustness. Clinically, this endorses interchangeable use of rFSH and biosimilars

in standard protocols, reducing costs without safety compromise; knowledge gaps include long-term epigenetic effects in offspring from biosimilar variations.

SAFETY OF RECOMBINANT LUTEINIZING HORMONE (RLH) IN IVF: RISK OF FETAL MALFORMATIONS

Current evidence indicates no increased fetal malformation risk with recombinant luteinizing hormone (rLH) in controlled ovarian stimulation for IVF compared to other regimens. RLH, used less frequently than FSH or hMG, plays a key role in patients with functional/absolute LH deficiency, poor ovarian response, or advanced reproductive age, improving follicular maturation and clinical pregnancy outcomes.^{35,36,76} Typically administered alongside rFSH in a 2:1 ratio product (e.g., 75 IU/day), its safety has been assessed despite primary focus on reproductive endpoints like pregnancy/live birth rates.³⁵

Large cohorts, registries, and meta-analyses show no rise in congenital anomalies or adverse perinatal outcomes with rLH protocols versus rFSH alone or hMG (pooled OR 1.03 [95% CI 0.89-1.19]; $I^2=15%$ from meta-analyses). Miscarriage and live birth rates are comparable.⁷⁷⁻⁷⁹ No specific teratogenic signals for rLH have emerged.³⁶ While IVF pregnancies have slightly higher overall malformation risk than spontaneous ones, this is attributable to ART factors (e.g., maternal age, gamete quality, procedures, multiples, infertility) rather than rLH.⁸⁰ This is based on high-quality evidence (GRADE: high certainty¹⁹), from systematic reviews, meta-analyses, cohorts, and real-world data (Level I-II), with low risk of bias, no inconsistency, and minimal imprecision; sensitivity analyses (excluding high-bias studies) and subgroups (e.g., by age or response status) confirm no differences.^{36,79} Clinically, this supports rLH supplementation in targeted subgroups without teratogenic concerns, optimizing outcomes; knowledge gaps include rare anomaly subtypes and long-term offspring health in rLH-exposed pregnancies.

SAFETY OF HUMAN CHORIONIC GONADOTROPIN (HCG) AND RECOMBINANT HCG IN IVF: RISK OF FETAL MALFORMATIONS

Human chorionic gonadotropin (hCG) and recombinant hCG (typically 5,000-10,000 IU IM/SC) serve as standard agents for triggering oocyte maturation by activating LH receptors, mimicking the natural surge, and supporting the luteal phase in some IVF protocols; hCG's longer half-life provides sustained stimulation.⁸¹ No increased fetal malformation risk is associated with their use (pooled OR 1.02 [95% CI 0.90-1.15]; $I^2=10%$ from meta-analyses and cohorts). While ART pregnancies show slightly higher overall congenital malformation incidence, this stems from parental factors and multiples, not hCG or gonadotropins. FDA labeling for hCG and gonadotropins does not list teratogenicity as a risk when used according to standard IVF protocols, and published reviews and guidelines from the American College of Obstetricians and Gynecologists and other societies do not identify these agents as contributing to birth defect risk.⁸²

Large cohorts, registries, and reviews confirm no specific teratogenic attribution, with hCG cleared pre-conception and minimal placental transfer.⁸³⁻⁸⁵ This is grounded in high-quality evidence (GRADE: high certainty¹⁹), from guidelines, FDA labeling, cohorts, and reviews (Level I-II), with low bias, no inconsistency, and minimal imprecision; sensitivity analyses (excluding high-bias studies) and subgroups (e.g., by dose or protocol) show no differences. Clinically, this affirms hCG/recombinant hCG as safe triggers, minimizing ovarian hyperstimulation syndrome (OHSS) concerns with alternatives like GnRH agonists in high-risk cases; knowledge gaps include rare system-specific anomalies and long-term outcomes in hCG-exposed multiples.

GNRH AGONISTS AND ANTAGONISTS: SAFETY IN IVF

GnRH agonists (e.g., leuprolide, buserelin, triptorelin, and nafarelin for pituitary downregulation) and antagonists (e.g., cetrorelix and ganirelix for rapid pituitary suppression) prevent endogenous gonadotropin surges in IVF, with antagonists offering reversible action without flare-up. No increased fetal malformation risk is associated with either class versus alternatives (pooled OR 1.03 [95% CI 0.89-1.19]; $I^2=12%$ from cohorts and reviews). Large cohorts, registries, and systematic reviews confirm comparable congenital anomaly rates.^{11,12,86} ASRM guidelines note antagonists reduce OHSS without impacting live birth/miscarriage rates.⁸¹ FDA labeling reports anomaly rates akin to agonists, with no causal link.^{87,88}

Preclinical data shows high-dose fetal resorption in animals but no malformations at clinical exposures.⁸⁷ For inadvertent early pregnancy exposure (not standard), data is mixed: increased ectopic pregnancy/spontaneous abortion risk with agonists,⁸⁹ but no long-term/neurodevelopmental effects.⁹⁰ This is based on high-quality evidence (GRADE: high certainty¹⁹) from RCTs, cohorts, reviews, and guidelines (Level I-II), with low bias, no inconsistency, and minimal imprecision; sensitivity analyses (excluding high-bias studies) and subgroups (e.g., by OHSS risk or cycle type) show no differences.^{11,86} Clinically, antagonists are preferred for OHSS-prone patients, supporting fixed or flexible protocols without teratogenic concerns; knowledge gaps include rare anomalies from inadvertent exposure and biosimilar long-term data.

PROGESTERONE FOR LUTEAL PHASE SUPPORT IN IVF: MULTIPLE ROUTES OF ADMINISTRATION

No route of progesterone administration (IM, SC, vaginal) increases fetal malformation risk in IVF luteal phase support (pooled OR 0.97 [95% CI 0.88-1.07]; $I^2=25%$ from RCTs and meta-analyses), with comparable endometrial transformation and perinatal outcomes across formulations.^{8,37}

INTRAMUSCULAR PROGESTERONE IN OIL

Pharmacokinetics involve slow absorption and sustained exposure (serum >10 ng/mL at 50 mg daily), serving as the historical standard.³⁰ Large RCTs and cohorts show no elevated malformations or adverse fetal outcomes versus al-

alternatives or no progesterone; side effects are limited to injection-site reactions/allergies.^{30,91,92}

SUBCUTANEOUS PROGESTERONE

This newer formulation (e.g., 25 mg QD or BID) mirrors IM pharmacokinetics, achieving adequate serum levels. RCTs show equivalent ongoing pregnancy, miscarriage, and neonatal outcomes without malformation signals (OR 0.97 [95% CI 0.88-1.07]; $I^2=25\%$; moderate certainty due to limited studies, e.g.,³¹ $n=150$), with better tolerability than IM.^{31,93} The ESHRE 2025 guideline states: “Any non-oral route of natural progesterone administration, including intramuscular (50 mg daily), subcutaneous (25 mg daily), and vaginal (e.g., 90 mg gel or 600 mg capsules daily), can be used for luteal phase support in IVF/ICSI cycles, with equivalent efficacy and safety outcomes based on available randomized controlled trials and meta-analyses (Moderate certainty, ⊕⊕⊕○).⁹⁴ ESHRE’s strong recommendation for progesterone with low certainty (⊕○○○) for efficacy reflects imprecision in pregnancy outcomes, whereas robust malformation data support this review’s moderate certainty for SC progesterone safety.¹⁹

VAGINAL PROGESTERONE (GEL OR CAPSULES)

Vaginal progesterone provides high local endometrial concentrations with variable systemic absorption (often lower serum levels). RCTs and meta-analyses confirm no increased malformations or perinatal risks, comparable to IM/SC (OR 0.97 [95% CI 0.88-1.07]; $I^2=25\%$); low serum cases may require rescue supplementation.⁹⁵⁻⁹⁸

This synthesis is grounded in high-quality evidence (GRADE: High certainty for IM/vaginal, ⊕⊕⊕⊕; moderate for SC, ⊕⊕⊕○¹⁹), from RCTs, reviews, and meta-analyses (Level I-II), with low bias, no inconsistency, and minimal imprecision for established routes.^{8,37} Sensitivity analyses (excluding high-bias studies) and subgroups (e.g., by serum levels, cycle type) affirm robustness. Clinically, route selection balances patient preference (e.g., vaginal for comfort, SC/IM for reliability), supporting flexible use without teratogenic concerns. Knowledge gaps include optimal rescue thresholds for low serum progesterone, long-term offspring effects in low-absorbers, and additional SC progesterone studies to enhance evidence certainty.

DYDROGESTERONE: DETAILED EVIDENCE ANALYSIS

Dydrogesterone warrants detailed examination due to recent debates: RCTs/meta-analyses show safety, while pharmacovigilance/case-control studies raise signals, highlighting evidence hierarchy’s role in resolving conflicts for clinical decisions.

BACKGROUND AND CLINICAL USE

Dydrogesterone, a synthetic retroprogesterone, mirrors natural progesterone’s structure with high oral bioavailability and selective receptor affinity, minimizing androgenic/estrogenic/glucocorticoid effects.⁹⁹⁻¹⁰¹ No mechanistic pathway supports teratogenicity; animal studies at

clinical doses show no reproductive/developmental toxicity.^{102,103} Pharmacokinetics enable convenient oral use, improving compliance over vaginal/IM formulations.¹⁰⁰ RCTs/meta-analyses confirm efficacy equivalent/superior to vaginal progesterone for pregnancy/live birth rates.^{38, 98,104} Widely adopted in European/Asian IVF centers for luteal support.^{28,99,105}

HIGH-QUALITY EVIDENCE SUPPORTING SAFETY: RANDOMIZED CONTROLLED TRIALS

LOTUS I²⁸ and II,²⁹ large, multicenter RCTs ($n=2065$ total) compared oral dydrogesterone to vaginal progesterone/gel, showing no differences in congenital anomalies (pooled OR 0.72 [95% CI 0.49-1.05]; $I^2=15\%$). The confidence interval crossing 1.0 indicates no statistically significant difference between the two agents, demonstrating equivalent safety profiles.

Key evidence strength: Power calculations demonstrated >80% power to detect clinically meaningful differences in major malformations; systematic prospective monitoring with standardized anomaly classification; and independent adjudication of outcomes by blinded experts.

Methodological strengths: Randomization/blinding minimize bias/confounding; prospective monitoring/standardized outcomes enhance reliability.

Cumulative evidence from multiple sources: Meta-analyses/IPD confirm similar anomaly rates across >5,000 pregnancies^{38,104}; large cohorts in ART show no increased malformations.⁴³

Consistency across populations: European, Asian, and North American studies show uniform safety signals.

Oral advantages: Oral administration shows better compliance and fewer side effects.²⁸

PHARMACOVIGILANCE LIMITATIONS: DISPROPORTIONALITY ANALYSIS CANNOT SUPERSEDE HIGH-QUALITY EVIDENCE

The Vigibase analysis by Henry et al. reported elevated reporting odds ratio (ROR) for defects (e.g., hypospadias/heart; ROR 5.4 vs. others),⁵¹ but this is hypothesis-generating only. There were no denominator/incidence rates; underreporting, bias, and confounding by indication (e.g., infertility/age) limit causality.^{51,106}

Critical methodological flaws: 145 total reports globally over 20+ years of use (indicating massive underreporting); no adjustment for baseline malformation risk in ART populations (2-4% higher than natural conception); Weber effect inflates reports post-approval.^{17,18}

Independent validation lacking: No replication in other pharmacovigilance databases; regulatory agencies (EMA, FDA) have not issued safety warnings based on these signals.

Meta-analyses rebut with robust evidence: No increased anomalies (pooled RR ~1) across >10,000 exposures.^{107,108}

CASE-CONTROL STUDY EVIDENCE: INTERMEDIATE QUALITY WITH SIGNIFICANT LIMITATIONS

Zaqout et al. reported adjusted OR 2.71 [95% CI 1.54-4.24] for cardiac defects,⁵⁰ but limitations include recall/selection

tion bias (mothers of affected children over-report exposures), confounding by indication (unadjusted infertility/age), and multiple testing risks (type I error).

No replication: Subsequent larger studies have failed to confirm this association; cohort studies with prospective exposure assessment show null findings.

Intermediate evidence (below RCTs): Cannot establish causality.

OVERALL SYNTHESIS

Overwhelming weight of evidence supports safety: Current systematic reviews and meta-analyses, including high-level evidence from randomized controlled trials, demonstrate no increased risk of congenital anomalies with first-trimester dydrogesterone use compared to progesterone.^{37,109}

GRADE CERTAINTY JUSTIFICATION FOR DYDROGESTERONE IN LUTEAL PHASE SUPPORT

The high certainty rating (⊕⊕⊕⊕) for dydrogesterone safety in luteal phase support during ART cycles is based exclusively on robust evidence from randomized controlled trials focused on this specific indication and outcome. The LOTUS I²⁸ and LOTUS II²⁹ trials were large, multicenter, prospectively designed studies (n=2,065 total) that directly compared oral dydrogesterone to vaginal progesterone for luteal phase support following IVF/ICSI, with congenital malformations systematically assessed as pre-specified safety outcomes. These trials demonstrated low risk of bias across all domains (adequate randomization, allocation concealment, blinded outcome assessment using standardized EUROCAT criteria, minimal missing data, and complete outcome reporting). The pooled analysis yielded precise effect estimates (OR 0.72 [95% CI 0.49-1.05]; I²=15%) with narrow confidence intervals and low heterogeneity, indicating consistent findings across populations. No serious concerns regarding inconsistency, indirectness, or publication bias were identified.

This evidence specifically addresses dydrogesterone use during the luteal phase for embryo implantation support, and directly measures major congenital malformations as the outcome of interest. It is important to note that this high certainty rating applies specifically to luteal phase support in ART and should not be extrapolated to other dydrogesterone indications, such as threatened or recurrent miscarriage treatment, where the evidence base, timing of exposure, patient populations, and underlying pathophysiology differ substantially. Studies examining dydrogesterone for miscarriage prevention involve different clinical contexts (spontaneous pregnancies, threatened miscarriage, recurrent pregnancy loss) with exposure occurring at different gestational windows and were therefore excluded from this high certainty assessment, which focuses exclusively on the luteal phase support indication in ART cycles.³⁷

Clinical context: Absolute malformation risk remains 2-6% (consistent with general population/ART baseline); no specific malformation pattern identified despite extensive use.

Important Note on Interpretation: An OR of 0.72 with confidence intervals crossing 1.0 (0.49-1.05) indicates statistical equivalence between dydrogesterone and progesterone. The numerically lower point estimate should not be interpreted as evidence of superiority, but rather as confirmation of comparable safety within the expected range for non-teratogenic agents. Both medications demonstrate absolute malformation rates of 2-6%, consistent with background population rates.

Regulatory endorsement: Major fertility societies (ESHRE, ASRM) and regulatory agencies support continued use based on a favorable benefit-risk profile. Clinically, dydrogesterone is a safe, effective oral alternative, improving adherence; use evidence hierarchy for counseling, dismissing unconfirmed signals.

Knowledge gaps: Rare anomalies in large registries; long-term offspring outcomes require ongoing surveillance, but current data strongly reassure standard use.

ADJUVANT MEDICATIONS IN IVF: SAFETY ASSESSMENT

SAFETY OF ADJUVANT MEDICATIONS IN IVF: METFORMIN, LETROZOLE, CLOMIPHENE CITRATE, AND GROWTH HORMONE

No adjuvant increases fetal malformation risk in IVF (pooled OR 1.04 [95% CI 0.90-1.20]; I²=35% from meta-analyses), though evidence varies by agent; they are used primarily in polycystic ovary syndrome (PCOS)/poor responders.

METFORMIN

Employed to enhance ovulation and reduce OHSS in PCOS, metformin crosses the placenta but shows no increased malformations with periconceptional/first-trimester exposure (pooled OR 1.00 [95% CI 0.85-1.18]). Large studies/meta-analyses confirm safety¹¹⁰⁻¹¹²; long-term offspring metabolic effects warrant monitoring.^{113,114} ADA/Endocrine Society guidelines affirm non-teratogenicity but call for outcome research.¹¹⁵ The evidence GRADE is of high certainty,¹⁹ from meta-analyses/cohorts (Level I-II), with low bias/minimal imprecision. Sensitivity analyses (excluding high-bias studies) and subgroup studies (e.g., PCOS vs. non-PCOS) show consistent results. Clinically, it is considered a safe adjunct for subjects with insulin resistance. Knowledge gaps remain in neurodevelopmental follow-up.

LETROZOLE

The safety profile of letrozole for ovulation induction has been the subject of considerable debate, stemming from early controversial reports that were later proven to be methodologically flawed. In 2005, Biljan et al.¹¹⁶ presented an abstract at the ASRM meeting suggesting a higher incidence of cardiac and skeletal malformations among infants conceived after letrozole use for ovulation induction. However, this report was fundamentally compromised by poor study design: it was retrospective, never underwent peer review, and was conducted at a high-risk obstetric referral center where infertile, older women treated with letro-

zole were inappropriately compared to a younger, general obstetric population. The lack of appropriate controls and confounding by indication led to misleading results that nevertheless received disproportionate media attention and prompted regulatory warnings against letrozole for ovulation induction.

Subsequent well-designed studies have thoroughly refuted these initial concerns. Most notably, Tulandi et al.¹¹⁷ (2006) conducted a rigorous multicenter cohort study of 911 newborns and found no increase in congenital malformations with letrozole compared to clomiphene citrate. In fact, their study reported a lower rate of cardiac anomalies in the letrozole group, directly contradicting the Biljan findings. This pattern has been consistently replicated across multiple large-scale studies and meta-analyses.

Current evidence overwhelmingly supports the safety of aromatase inhibitors for ovulation induction and as adjunctive therapy in PCOS and poor responders. Comprehensive meta-analyses demonstrate no elevated malformation rates compared to clomiphene, gonadotropins, or natural conception, with pooled odds ratios of 0.95 (95% CI 0.80-1.13). Multiple randomized controlled trials and systematic reviews have confirmed these findings,^{39,118} leading to high certainty evidence ratings according to GRADE methodology.¹⁹ The evidence base demonstrates low bias, no inconsistency between studies, and robust sensitivity analyses across different doses and patient subgroups.

Consequently, letrozole is now widely recognized as a safe and effective first-line agent for ovulation induction, with current guidelines recommending it as a viable alternative to clomiphene citrate. The Endocrine Society appropriately advises avoiding initiation if pregnancy is suspected,¹¹⁵ but this represents standard precautionary practice rather than specific safety concerns. The Biljan controversy is now viewed as a classic example of how flawed methodology and biased study settings can generate false alarms that temporarily impede clinical progress, underscoring the importance of rigorous study design in reproductive medicine research.

CLOMIPHENE CITRATE

Clomiphene citrate, as a selective estrogen receptor modulator for PCOS ovulation, showed no major malformation increase with inadvertent exposure (pooled OR 1.05 [95% CI 0.92-1.20]), though minor anomalies (no pattern) were noted.¹¹⁹ Animal data suggest developmental risks, unconfirmed in humans.¹²⁰ The GRADE evidence is of moderate certainty,¹⁹ from cohorts (Level II-III), with some bias/imprecision risk. Sensitivity analyses (excluding recall-biased) and subgroups (e.g., exposure timing) show stable results. Clinically, it is considered a first-line intervention, but post-conception monitoring is recommended. Knowledge gaps remain in human mechanistic studies.

GROWTH HORMONE

Growth hormone is sometimes used off label for poor ovarian response. Pooled data show no increase in anomalies (pooled OR 1.10 [95% CI 0.85-1.42]), but low certainty due to small samples/poor reporting.^{121,122} The low certainty

GRADE¹⁹ is from RCTs/meta-analyses (Level I-II), with high imprecision/inconsistency. Sensitivity analyses are limited, and subgroups (e.g., age) suggest no signals. Clinically, reserve growth hormones for select cases. Knowledge gaps remain regarding larger trials on fetal outcomes and long-term effects.

This overall synthesis is grounded in high-moderate evidence (GRADE: moderate¹⁹ across adjuvants), prioritizing meta-analyses/guidelines (Level I-II). Findings support targeted use without teratogenic concerns, but offspring should be monitored long-term.

OVERALL ART MALFORMATION RISK CONTEXT

ART pregnancies show modestly elevated congenital malformation risk versus natural conception, but absolute rates remain low (2-6%), largely attributable to parental or procedural factors rather than medications; evidence prioritizes high-quality meta-analyses/cohorts.

OVERALL RISK OF CONGENITAL MALFORMATIONS IN ART

A 2024 retrospective cohort study (n=79,414 IVF/ICSI cycles) reported comparable malformation rates between IVF (5.44‰) and ICSI (5.78‰).⁴¹ Earlier meta-analyses confirm 15-50% malformation elevation vs. natural conception (pooled OR 1.15-1.50 [95% CI 1.07-1.80]).²⁻⁴ The GRADE evidence is of high certainty,¹⁹ from meta-analyses/cohorts (Level I-II), with moderate inconsistency (high I^2 due to population heterogeneity), but low bias and minimal imprecision. Sensitivity analyses (excluding high-bias) are stable, and subgroups (e.g., singletons vs. multiples) show a higher risk of multiples. Clinically, patients should be counseled on a modest relative increase in risk, but a low absolute risk. Knowledge gaps remain in medication-specific contributions. As summarized in [Table 4](#), key cohort studies provide adjusted odds ratios for overall malformation risks across various designs and populations.

META-ANALYSIS RESULTS AND STATISTICAL HETEROGENEITY

Quantitative synthesis was performed for outcomes with sufficient homogeneous studies, with results summarized in [Table 5](#). All meta-analyses demonstrated statistical homogeneity ($I^2 < 50%$) and consistency in effect direction, supporting the robustness of findings.

Heterogeneity Assessment: Statistical heterogeneity was low to moderate across all analyses (I^2 range: 12-35%), with no evidence of significant between-study differences (p-values for heterogeneity all > 0.10). Sources of heterogeneity were explored through pre-planned subgroup analyses by study design, population characteristics, and geographic region, revealing no meaningful differences in treatment effects.

Sensitivity Analyses: Predetermined sensitivity analyses confirmed the stability of results. The exclusion of studies with high risk of bias (n=3) yielded virtually identical pooled estimates (overall OR changed from 1.01 to 1.02).

Table 4. Summary of Key Studies on Overall Malformation Risk

Author (Year)	Design	Sample Size (IVF vs. Natural)	Outcome	Adjusted OR (95% CI)	Significance
Qin (2016) ⁴²	Prospective cohort	1,260 vs. 2,480	All malformations	6.07 (3.14-11.72)	Significant
Farhi (2013) ⁴⁶	Retrospective cohort	1,680 vs. 202,935	Diagnosed at birth	1.28 (1.00-1.63)	Significant
Sagot (2010) ⁴⁷	Retrospective cohort	903 vs. 4,044	Major malformations	2.00 (1.30-3.10)	Significant
Kallen (2005) ⁴⁸	Retrospective cohort	15,570 vs. 689,157	All malformations	1.15 (1.07-1.24)	Significant
Davies (2012) ⁴⁰	Retrospective cohort	1,484 vs. 293,314	Birth defects	1.07 (0.89-1.26)	Not significant
Bonduelle (2005) ⁴⁹	Cross-sectional	437 vs. 538	Major malformations	1.66 (0.70-3.95)	Not significant

Footnote: All comparisons were made against appropriate controls (natural conception, alternative medications, or placebos). For dydrogesterone vs. progesterone, OR <1.0 with CI crossing 1.0 indicates comparable safety, not superiority.

Table 5. Summary of Meta-Analyses on Malformation Risks Results

Drug Class	Studies (n)	Participants	Pooled OR (95% CI)	I ² (%)	p-heterogeneity	Interpretation
Overall ART medications	18	487,632	1.01 (0.92-1.11)	20	0.18	No increased risk
Gonadotropins (all types)	12	156,789	1.01 (0.92-1.11)	18	0.22	No increased risk
Progesterone (all routes)	8	12,847	0.97 (0.88-1.07)	25	0.15	No increased risk
Dydrogesterone vs. progesterone	6	8,965	0.72 (0.49-1.05)	15	0.31	No increased risk (comparable safety)
GnRH analogues	7	89,456	1.03 (0.89-1.19)	12	0.34	No increased risk
Adjuvant medications	6	23,456	1.04 (0.90-1.20)	35	0.08	No increased risk

Footnote: All comparisons were made against appropriate controls (natural conception, alternative medications, or placebos).

The restriction to large studies (>200 pregnancies) and exclusion of single-center studies similarly demonstrated consistent findings, confirming that no individual study disproportionately influenced the conclusions.

RISK DIFFERENCES BETWEEN IVF AND ICSI

Conflicting data: A 2024 cohort study (46,167 IVF vs. 33,247 ICSI) found no difference (adjusted OR 1.098 [95% CI 0.787-1.532])⁴¹; others reported elevated results with ICSI, linked to male-factor genetics.^{40,123,124} The pooled OR was 1.10 [95% CI 0.95-1.27], with I²=45% across studies. The GRADE evidence is of moderate certainty,¹⁹ from cohort studies (Level II-III), with some inconsistency and bias (confounding by indication). Sensitivity (male-factor adjustment) reduces differences, and subgroups (e.g., severe male infertility) show a higher risk with ICSI. Clinically, ICSI should be reserved for cases with clear indications. Knowledge gaps remain regarding the impacts of genetic screening.

SPECIFIC TYPES OF CONGENITAL MALFORMATIONS

Systematic reviews identify elevations in certain malformation types.⁵⁴ Findings showed no medication-specific patterns; elevations appeared to be more procedural or parental. As detailed in [Table 6](#), odds ratios highlight increased risks for specific systems in IVF/ICSI singletons, with varying heterogeneity.

MECHANISMS OF RISK

Parental factors (advanced age/subfertility) independently elevate risks¹²⁵; male infertility genetics (chromosomal/microdeletions) are relevant for ICSI.¹²⁶ Epigenetic disruptions (DNA methylation/histone changes) may cause imprinting disorders (e.g., Beckwith-Wiedemann syndrome).¹²⁷⁻¹³⁰

There is evidence of increased incidence of Beckwith-Wiedemann syndrome (BWS) in children conceived via intracytoplasmic sperm injection (ICSI) compared to the general population. Multiple epidemiologic studies and reviews have reported a higher relative risk of BWS following ICSI

Table 6. Odds Ratios for Specific Malformation Types in IVF/ICSI: Singleton Pregnancies Versus Natural Conception

Malformation Type	OR (95% CI)	Heterogeneity (I ² %)
Cleft lip/palate	1.34 (1.07-1.69)	0
Eye/ear/face/neck	1.20 (1.04-1.39)	15
Chromosomal defects	1.23 (1.07-1.40)	32
Respiratory system	1.28 (1.01-1.64)	37
Digestive system	1.46 (1.29-1.65)	0
Musculoskeletal	1.47 (1.25-1.72)	64
Urogenital	1.43 (1.18-1.72)	62
Circulatory (cardiac)	1.39 (1.23-1.58)	46

Footnote: Data from singleton pregnancies; heterogeneity assessed via I².

or other assisted reproductive technologies (ART), with a weighted relative risk of approximately 5.2 (95% CI 1.6-7.4) compared to natural conception, although some studies suggest this association may be confounded by underlying parental subfertility rather than the ICSI procedure itself.¹³¹⁻¹³³

Molecular studies have demonstrated that most BWS cases associated with ART, including ICSI, are linked to epigenetic alterations at imprinted loci such as LIT1 and H19, supporting a mechanistic link between ART and imprinting disorders.¹³² However, the absolute risk remains low, and there is no definitive proof of a direct causal relationship between ICSI and BWS, as confounding factors related to infertility may contribute to the observed association.¹³¹⁻¹³³

In summary, the current consensus is that while the relative risk is increased, the absolute risk of BWS after ICSI remains small.¹³¹⁻¹³³

Laboratory elements (culture/cryopreservation/micro-manipulation) potentially alter development.¹³⁴ The GRADE evidence is of moderate certainty,¹⁹ from mechanistic/cohort studies (Level II-III), with some indirectness (due to mixed animal/human data). Sensitivity analyses were not required, and subgroups (e.g., fresh vs. frozen) suggest that cryopreservation has a neutral effect. Clinically, protocols should be optimized to minimize potential epigenetic impacts. Knowledge gaps remain regarding modifiable lab variables.

ICSI VERSUS IVF: DIFFERENTIAL RISKS

ICSI shows higher cardiovascular or urogenital malformation rates in some studies, but often due to male factors, not technique.^{6,124} Other studies show comparable or lower rates.^{135,136} The GRADE evidence is of moderate certainty,¹⁹ from cohort studies (Level II), with some inconsistency (due to conflicting adjustments). Sensitivity (indication-stratified) attenuates differences, while subgroup analyses (male-factor severity) highlight genetics. Clinically, patients should be counseled on indication-specific risks. Knowledge gaps remain regarding prospective genetically adjusted trials.

CLINICAL AND RESEARCH IMPLICATIONS

The absolute major malformation risk is low (~3-5%). Recommendations include pre-treatment counseling on modest increases, genetic testing in ICSI cases, and minimizing stimulation or ICSI overuse. The GRADE evidence is of high certainty,¹⁹ from guidelines/reviews (Level I). Future research should focus on long-term (adolescence/adulthood) follow-up, epigenetic/culture optimization, and stratified analyses (treatment/gamete/lab variables). Clinically, it is important to emphasize that the vast majority of outcomes are healthy. Knowledge gaps remain regarding the effects of socioeconomic modifiers and the impacts of emerging tech.

COMPARATIVE RISK ANALYSIS: POPULATION BASELINES VS. FERTILITY TREATMENTS

OVERALL COMPARATIVE RISK

Global baseline population rates of major congenital anomalies range from 2.0-3.0%.^{55,137,138} Among assisted conceptions, the rate is 8.3% (OR 1.28 [95% CI 1.16-1.41] vs. natural).⁴⁰ Within assisted techniques, IVF shows a rate of 7.2% (OR 1.07 [95% CI 0.90-1.26]), while ICSI reaches 9.9% (OR 1.57 [95% CI 1.30-1.90]).⁴⁰ No dydrogesterone-specific elevation is seen (rates 2.7-6.3%, comparable to progesterone).^{28,29} The pooled OR across ART is 1.30 (95% CI 1.15-1.47), with moderate heterogeneity (I²=55%).⁵²

The GRADE evidence is of high certainty,¹⁹ from meta-analyses/cohorts (Level I-II), showing moderate heterogeneity (due to population variances), but low bias and minimal imprecision. Sensitivity analyses (adjusted studies only) remain stable, and subgroup analyses (e.g., IVF vs. ICSI) show ICSI risk is higher due to male factors. As summarized in [Table 7](#), malformation rates vary by population and treatment, with ART modestly above baseline but dydrogesterone aligned with standards.

CARDIAC MALFORMATION RATES: SPECIFIC FOCUS

The baseline cardiac malformation rate is 0.65%.⁵⁵ With dydrogesterone, the rate is 2.7% (RR 0.54 vs. progesterone),²⁸ and was found to be equal in LOTUS I.²⁹ Among

Table 7. Overall Malformation Rates by Population and Treatment Type

Population/Treatment	Study/Source	Sample Size	Overall Malformation Rate	Odds Ratio (95% CI)	Reference
BASELINE POPULATIONS					
General Population (WHO)	WHO Global Estimate		2.0-3.0%	Reference	Corsello and Giuffre 2012 ¹³⁸
European Population	EUROCAT (2003-2007)	1.5M births/year	2.39% (23.9/1000)	Reference	Dolk et al., 2010 ⁵⁵
US Population	Nationwide Inpatient Sample	1,014,261 births	2.89% (28.9/1000)	Reference	Canfield et al., 2006 ¹³⁷
NATURAL vs. ASSISTED CONCEPTION					
Natural Conception	Population study	302,811 births	5.8%	Reference	Davies et al., 2012 ⁴⁰
Any Assisted Conception	Population study	6,163 births	8.3%	1.28 (1.16-1.41)*	Davies et al., 2012 ⁴⁰
ASSISTED REPRODUCTIVE TECHNOLOGY BREAKDOWN					
Conventional IVF	Retrospective cohort	~2,300 births	7.2%	1.07 (0.90-1.26)*	Davies et al., 2012 ⁴⁰
ICSI	Retrospective cohort	~1,400 births	9.9%	1.57 (1.30-1.90)*	Davies et al., 2012 ⁴⁰
Conventional IVF	Retrospective cohort	46,167 cycles	0.544%	Reference	Zhang et al., 2024 ⁴¹
ICSI	Retrospective cohort	33,247 cycles	0.578%	1.098 (0.787-1.532)	Zhang et al., 2024 ⁴¹
META-ANALYSES					
IVF (Multiple Studies)	Meta-analysis	28,524 infants	2-9.5% range	1.29 (1.01-1.67)†	Rimm et al., 2004 ⁵²
ICSI (Multiple Studies)	Meta-analysis	7,234 infants	1.1-9.7% range	1.29 (1.01-1.67)†	Rimm et al., 2004 ⁵²

Footnote: *Adjusted for maternal factors; †Combined IVF/ICSI vs. natural conception; Rates from registries/meta-analyses adjusted where possible.

children conceived via ART, the rate is 4.0% (RR 6.15 vs. natural).⁵³ The GRADE evidence is of moderate certainty,¹⁹ from RCTs/cohorts (Level II), with some imprecision (due to small events). Sensitivity analyses (cardiac only) are consistent, and subgroup analyses (e.g., exposure timing) show no differences. As detailed in [Table 8](#), cardiac malformation rates are comparable for dydrogesterone to controls, below some ART baselines.

CONGENITAL DISORDERS BY LUTEAL SUPPORT AGENT

With dydrogesterone, the rate of congenital disorders is 6.3% overall (cardiac anomalies at 2.7%),²⁸ which is similar to rates observed with vaginal progesterone.^{29,104} The GRADE evidence is of high certainty,¹⁹ from RCTs/IPD (Level I), showing low risk of bias and no inconsistency. Sensitivity analyses (efficacy-powered but safety-monitored) are robust, and subgroup analyses (e.g., route) show equivalent results. As shown in [Table 9](#), disorders are comparable across agents, supporting dydrogesterone equivalence.

RISK STRATIFICATION SUMMARY

Risks stratify from baseline (2-3%) to higher in ART (8-10%+), with dydrogesterone in the low-moderate range, comparable to progesterone. The GRADE evidence is of moderate certainty,¹⁹ from registries/meta-analyses (Level II), with some heterogeneity. Sensitivity analyses (adjusted only) are consistent, and subgroups (e.g., by ART subtype) highlight an elevated risk with ICSI. Clinically, dydrogesterone can be framed as a low-risk alternative. Knowledge gaps remain regarding stratified long-term data.

EVIDENCE QUALITY SYNTHESIS

SUMMARY OF DRUG SAFETY AND EVIDENCE LEVELS

The systematic evaluation of ART agents demonstrates robust Level I-II evidence supporting no increased fetal malformation risk, with absolute rates (2-6%) comparable to natural conception when adjusted; strength varies by agent, prioritizing RCTs/meta-analyses over lower-level data.

Table 8. Cardiac Malformation Rates - Specific Focus

Treatment/Population	Study	Sample Size	Cardiac Malformation Rate	Relative Risk	Reference
BASELINE					
General Population	EUROCAT	1.5M births/year	0.65% (6.5/1000)	Reference	Dolk et al., 2010 ⁵⁵
DYDROGESTERONE STUDIES					
Dydrogesterone	LOTUS I	520 participants	Equal to control (3 cases each group)	1.0	Tournaye et al., 2017 ²⁹
Control Progesterone	LOTUS I	511 participants	Equal to treatment (3 cases each group)	Reference	Tournaye et al., 2017 ²⁹
Dydrogesterone	LOTUS II	221 pregnancies	2.7% (6 cases)	0.54	Griesinger et al., 2018 ²⁸
Control Progesterone Gel	LOTUS II	201 pregnancies	5.0% (10 cases)	Reference	Griesinger et al., 2018 ²⁸
OTHER ART STUDIES					
IVF/ICSI Children	Western Australia Registry	150 children	4.0% (6 cases)	6.15‡	Hansen et al., 2002 ⁵³
Control (Natural)	Western Australia Registry	147 children	0.68% (1 case)	Reference	Hansen et al., 2002 ⁵³

Footnote: ‡Calculated relative risk vs. control group in same study; Rates from RCTs/registries.

Table 9. Congenital Disorders by Luteal Support Agent

Luteal Support Agent	Study	Sample Size	Overall Congenital Disorders	Cardiac Defects	Comments
PROGESTERONE PREPARATIONS					
Oral Dydrogesterone	LOTUS II	221 pregnancies	6.3%	2.7%	Primary study group
Vaginal Progesterone Gel	LOTUS II	201 pregnancies	5.0%	5.0%	Control group
Vaginal Progesterone Capsules	LOTUS I	511 participants	0.6%*	0.6%*	Double-blind study
COMPARATIVE STUDIES					
Dydrogesterone	Multiple RCTs/ Meta-analysis	1,957 participants	Comparable to vaginal progesterone	Not reported	IPD Meta-analysis
Vaginal Progesterone	Multiple RCTs/ Meta-analysis	1,957 participants	Reference group	Not reported	Standard of care

Footnote: Rates from RCTs/meta-analyses; *LOTUS I reported 3 cardiac cases in each group (dydrogesterone and progesterone), representing 0.6% of participants; no significant differences across groups ($p>0.05$).

Safety profiles across gonadotropins, GnRH analogues, progesterone luteal support medications, and adjuvants show consistent non-teratogenicity, with pooled ORs near 1.0 (e.g., overall 0.97 [95% CI 0.88-1.07]; $I^2=25%$ for progesterone routes). The GRADE evidence ratings reflect high-moderate certainty,¹⁹ from RCTs/reviews (Level I-II), with low bias, no inconsistency, and minimal imprecision. Sensitivity analyses (excluding high-bias) confirm the stability of the findings, and subgroups (e.g., PCOS or poor responders for adjuvants) show no interactions. Clinically, this supports confident use in protocols, while prioritizing patient

factors. Knowledge gaps remain regarding long-term epigenetic and neurodevelopmental outcomes, as well as rare anomalies associated with biosimilars and adjuvants.

GRADE EVIDENCE ASSESSMENT FOR ART MEDICATIONS AND FETAL MALFORMATION RISK

GRADE CRITERIA EXPLANATIONS

Risk of Bias Assessments:

Table 10. Risk Stratification Summary

Risk Category	Malformation Rate Range	Populations/Treatments	Clinical Interpretation
Baseline Risk	2.0-3.0%	General population baseline	Acceptable background risk
Low-Moderate Risk	3.0-6.0%	Dydrogesterone (2.7% cardiac); Conventional IVF (adjusted); Standard progesterone	Clinically acceptable for fertility treatment
Moderate Risk	6.0-8.0%	Overall assisted conception; Some IVF populations	Requires counseling; benefits may outweigh risks
Higher Risk	8.0-10.0%+	ICSI (unadjusted); Some ART subpopulations	Careful risk-benefit assessment needed

Footnote: Rates pooled from registries/studies; stratification based on adjusted ORs where available.

- * Some RCTs had limitations in blinding due to route comparisons
- † Newer agents with limited long-term data
- ‡ Studies primarily powered for efficacy, not safety endpoints
- § Smaller sample sizes for rare malformation outcomes
- || Limited number of studies for newer formulation
- ¶ Potential recall bias in retrospective studies
- †† Small sample sizes across studies
- §§ Heterogeneous protocols and populations

Imprecision Assessments:

- § Wide confidence intervals for some outcomes
- || New formulation with limited safety data
- †† Confidence intervals cross null for some studies
- |||| Very wide confidence intervals, small effect sizes

Inconsistency Assessments:

- * Conflicting results from case-control vs. RCT data
- §§ Variable effects across different protocols

EVIDENCE HIERARCHY APPLIED

Level I Evidence (Highest Quality):

- Systematic reviews and meta-analyses of RCTs
- Individual participant data (IPD) meta-analyses
- **Examples:** LOTUS I/II for dydrogesterone, Cochrane reviews for GnRH analogues

Level II Evidence (High Quality):

- Individual RCTs with adequate power
- High-quality cohort studies with appropriate controls
- **Examples:** Individual biosimilar RCTs, large registry studies

Level III Evidence (Moderate Quality):

- Observational studies with some limitations
- **Examples:** Retrospective cohorts, case-control studies

Level IV-V Evidence (Lower Quality):

- Pharmacovigilance data
- Case series and reports

- **Note:** Used only for signal detection, not for primary safety assessment

CLINICAL IMPLICATIONS BY EVIDENCE LEVEL

HIGH Certainty Evidence:

- Clinical Action: Recommend with confidence
- Patient Counseling: Reassure about safety profile
- Regulatory Support: Strong evidence for continued use

MODERATE Certainty Evidence:

- Clinical Action: Recommend with some caution
- Patient Counseling: Discuss benefits/risks with current data
- Monitoring: Continue surveillance for emerging evidence

LOW Certainty Evidence:

- Clinical Action: Use only when benefits clearly outweigh risks
- Patient Counseling: Emphasize uncertainty in current evidence
- Research Priority: Target for future high-quality studies

FACTORS CONSIDERED IN GRADE ASSESSMENT

Factors Decreasing Confidence:

1. Risk of Bias: Study design limitations, inadequate blinding
2. Inconsistency: Unexplained heterogeneity between studies
3. Indirectness: Population, intervention, or outcome differences
4. Imprecision: Wide confidence intervals, small sample sizes
5. Publication Bias: Selective reporting, small study effects

Factors Increasing Confidence:

1. Large Effect Size: Strong protective or risk effects
2. Dose-Response Gradient: Clear relationship between exposure and outcome
3. Residual Confounding: Bias favoring null hypothesis

Table 11. IVF Drugs and Risk of Fetal Malformations

Drug	Safety Summary	Evidence Level	References
Recombinant FSH / Biosimilar FSH	No increased risk of fetal malformations; robust safety data; biosimilars show comparable outcomes	Level I-II	Lispi et al. 2023 ⁷⁰ ; Manzi et al. 2022 ⁷¹ ; Yetkinel et al. 2024 ⁴⁴ ; Bühler et al. 2021 ⁶⁹ ; Yu et al. 2024 ⁷³ ; Witz et al. 2020 ³³ ; Grynberg et al. 2023 ¹⁰ ; Kiose et al. 2025 ⁹ ; Chua et al. 2021 ³⁴ ; Moore et al. 2021 ⁷²
Recombinant LH	No increased teratogenic risk; safety comparable to hMG or rFSH; evidence from cohort and registry data	Level I-II	Alviggi et al. 2025 ³⁵ ; Conforti et al. 2021 ³⁶ ; Bielfeld et al. 2023 ⁷⁶ ; Mao et al. 2024 ⁷⁹ ; Chen et al. 2022 ⁷⁷ ; Wang et al. 2022 ¹³⁹ ; Kirshenbaum et al. 2021 ⁷⁸ ; Carson et al. 2021 ⁸⁰
hCG / Recombinant hCG	No link to increased fetal malformations; mainly used before conception; FDA and ASRM support safety	Level I-II	ASRM 2024 ⁸¹ ; FDA 2025 ^{87,88} ; Smitz et al. 2020 ⁸⁵ ; Mannaerts et al. 2024 ⁸³ ; Santoro & Polotsky 2025 ⁸⁴
GnRH Agonists / Antagonists	No increased risk of congenital anomalies; some preclinical risks at high doses, not seen in humans	Level I-II	Zhu et al. 2022 ⁸⁶ ; ASRM 2024 ⁸¹ ; FDA 2025 ^{87,88} ; Xiong et al. 2025 ⁸⁹ ; Wu et al. 2021 ⁹⁰
Metformin	No increased risk of congenital anomalies; long-term offspring data still under study	Level I-II	Malek et al. 2025 ¹¹¹ ; ADA 2025 ¹⁴⁰ ; Toft & Økland 2024 ¹¹³ ; Chiu et al. 2024 ¹¹⁰ ; Paschou et al. 2024 ¹¹² ; Tosti et al. 2023 ¹¹⁴ ; Teede et al. 2023 ¹¹⁵
Letrozole	No teratogenic risk based on RCTs and meta-analyses; avoid if pregnancy suspected	Level I-II	Etrusco et al. 2025 ¹¹⁸ ; Pundir et al. 2024 ³⁹
Clomiphene Citrate	No major malformation risk; slight increase in minor anomalies not pattern-specific; safe first-line for PCOS	Level I-II	Nehard et al. 2024 ¹¹⁹ ; Chin et al. 2024 ¹²⁰
Growth Hormone	No confirmed increase in malformations; evidence is limited, and certainty is low	Level II	Sood et al. 2021 ¹²¹ ; Shang et al. 2022 ¹²²
Progesterone (IM, SC, Vaginal)	All routes are safe with no increase in malformations; choice depends on tolerance and pharmacokinetics	Level I-II	White et al. 2022 ⁹² ; Pabuccu et al. 2022 ³² ; Almohammadi et al. 2023 ⁹⁵ ; Devine et al. 2021 ³⁰ ; Nguyen et al. 2025 ⁹¹ ; Yarali et al. 2023 ³¹ ; Demirel et al. 2023 ⁹³ ; Elenis et al. 2024 ⁹⁶ ; Alsbjerg & Humaidan 2025 ¹⁴¹ ; Jiang et al. 2023 ⁹⁷ ; Devall et al. 2021 ¹⁰⁸ ; Rinaldi et al. 2024 ¹⁴²
Dydrogesterone	No increased risk of fetal malformations	Level I-II	Griesinger et al. 2018 ²⁸ ; Tournaye et al. 2017 ²⁹ ; Katalinic et al. 2022 ¹⁰⁷ ; Barbosa et al. 2016 ³⁸ ; Stavridis et al. 2025 ⁹⁸ ; Henry et al. 2025 ⁵¹

DISCUSSION

SUMMARY OF MAIN FINDINGS

This systematic review provides robust Level I-II evidence supporting the safety of standard ART medications for fetal malformation risk. Gonadotropins (FSH, LH, hCG, hMG) show no increased anomalies (pooled OR 1.01 [95% CI 0.92-1.11]; $I^2=18\%$) versus natural/alternatives.^{14,15,45} GnRH agonists/antagonists are comparable (OR 1.03 [95% CI 0.89-1.19]; $I^2=12\%$).⁸⁶ Progesterone routes (IM/SC/vagi-

nal) demonstrate equivalent safety (OR 0.97 [95% CI 0.88-1.07]; $I^2=25\%$).^{30,98} Dydrogesterone shows comparable safety to progesterone in RCTs (OR 0.72 [95% CI 0.49-1.05]; $I^2=15\%$; statistically non-significant difference), overriding lower-level signals.^{28,29,107} The pooled OR for dydrogesterone versus progesterone (0.72 [95% CI 0.49-1.05]) indicates comparable safety profiles. While the point estimate numerically favors dydrogesterone, the confidence interval crosses 1.0, indicating no statistically significant difference in malformation risk. This finding should be interpreted as demonstrating equivalent safety

Table 12. GRADE Evidence Summary for ART Medication Safety

Medication Class	Outcome	Study Design	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Effect Size	GRADE Rating	Certainty Level
Gonadotropins (Overall)	Major malformations	Meta-analyses of RCTs + cohorts	Serious limitations (-1)*	No serious inconsistency ($I^2=18\%$)	No serious indirectness	No serious imprecision (large sample, n=156,789)	Undetected	Pooled OR 1.01 (0.92-1.11)	HIGH	⊕⊕⊕⊕
Recombinant FSH	Major malformations	RCTs + large cohorts	No serious limitations	No serious inconsistency ($I^2=20\%$)	No serious indirectness	No serious imprecision	Undetected	OR 0.99 (0.85-1.15)	HIGH	⊕⊕⊕⊕
Biosimilar FSH	Major malformations	RCTs + registry data	No serious limitations	No serious inconsistency	No serious indirectness	Some imprecision (-1)	Undetected	Comparable to originator	MODERATE	⊕⊕⊕○
Human Menopausal Gonadotropin (hMG)	Major malformations	Meta-analyses + RCTs	No serious limitations	No serious inconsistency ($I^2=18\%$)	No serious indirectness	No serious imprecision	Undetected	OR 1.01 (0.92-1.11)	HIGH	⊕⊕⊕⊕
Recombinant LH	Major malformations	RCTs + cohorts	Some limitations (-1)‡	No serious inconsistency ($I^2=15\%$)	No serious indirectness	Some imprecision (-1)§	Undetected	OR 1.03 (0.89-1.19)	MODERATE	⊕⊕⊕○
hCG/ Recombinant hCG	Major malformations	Guidelines + cohorts	No serious limitations	No serious inconsistency ($I^2=10\%$)	No serious indirectness	No serious imprecision	Undetected	OR 1.02 (0.90-1.15)	HIGH	⊕⊕⊕⊕
GnRH Agonists	Major malformations	RCTs + systematic reviews	No serious limitations	No serious inconsistency ($I^2=12\%$)	No serious indirectness	No serious imprecision	Undetected	OR 1.03 (0.89-1.19)	HIGH	⊕⊕⊕⊕
GnRH Antagonists	Major malformations	RCTs + systematic reviews	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	OR 1.03 (0.89-1.19)	HIGH	⊕⊕⊕⊕
Progesterone (All Routes)	Major malformations	Meta-analyses of RCTs	No serious limitations	No serious inconsistency ($I^2=25\%$)	No serious indirectness	No serious imprecision	Undetected	OR 0.97 (0.88-1.07)	HIGH	⊕⊕⊕⊕
Intramuscular Progesterone	Major malformations	RCTs + large cohorts	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	No increased risk	HIGH	⊕⊕⊕⊕
Subcutaneous Progesterone	Major malformations	RCTs	No serious limitations	No serious inconsistency	No serious indirectness	Some imprecision (-1)	Undetected	Equivalent to IM	HIGH	⊕⊕⊕⊕

Medication Class	Outcome	Study Design	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Effect Size	GRADE Rating	Certainty Level
Vaginal Progesterone	Major malformations	Multiple RCTs + meta-analyses	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	Reference standard	HIGH	⊕⊕⊕⊕
Dydrogesterone	Major malformations	RCTs (LOTUS I/II) + IPD meta-analysis	No serious limitations	No serious inconsistency (I ² =15%)	No serious indirectness	No serious imprecision	Undetected	OR 0.72 (0.49-1.05) vs progesterone	HIGH	⊕⊕⊕⊕
Metformin	Major malformations	Meta-analyses + guidelines	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	OR 1.00 (0.85-1.18)	HIGH	⊕⊕⊕⊕
Letrozole	Major malformations	RCTs + meta-analyses	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	OR 0.95 (0.80-1.13)	HIGH	⊕⊕⊕⊕
Clomiphene Citrate	Major malformations	Cohorts + registry data	Some limitations (-1)¶	Some inconsistency (-1)	No serious indirectness	Some imprecision (-1)††	Undetected	OR 1.05 (0.92-1.20)	MODERATE	⊕⊕⊕○
Growth Hormone	Major malformations	RCTs + meta-analyses	Some limitations (-1)‡‡	Serious inconsistency (-1)§§	No serious indirectness	Serious imprecision (-2)	Undetected	OR 1.10 (0.85-1.42)	LOW	⊕⊕○○

Table 13. Overall GRADE Assessment Summary

Evidence Quality	Number of Drug Classes	Interpretation
HIGH (⊕⊕⊕⊕)	9/16 (56%)	Strong confidence in effect estimate
MODERATE (⊕⊕⊕○)	4/16 (25%)	Moderate confidence; likely close to true effect
LOW (⊕⊕○○)	1/16 (6%)	Limited confidence; may be substantially different
VERY LOW (⊕○○○)	0/16 (0%)	Very limited confidence in effect estimate

rather than suggesting a protective effect, consistent with both agents being non-teratogenic. Adjuvants (metformin, letrozole, clomiphene) lack teratogenicity (OR 1.04 [95% CI 0.90-1.20]; $I^2=35\%$), with varying certainty.^{39,110}

Absolute major malformation risk: 2-6%, aligned with natural conception adjusted for factors.²⁻⁴ The GRADE evidence is of high certainty overall,¹⁹ from RCTs/meta-analyses (Level I-II), with low bias, no inconsistency, and minimal imprecision. Clinically, this supports protocol flexibility. Knowledge gaps remain regarding rare events in subgroups.

INTERPRETATION IN THE CONTEXT OF EVIDENCE HIERARCHY

The evidence hierarchy helps resolve conflicts: Higher-quality RCTs and meta-analyses (Level I-II) consistently affirm safety, while concerns stem from observational and pharmacovigilance data (Level III-V), which are more vulnerable to bias and confounding.^{19,20} The randomization and prospective design of RCTs minimize selection and recall bias, supporting causal inference, while limitations in observational studies (e.g., indication bias) persist despite adjustments.²³

Dydrogesterone serves as an illustrative example: LOTUS RCTs and IPD demonstrate comparable safety to progesterone.¹⁰⁴ However, VigiBase signals⁵¹ reflect the Weber effect and reporting bias,^{17,18} and case-control studies⁵⁰ are limited by recall bias and confounding.¹⁴³ There is no biological plausibility for teratogenicity.¹⁰⁷ Sensitivity analyses (excluding biased studies) and subgroup analyses (e.g., exposure timing) confirm the dominance of RCTs. Clinically, RCT data should be prioritized for counseling. Knowledge gaps remain in applying the evidence hierarchy to real-time pharmacovigilance.

The timing of medication administration relative to embryonic development is crucial for interpreting safety data. Gonadotropins and GnRH analogues are typically administered pre-conception and cleared before organogenesis, while luteal phase support occurs during early embryonic development when organ formation begins. This temporal distinction may explain why concerns often focus on progestogens despite robust RCT evidence supporting their safety.

CLINICAL IMPLICATIONS FOR PATIENT COUNSELING AND TREATMENT SELECTION

Clinicians can confidently counsel on low teratogenic risk from ART medications, focusing on parental and procedural factors for the modest elevation (absolute 2-6%; adjusted OR 1.15-1.50).^{5,6} Treatment choices should be individualized: Oral dydrogesterone for compliance²⁸; antagonists for OHSS-prone⁸¹; and adjuvants in PCOS.¹¹⁵ The benefits outweigh the risks for most patients. The GRADE evidence is of high certainty,¹⁹ from guidelines and reviews. Clinically, patient preferences should be incorporated to enhance adherence. Knowledge gaps remain regarding personalized risk calculators.

COMPARISON WITH PREVIOUS SYSTEMATIC REVIEWS

This work aligns with and extends prior work. Elizur and Tulandi (2008)¹⁴⁴ provided foundational but outdated findings; Katalinic et al. (2022)¹⁰⁷ focused on dydrogesterone, matching our conclusions, though their scope was narrower. Our hierarchy emphasis differentiates this work, weighting RCTs over observational studies.^{21,22} The GRADE evidence is of high certainty,¹⁹ based on comparative synthesis. Clinically, these findings update counseling information with current data. Knowledge gaps remain in integrated reviews of multiple agents.

STRENGTHS OF THIS REVIEW

A comprehensive search from 1990-2025 across multiple databases and registries captured diverse evidence. The use of GRADE and risk tools ensured transparency^{56,57}; the evidence hierarchy resolved conflicts¹⁹ and addresses the challenge of attributing effects to specific agents in multi-drug protocols. Analysis of dydrogesterone illustrates this approach to evaluating evidence. The large sample size (~1.2 million pregnancies) provides sufficient power to assess rare outcomes. The GRADE evidence is of high certainty,¹⁹ with low risk of bias. Clinically, this work provides a robust framework for decision-making. Knowledge gaps remain because non-English research was excluded from the analysis.

LIMITATIONS

This systematic review has several limitations that warrant consideration. Heterogeneity in study designs, populations, and outcome definitions precluded meta-analyses for some

outcomes, necessitating reliance on narrative synthesis, which may limit precision in those areas.²⁷ Most of the trials included were powered for efficacy endpoints rather than rare safety outcomes such as congenital malformations, potentially missing ultra-rare events. Publication bias, although not strongly evident (funnel plot symmetry, Egger's $p=0.42$),⁵⁹ remains a concern due to the English-language restriction and potential underreporting of negative studies. The focus on short-term malformation outcomes also limits insights into long-term neurodevelopmental or metabolic effects, which require further investigation.⁶ Additionally, pooling data across studies to achieve precise estimates for rare events (e.g., pooled OR 1.01 [95% CI 0.92-1.11]; $I^2=20\%$ for overall ART medication safety) may obscure subtle differences between specific agents, protocols, or populations, despite low heterogeneity ($I^2=12-35\%$) supporting the validity of these pooled results.¹⁹ While this approach aligns with PRISMA and GRADE standards to maximize statistical power and provide high-certainty evidence ($\oplus\oplus\oplus\oplus$) for clinical guidance, it could mask study-specific variations that might be relevant in certain clinical contexts. To mitigate this, individual study characteristics are detailed in [Table 1](#), allowing readers to assess specific results alongside pooled estimates. A fundamental limitation in ART-safety research remains the difficulty of attributing malformations to specific agents within multi-drug protocols, partially addressed by emphasizing randomized controlled trials but not fully resolved. Future studies should prioritize prospective registries with standardized assessments and long-term follow-up to address these gaps.^{6,26}

RECOMMENDATIONS FOR FUTURE RESEARCH

Future studies should prioritize prospective registries with standardized assessments²⁶ and include long-term follow-up into childhood and adolescence for neurodevelopmental and metabolic outcomes.⁶ Research on gene-drug interaction and targeted dydrogesterone trials is needed to clarify existing signals.¹⁰⁷ International data should be collected, harmonized, and analyzed for new insights.²⁵ The GRADE evidence supporting these priorities is of high certainty,¹⁹ based on the gap analysis. Clinically, these recommendations will inform evidence-based updates. Knowledge gaps remain regarding modifiable procedural risks.

PUBLIC HEALTH IMPLICATIONS

With global ART growth (>13 million births),^{1,24} these findings encourage access to safe medications and promote individualizing care without undue anxiety. Emphasizing evidence quality in regulations and communication is essential to preventing restrictions on preliminary signals.⁶⁹ Continued support for registries and pharmacovigilance systems¹⁰⁷ along with investment in quality research will strengthen public and professional confidence.²⁵ The GRADE evidence is of high certainty,¹⁹ based on utilization data. Clinically, these insights promote equitable access to treatment. Gaps remain in our understanding of the impact of socioeconomic disparities in outcomes.

CONCLUSIONS AND CLINICAL IMPLICATIONS

This systematic review of 32 primary studies (~1.2 million pregnancies), drawing from a broader evidence base of 89 total cited works for contextual support, provides robust Level I-II evidence supporting the safety of standard ART pharmacological agents for fetal malformation risk, with no medication-specific teratogenic signals (overall pooled OR 1.01 [95% CI 0.92-1.11]; $I^2=20\%$ across classes).²⁻⁴

Gonadotropins, GnRH analogues, progesterone formulations (all routes), and adjuvants show absolute major malformation rates of 2-6%, comparable to adjusted natural conception.^{5,6} The GRADE evidence is of high certainty,¹⁹ based on RCTs and meta-analyses with low bias, no inconsistency, and minimal imprecision. Sensitivity analyses and subgroup evaluations (e.g., by PCOS diagnosis or age) confirm robustness.

Dydrogesterone serves as a clear example of this safety. RCTs (LOTUS I/II) and meta-analyses demonstrate statistically equivalent anomaly rates to progesterone (pooled OR 0.72 [95% CI 0.49-1.05]; $I^2=15\%$, $p>0.05$), confirming comparable safety profiles and equivalent efficacy.^{28,29,107} Lower-level signals (pharmacovigilance and case-control)^{50,51} are limited by biases (e.g., Weber effect and confounding),^{17,145} but outweighed by higher-quality evidence in the hierarchy.¹⁹ The GRADE evidence indicates with high certainty¹⁹ that prioritizing RCTs and IPD supports oral dydrogesterone use for compliance.⁹⁹ Knowledge gaps remain in registry data regarding rare anomaly subtypes.

Clinically, practitioners should counsel patients on low absolute risks, individualize protocols (e.g., antagonists for OHSS-prone patients⁸¹ or adjuvants in PCOS¹¹⁵). Practitioners should apply the evidence hierarchy when findings conflict, preserving access amid ART growth (>13 million births).¹ For the future, research should include long-term neurodevelopmental and metabolic follow-up, epigenetic studies, and harmonized registries.^{25,26} From a public health perspective, ongoing surveillance should be maintained without imposing undue restrictions based on preliminary signals,⁶⁹ promoting evidence-based confidence in ART safety.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

To prepare this manuscript, the author(s) used artificial intelligence software, including Grok, Claude AI, ChatGPT, and Open Science to organize tables and references. The authors reviewed and edited the content as needed after tool use and take full responsibility for the article content.

FUNDING STATEMENT

No specific funding was received for this study.

DISCLOSURE STATEMENT

Z.S. is a co-chairman of the online IVF-Worldwide Congress, which receives unrestricted educational grants from Merck, Organon, GE, Abbott, Vitrolife, and Besins. A.W. and Y.Y. have no competing interests.

ATTESTATION STATEMENT

This review does not involve human participants or patient data; therefore, ethics approval by an institutional review board was not required.

DATA SHARING STATEMENT

The datasets analyzed during the current study are available from the corresponding author upon reasonable request. Policy documents and reports cited in this analysis are publicly available from the sources referenced.

TRIAL REGISTRATION

PROSPERO CRD 420251118713

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Conceptualization: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Supporting). Data curation: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Supporting). Formal Analysis: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Supporting). Investigation: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Supporting). Methodology: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Supporting). Project administration: Zeev Shoham (Lead), Ariel Weissman (Supporting). Resources: Zeev Shoham (Lead), Ariel Weissman (Supporting), Yuval Yaron (Supporting). Software: Zeev

Shoham (Supporting), Ariel Weissman (Supporting). Supervision: Zeev Shoham (Lead), Ariel Weissman (Supporting). Validation: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Supporting). Visualization: Zeev Shoham (Lead), Ariel Weissman (Supporting). Writing – original draft: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Equal). Writing – review & editing: Zeev Shoham (Lead), Ariel Weissman (Equal), Yuval Yaron (Equal).

EQUATOR REPORTING GUIDELINES

The manuscript follows EQUATOR guidelines. See the Results section for the Preferred Reporting Items for Systematic Reviews with Selective Meta-Analyses (PRISMA) flow table.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contribution of Mr. Jaromir Tomasik (Statistical Consultant, Warsaw, Poland) for his expert support in reviewing and validating the statistical analyses of this work.

CAPSULE

This systematic review of 32 studies (~1.2 million pregnancies) demonstrates robust evidence that standard assisted reproductive technology medications carry no increased fetal malformation risk, with absolute rates of 2-6%, comparable to natural conception when adjusted for parental factors.

Submitted: October 23, 2025 CDT. Accepted: October 30, 2025 CDT. Published: January 29, 2026 CDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

REFERENCES

1. Adamson GD, Creighton P, de Mouzon J, Zegers-Hochschild F, Dyer S, Chambers GM. How many infants have been born with the help of assisted reproductive technology? *Fertil Steril*. 2025;124(1):40-50. doi:[10.1016/j.fertnstert.2025.02.009](https://doi.org/10.1016/j.fertnstert.2025.02.009)
2. Chen L, Yang T, Zheng Z, Yu H, Wang H, Qin J. Birth prevalence of congenital malformations in singleton pregnancies resulting from in vitro fertilization/intracytoplasmic sperm injection worldwide: a systematic review and meta-analysis. *Arch Gynecol Obstet*. 2018;297(5):1115-1130. doi:[10.1007/s00404-018-4712-x](https://doi.org/10.1007/s00404-018-4712-x)
3. Lu YH, Wang N, Jin F. Long-term follow-up of children conceived through assisted reproductive technology. *J Zhejiang Univ Sci B*. 2013;14(5):359-371. doi:[10.1631/jzus.B1200348](https://doi.org/10.1631/jzus.B1200348)
4. Zhao J, Yan Y, Huang X, Li Y. Do the children born after assisted reproductive technology have an increased risk of birth defects? A systematic review and meta-analysis. *J Matern Fetal Neonatal Med*. 2020;33(2):322-333. doi:[10.1080/14767058.2018.1488168](https://doi.org/10.1080/14767058.2018.1488168)
5. Pinborg A, Wennerholm UB, Romundstad LB, Loft A, Aittomaki K, Söderström-Anttila V, et al. Why do singletons conceived after assisted reproduction technology have adverse perinatal outcome? Systematic review and meta-analysis. *Hum Reprod Update*. 2013;19(2):87-104. doi:[10.1093/humupd/dms044](https://doi.org/10.1093/humupd/dms044)
6. Luke B, Brown MB, Wantman E, Forestieri NE, Browne ML, Fisher SC, et al. The risk of birth defects with conception by ART. *Hum Reprod*. 2020;35(1):116-123. doi:[10.1093/humrep/deaa272](https://doi.org/10.1093/humrep/deaa272)
7. Macklon NS, Stouffer RL, Giudice LC, Fauser BC. The science behind 25 years of ovarian stimulation for in vitro fertilization. *Endocr Rev*. 2006;27(2):170-207. doi:[10.1210/er.2005-0015](https://doi.org/10.1210/er.2005-0015)
8. van der Linden M, Buckingham K, Farquhar C, Kremer JA, Metwally M. Luteal phase support for assisted reproduction cycles. *Cochrane Database Syst Rev*. 2015;7:CD009154. doi:[10.1002/14651858.CD009154.pub3](https://doi.org/10.1002/14651858.CD009154.pub3)
9. Kiose KI, Storr A, Kolibianakis EM, Mol BW, Venetis CA. Biosimilars versus the originator of follitropin alfa for ovarian stimulation in ART: a systematic review and meta-analysis. *Hum Reprod*. 2025;40(2):343-359. doi:[10.1093/humrep/deae274](https://doi.org/10.1093/humrep/deae274)
10. Grynberg M, Cedrin-Durnerin I, Raguideau F, Herquelot E, Luciani L, Porte F, et al. Comparative effectiveness of gonadotropins used for ovarian stimulation during assisted reproductive technologies (ART) in France: a real-world observational study from the French nationwide claims database (SNDS). *Best Pract Res Clin Obstet Gynaecol*. 2023;88:102308. doi:[10.1016/j.bpobgyn.2022.102308](https://doi.org/10.1016/j.bpobgyn.2022.102308)
11. Glujovsky D, Pesce R, Miguens M, Sueldo C, Ciapponi A. Progestogens for prevention of luteinising hormone (LH) surge in women undergoing controlled ovarian hyperstimulation as part of an assisted reproductive technology (ART) cycle. *Cochrane Database Syst Rev*. 2023;11(11):CD013827. doi:[10.1002/14651858.CD013827.pub2](https://doi.org/10.1002/14651858.CD013827.pub2)
12. Ata B, Capuzzo M, Turkgeldi E, Yildiz S, La Marca A. Progestins for pituitary suppression during ovarian stimulation for ART: a comprehensive and systematic review including meta-analyses. *Hum Reprod Update*. 2021;27(1):48-66. doi:[10.1093/humupd/dmaa040](https://doi.org/10.1093/humupd/dmaa040)
13. Koren G, Gilboa D, Katz R. RETRACTED ARTICLE: Fetal safety of dydrogesterone exposure in the first trimester of pregnancy. *Clin Drug Investig*. 2019;39(11):1085-1090. doi:[10.1007/s40261-019-00862-w](https://doi.org/10.1007/s40261-019-00862-w)
14. Koren G, Gilboa D, Katz R. Expression of concern to: fetal safety of dydrogesterone exposure in the first trimester of pregnancy. *Clin Drug Investig*. 2020;40(2):199-200. doi:[10.1007/s40261-019-00884-4](https://doi.org/10.1007/s40261-019-00884-4)
15. Atarieh M, Javadian M, Basirat Z, Kashifard M, Yazdani S, Adib-Rad H, et al. RETRACTED ARTICLE: Comparison of the effect of dydrogesterone and natural micronized progesterone for luteal-phase support in assisted reproductive technology cycles: a single-blind randomized clinical trial study. *Health Sci Rep*. 2024;7(8):e2296. doi:[10.1002/hsr2.2296](https://doi.org/10.1002/hsr2.2296)
16. Retraction: Comparison of the effect of dydrogesterone and natural micronized progesterone for luteal-phase support in assisted reproductive technology cycles: a single-blind randomized clinical trial study. *Health Sci Rep*. 2025;8(8):e71139. doi:[10.1002/hsr2.71139](https://doi.org/10.1002/hsr2.71139)
17. Weber JCP. Epidemiology of adverse reactions to nonsteroidal anti-inflammatory drugs. In: Rainsford KD, Velo GD, eds. *Side-Effects of Anti-Inflammatory Drugs*. Advances in inflammation research. Raven Press; 1984:1-7.

18. Hartnell NR, Wilson JP. Replication of the Weber effect using postmarketing adverse event reports voluntarily submitted to the United States Food and Drug Administration. *Pharmacotherapy*. 2004;24(6):743-749. doi:[10.1592/phco.24.8.743.36068](https://doi.org/10.1592/phco.24.8.743.36068)
19. Guyatt G, Agoritsas T, Brignardello-Petersen R, Mustafa RA, Rylance J, Foroutan F, et al. Core GRADE 1: overview of the Core GRADE approach. *BMJ*. 2025;389:e081903. doi:[10.1136/bmj-2024-081903](https://doi.org/10.1136/bmj-2024-081903)
20. Brignardello-Petersen R, Carrasco-Labra A, Guyatt GH. How to interpret and use a clinical practice guideline or recommendation: Users' guides to the medical literature. *JAMA*. 2021;326(15):1516-1523. doi:[10.1001/jama.2021.15319](https://doi.org/10.1001/jama.2021.15319)
21. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred reporting items for a systematic review and meta-analysis of individual participant data: the PRISMA-IPD statement. *JAMA*. 2015;313(16):1657-1665. doi:[10.1001/jama.2015.3656](https://doi.org/10.1001/jama.2015.3656)
22. Veroniki AA, Seitidis G, Tsvigoulis G, Katsanos AH, Mavridis D. An introduction to individual participant data meta-analysis. *Neurology*. 2023;100(23):1102-1110. doi:[10.1212/WNL.0000000000207078](https://doi.org/10.1212/WNL.0000000000207078)
23. Loke YK, Price D, Derry S, Aronson JK. Case reports of suspected adverse drug reactions--systematic literature survey of follow-up. *BMJ*. 2006;332(7537):335-339. doi:[10.1136/bmj.38701.399942.63](https://doi.org/10.1136/bmj.38701.399942.63)
24. Faddy MJ, Gosden MD, Gosden RG. A demographic projection of the contribution of assisted reproductive technologies to world population growth. *Reprod Biomed Online*. 2018;36(4):455-458. doi:[10.1016/j.rbmo.2018.01.006](https://doi.org/10.1016/j.rbmo.2018.01.006)
25. De Geyter C, Wyns C, Calhaz-Jorge C, de Mouzon J, Ferraretti AP, Kupka M, et al. 20 years of the European in vitro fertilization-monitoring Consortium registry: what have we learned? A comparison with registries from two other regions. *Hum Reprod*. 2020;35(12):2832-2849. doi:[10.1093/humrep/deaa250](https://doi.org/10.1093/humrep/deaa250)
26. Berntsen S, Söderström-Anttila V, Wennerholm UB, Laivuori H, Loft A, Oldereid NB, et al. The health of children conceived by ART: "the chicken or the egg?" *Hum Reprod Update*. 2019;25(2):137-158. doi:[10.1093/humupd/dmz001](https://doi.org/10.1093/humupd/dmz001)
27. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:[10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)
28. Griesinger G, Blockeel C, Sukhikh GT, Patki A, Dhorepatil B, Yang DZ, et al. Oral dydrogesterone versus intravaginal micronized progesterone gel for luteal phase support in IVF: a randomized clinical trial. *Hum Reprod*. 2018;33(12):2212-2221. doi:[10.1093/humrep/dey306](https://doi.org/10.1093/humrep/dey306)
29. Tournaye H, Sukhikh GT, Kahler E, Griesinger G. A Phase III randomized controlled trial comparing the efficacy, safety and tolerability of oral dydrogesterone versus micronized vaginal progesterone for luteal support in in vitro fertilization. *Human Reproduction (Oxford, England)*. 2017;32(5):1019-1027. doi:[10.1093/humrep/dex023](https://doi.org/10.1093/humrep/dex023)
30. Devine K, Richter KS, Jahandideh S, Widra EA, McKeeby JL. Intramuscular progesterone optimizes live birth from programmed frozen embryo transfer: a randomized clinical trial. *Fertil Steril*. 2021;116(3):633-643. doi:[10.1016/j.fertnstert.2021.04.013](https://doi.org/10.1016/j.fertnstert.2021.04.013)
31. Yarali H, Mumusoglu S, Polat M, Erden M, Yarali Ozbek I, Esteves SC, et al. Comparison of the efficacy of subcutaneous versus vaginal progesterone using a rescue protocol in vitrified blastocyst transfer cycles. *Reprod Biomed Online*. 2023;47(3):103233. doi:[10.1016/j.rbmo.2023.05.005](https://doi.org/10.1016/j.rbmo.2023.05.005)
32. Pabuccu E, Kovanci E, Israfilova G, Tulek F, Demirel C, Pabuccu R, et al. Oral, vaginal or intramuscular progesterone in programmed frozen embryo transfer cycles: a pilot randomized controlled trial. *Reprod Biomed Online*. 2022;45(6):1145-1151. doi:[10.1016/j.rbmo.2022.06.027](https://doi.org/10.1016/j.rbmo.2022.06.027)
33. Witz CA, Daftary GS, Doody KJ, Park JK, Seifu Y, Yankov VI, et al. Randomized, assessor-blinded trial comparing highly purified human menotropin and recombinant follicle-stimulating hormone in high responders undergoing intracytoplasmic sperm injection. *Fertil Steril*. 2020;114(2):321-330. doi:[10.1016/j.fertnstert.2020.03.029](https://doi.org/10.1016/j.fertnstert.2020.03.029)
34. Chua SJ, Mol BW, Longobardi S, Orvieto R, Venetis CA, Lispi M, et al. Biosimilar recombinant follitropin alfa preparations versus the reference product (Gonal-F®) in couples undergoing assisted reproductive technology treatment: a systematic review and meta-analysis. *Reprod Biol Endocrinol*. 2021;19(1):51. doi:[10.1186/s12958-021-00727-y](https://doi.org/10.1186/s12958-021-00727-y)
35. Alviggi C, Vigilante L, Cariati F, Conforti A, Humaidan P. The role of recombinant LH in ovarian stimulation: what's new? *Reprod Biol Endocrinol*. 2025;23(Suppl 1):38. doi:[10.1186/s12958-025-01361-8](https://doi.org/10.1186/s12958-025-01361-8)

36. Conforti A, Esteves SC, Humaidan P, Longobardi S, D'Hooghe T, Orvieto R, et al. Recombinant human luteinizing hormone co-treatment in ovarian stimulation for assisted reproductive technology in women of advanced reproductive age: a systematic review and meta-analysis of randomized controlled trials. *Reprod Biol Endocrinol*. 2021;19(1):91. doi:[10.1186/s12958-021-00759-4](https://doi.org/10.1186/s12958-021-00759-4)
37. Katalinic A, Noftz MR, Garcia-Velasco JA, Shulman LP, van den Anker JN, Strauss JF 3rd. No additional risk of congenital anomalies after first-trimester dydrogesterone use: a systematic review and meta-analysis. *Hum Reprod Open*. 2024;2024(1):hoae004. doi:[10.1093/hropen/hoae004](https://doi.org/10.1093/hropen/hoae004)
38. Barbosa MWP, Silva LR, Navarro PA, Ferriani RA, Nastri CO, Martins WP. Dydrogesterone vs progesterone for luteal-phase support: systematic review and meta-analysis of randomized controlled trials. *Ultrasound Obstet Gynecol*. 2016;48(2):161-170. doi:[10.1002/uog.15814](https://doi.org/10.1002/uog.15814)
39. Pundir J, Achilli C, Bhide P, Sabatini L, Legro RS, Rombauts L, et al. Risk of foetal harm with letrozole use in fertility treatment: a systematic review and meta-analysis. *Hum Reprod Update*. 2021;27(3):474-485. doi:[10.1093/humupd/dmaa055](https://doi.org/10.1093/humupd/dmaa055)
40. Davies MJ, Moore VM, Willson KJ, Van Essen P, Priest K, Scott H, et al. Reproductive technologies and the risk of birth defects. *N Engl J Med*. 2012;366(19):1803-1813. doi:[10.1056/NEJMoa1008095](https://doi.org/10.1056/NEJMoa1008095)
41. Zhang N, Tian T, Li J, Zhu X, Jiesisibieke D, Fang S, et al. A comparison of pregnancy outcomes and congenital malformations in offspring between patients undergoing intracytoplasmic sperm injection and conventional in vitro fertilization: a retrospective cohort study. *Fertil Steril*. 2024;121(6):982-990. doi:[10.1016/j.fertnstert.2024.01.025](https://doi.org/10.1016/j.fertnstert.2024.01.025)
42. Qin JB, Sheng XQ, Wang H, Chen GC, Yang J, Yu H, et al. Worldwide prevalence of adverse pregnancy outcomes associated with in vitro fertilization/ intracytoplasmic sperm injection among multiple births: a systematic review and meta-analysis based on cohort studies. *Arch Gynecol Obstet*. 2017;295(3):577-597. doi:[10.1007/s00404-017-4291-2](https://doi.org/10.1007/s00404-017-4291-2)
43. Huang J, Xie Q, Lin J, Lu X, Wang N, Gao H, et al. Neonatal outcomes and congenital malformations in children born after dydrogesterone application in progestin-primed ovarian stimulation protocol for IVF: a retrospective cohort study. *Drug Des Devel Ther*. 2019;13:2553-2563. doi:[10.2147/DDDT.S210228](https://doi.org/10.2147/DDDT.S210228)
44. Yetkinel S, Aytaç PÇ, Durdağ GD, Yağınç DA, Kılıçdağ EB, Şimşek E. Comparison of highly purified human menopausal gonadotropin and recombinant follicle stimulating hormone use in patients undergoing in vitro fertilization with progestin-primed ovarian stimulation protocol: a single center retrospective analysis. *Arch Gynecol Obstet*. 2024;310(5):2657-2662. doi:[10.1007/s00404-024-07756-z](https://doi.org/10.1007/s00404-024-07756-z)
45. Ni T, Zhou W, Liu Y, Xu C, Xu B, Li Y, et al. Excessive exogenous gonadotropins and genetic and pregnancy outcomes after euploidy embryo transfer: a secondary analysis of a randomized clinical trial. *JAMA Netw Open*. 2024;7(4):e244438. doi:[10.1001/jamanetworkopen.2024.4438](https://doi.org/10.1001/jamanetworkopen.2024.4438)
46. Farhi J, Ben-Haroush A, Andrawus N, Pinkas H, Sapir O, Fisch B, et al. High serum oestradiol concentrations in IVF cycles increase the risk of pregnancy complications related to abnormal placentation. *Reprod Biomed Online*. 2010;21(3):331-337. doi:[10.1016/j.rbmo.2010.04.022](https://doi.org/10.1016/j.rbmo.2010.04.022)
47. Sagot P, Bechoua S, Ferdynus C, Facy A, Flamm X, Gouyon JB, et al. Similarly increased congenital anomaly rates after intrauterine insemination and IVF technologies: a retrospective cohort study. *Hum Reprod*. 2012;27(3):902-909. doi:[10.1093/humrep/der443](https://doi.org/10.1093/humrep/der443)
48. Kallen B, Finnström O, Nygren KG, Otterblad Olausson P. In vitro fertilization (IVF) in Sweden: risk for congenital malformations after different IVF methods. *Birth Defects Res A Clin Mol Teratol*. 2005;73(3):162-169. doi:[10.1002/bdra.20107](https://doi.org/10.1002/bdra.20107)
49. Bonduelle M, Wennerholm UB, Loft A, Tarlatzis BC, Peters C, Henriët S, et al. A multi-centre cohort study of the physical health of 5-year-old children conceived after intracytoplasmic sperm injection, in vitro fertilization and natural conception. *Hum Reprod*. 2005;20(2):413-419. doi:[10.1093/humrep/deh592](https://doi.org/10.1093/humrep/deh592)
50. Zaqout M, Aslem E, Abuqamar M, Abughazza O, Panzer J, De Wolf D, et al. The impact of oral intake of dydrogesterone on fetal heart development during early pregnancy. *Pediatric Cardiology*. 2015;36(7):1483-1488. doi:[10.1007/s00246-015-1190-9](https://doi.org/10.1007/s00246-015-1190-9)
51. Henry A, Santulli P, Bourdon M, Maignien C, Chapron C, Treluyer JM, et al. Birth defects reporting and the use of dydrogesterone: a disproportionality analysis from the World Health Organization pharmacovigilance database (VigiBase). *Human Reproduction Open*. 2025;2025(1):hoae072. doi:[10.1093/hropen/hoae072](https://doi.org/10.1093/hropen/hoae072)

52. Rimm AA, Katayama AC, Diaz M, Katayama KP. A meta-analysis of controlled studies comparing major malformation rates in IVF and ICSI infants with naturally conceived children. *J Assist Reprod Genet.* 2004;21(12):437-443. doi:[10.1007/s10815-004-8760-8](https://doi.org/10.1007/s10815-004-8760-8)
53. Hansen M, Kurinczuk JJ, Bower C, Webb S. The risk of major birth defects after intracytoplasmic sperm injection and in vitro fertilization. *N Engl J Med.* 2002;346(10):725-730. doi:[10.1056/NEJMoa010035](https://doi.org/10.1056/NEJMoa010035)
54. Liang Y, Chen L, Yu H, Wang H, Li Q, Yu R, et al. Which type of congenital malformations is significantly increased in singleton pregnancies following in vitro fertilization/intracytoplasmic sperm injection: a systematic review and meta-analysis. *Oncotarget.* 2018;9(3):4267-4278. doi:[10.18632/oncotarget.23689](https://doi.org/10.18632/oncotarget.23689)
55. Dolk H, Loane M, Garne E. The prevalence of congenital anomalies in Europe. *Adv Exp Med Biol.* 2010;686:349-364. doi:[10.1007/978-90-481-9485-8_20](https://doi.org/10.1007/978-90-481-9485-8_20)
56. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366:l4898. doi:[10.1136/bmj.l4898](https://doi.org/10.1136/bmj.l4898)
57. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute. 2000. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
58. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177-188. doi:[10.1016/0197-2456\(86\)90046-2](https://doi.org/10.1016/0197-2456(86)90046-2)
59. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7109):629-634. doi:[10.1136/bmj.315.7109.629](https://doi.org/10.1136/bmj.315.7109.629)
60. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ.* 2015;350:g7647. doi:[10.1136/bmj.g7647](https://doi.org/10.1136/bmj.g7647)
61. Andrade C. A primer on individual participant data meta-analysis and its strengths and limitations. *J Clin Psychiatry.* 2025;86(3):25f16001. doi:[10.4088/ICP.25f16001](https://doi.org/10.4088/ICP.25f16001)
62. Tudur Smith C, Marcucci M, Nolan SJ, Iorio A, Sudell M, Riley R, et al. Individual participant data meta-analyses compared with meta-analyses based on aggregate data. *Cochrane Database Syst Rev.* 2016;9:MR000007. doi:[10.1002/14651858.MR000007.pub3](https://doi.org/10.1002/14651858.MR000007.pub3)
63. Younes N, Claude LA, Paoletti X. Reading, conducting, and developing systematic review and individual patient data meta-analyses in psychiatry for treatment issues. *Front Psychiatry.* 2021;12:644980. doi:[10.3389/fpsy.2021.644980](https://doi.org/10.3389/fpsy.2021.644980)
64. Jacobs AK, Kushner FG, Ettinger SM, Guyton RA, Anderson JL, Ohman EM, et al. ACCF/AHA clinical practice guideline methodology summit report: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* 2013;61(2):213-265. doi:[10.1016/j.jacc.2012.09.025](https://doi.org/10.1016/j.jacc.2012.09.025)
65. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *J Am Coll Cardiol.* 2014;63(25 Pt B):2985-3023. doi:[10.1016/j.jacc.2013.11.004](https://doi.org/10.1016/j.jacc.2013.11.004)
66. Capolupo A, Petrocchi S, Melchiorre M, Jonas K, D'Hooghe T, Hanyaloglu A, et al. Analytical investigation of the profile of human chorionic gonadotropin in highly purified human menopausal gonadotrophin preparations. *Int J Mol Sci.* 2024;25(17):9405. doi:[10.3390/ijms25179405](https://doi.org/10.3390/ijms25179405)
67. Casarini L, Riccetti L, Paradiso E, Limoncella S, Lazzaretti C, Sperduti S, et al. Two human menopausal gonadotrophin (hMG) preparations display different early signaling in vitro. *Mol Hum Reprod.* 2020;26(12):894-905. doi:[10.1093/molehr/gaaa070](https://doi.org/10.1093/molehr/gaaa070)
68. FDA. Menopur. FDA Drug Label. Food and Drug Administration.
69. Buhler KF, Fischer R, Verpillat P, Allignol A, Guedes S, Boutmy E, et al. Comparative effectiveness of recombinant human follicle-stimulating hormone alfa (R-hFSH-alfa) versus highly purified urinary human menopausal gonadotropin (hMG HP) in assisted reproductive technology (ART) treatments: a non-interventional study in Germany. *Reprod Biol Endocrinol.* 2021;19(1):90. doi:[10.1186/s12958-021-00768-3](https://doi.org/10.1186/s12958-021-00768-3)
70. Lispi M, Humaidan P, Bousfield GR, D'Hooghe T, Ulloa-Aguirre A. Follicle-stimulating hormone biological products: does potency predict clinical efficacy? *Int J Mol Sci.* 2023;24(10):9020. doi:[10.3390/ijms24109020](https://doi.org/10.3390/ijms24109020)

71. Manzi L, Sepe N, Migliaccio W, Lanzoni L, D'Angelo F, Colarusso L, et al. Comparative assessment of the structural features of originator recombinant human follitropin alfa versus recombinant human follitropin alfa biosimilar preparations approved in non-European regions. *Int J Mol Sci*. 2022;23(12):6762. doi:[10.3390/ijms23126762](https://doi.org/10.3390/ijms23126762)
72. Moore TJ, Mouslim MC, Blunt JL, Alexander GC, Shermock KM. Assessment of availability, clinical testing, and US Food and Drug Administration review of biosimilar biologic products. *JAMA Intern Med*. 2021;181(1):52-60. doi:[10.1001/jamainternmed.2020.3997](https://doi.org/10.1001/jamainternmed.2020.3997)
73. Yu Y, Zhang Q, Sun K, Xiu Y, Wang X, Yan L, et al. The therapeutic effects of rFSH versus uFSH/uHMG on ovarian stimulation in women undergoing assisted reproductive technology: a meta-analysis of randomized controlled trials. *Arch Gynecol Obstet*. 2024;309(6):2529-2555. doi:[10.1007/s00404-023-07095-5](https://doi.org/10.1007/s00404-023-07095-5)
74. Twells LK. Evidence-based decision-making 1: critical appraisal. *Methods Mol Biol*. 2021;2249:389-404. doi:[10.1007/978-1-0716-1138-8_21](https://doi.org/10.1007/978-1-0716-1138-8_21)
75. Hohmann E, Feldman M, Hunt TJ, Cote MP, Brand JC. Research pearls: how do we establish the level of evidence? *Arthroscopy*. 2018;34(12):3271-3277. doi:[10.1016/j.arthro.2018.10.002](https://doi.org/10.1016/j.arthro.2018.10.002)
76. Bielfeld AP, Schwarze JE, Verpillat P, Lispi A, Fischer R, Hayward B, et al. Effectiveness of recombinant human follicle-stimulating hormone (R-hFSH): recombinant human luteinizing hormone versus R-hFSH alone in assisted reproductive technology treatment cycles among women aged 35-40 years: a German database study. *Best Pract Res Clin Obstet Gynaecol*. 2023;89:102350. doi:[10.1016/j.bpobgyn.2023.102350](https://doi.org/10.1016/j.bpobgyn.2023.102350)
77. Chen MJ, Yi YC, Guu HF, Chen YF, Kung HF, Chang JC, et al. A retrospective, matched case-control study of recombinant LH versus hMG supplementation on FSH during controlled ovarian hyperstimulation in the GnRH-antagonist protocol. *Front Endocrinol (Lausanne)*. 2022;13:931756. doi:[10.3389/fendo.2022.931756](https://doi.org/10.3389/fendo.2022.931756)
78. Kirshenbaum M, Gil O, Haas J, Nahum R, Zilberberg E, Lebovitz O, et al. Recombinant follicular stimulating hormone plus recombinant luteinizing hormone versus human menopausal gonadotropins---does the source of LH bioactivity affect ovarian stimulation outcome? *Reprod Biol Endocrinol*. 2021;19(1):182. doi:[10.1186/s12958-021-00853-7](https://doi.org/10.1186/s12958-021-00853-7)
79. Mao R, Hou X, Feng X, Wang R, Fei X, Zhao J, et al. Recombinant human luteinizing hormone increases endometrial thickness in women undergoing assisted fertility treatments: a systematic review and meta-analysis. *Front Pharmacol*. 2024;15:1434625. doi:[10.3389/fphar.2024.1434625](https://doi.org/10.3389/fphar.2024.1434625)
80. Carson SA, Kallen AN. Diagnosis and management of infertility: a review. *JAMA*. 2021;326(1):65-76. doi:[10.1001/jama.2021.4788](https://doi.org/10.1001/jama.2021.4788)
81. Practice Committee of the American Society for Reproductive Medicine. Prevention of moderate and severe ovarian hyperstimulation syndrome: a guideline. *Fertil Steril*. 2024;121(2):230-245. doi:[10.1016/j.fertnstert.2023.11.013](https://doi.org/10.1016/j.fertnstert.2023.11.013)
82. Chung EH, Harris BS, Muasher SJ, Kuller JA. The risk of congenital anomalies by fertility treatment modality. *Obstet Gynecol Surv*. 2021;76(1):37-47. doi:[10.1097/OGX.0000000000000855](https://doi.org/10.1097/OGX.0000000000000855)
83. Mannaerts B, Yding Andersen C, Howles CM. Does hCG and LH supplementation during ovarian stimulation improve clinical outcome? An evaluation of 30 years of clinical research. *Reprod Biomed Online*. 2024;50(6):104782. doi:[10.1016/j.rbmo.2024.104782](https://doi.org/10.1016/j.rbmo.2024.104782)
84. Santoro N, Polotsky AJ. Infertility evaluation and treatment. *N Engl J Med*. 2025;392(11):1111-1119. doi:[10.1056/NEJMcp2311150](https://doi.org/10.1056/NEJMcp2311150)
85. Smits J, Platteau P. Influence of human chorionic gonadotrophin during ovarian stimulation: an overview. *Reprod Biol Endocrinol*. 2020;18(1):80. doi:[10.1186/s12958-020-00639-3](https://doi.org/10.1186/s12958-020-00639-3)
86. Zhu J, Xing W, Li T, Lin H, Ou J. GnRH antagonist protocol versus GnRH agonist long protocol: a retrospective cohort study on clinical outcomes and maternal-neonatal safety. *Front Endocrinol*. 2022;13:875779. doi:[10.3389/fendo.2022.875779](https://doi.org/10.3389/fendo.2022.875779)
87. FDA. Cetrotide. FDA Drug Label.
88. FDA. Fyremadel. FDA Drug Label.
89. Xiong M, Lu J, Dong N, Wu R, Zhang D, Li B, et al. Exposure to gonadotropin-releasing hormone agonist in early pregnancy leads to adverse pregnancy outcomes: a retrospective analysis. *Arch Gynecol Obstet*. 2025;311(3):801-809. doi:[10.1007/s00404-024-07914-3](https://doi.org/10.1007/s00404-024-07914-3)
90. Wu H, Xu X, Ma C, Zhou Y, Pei S, Geng H, et al. No significant long-term complications from inadvertent exposure to gonadotropin-releasing hormone agonist during early pregnancy in mothers and offspring: a retrospective analysis. *Reprod Biol Endocrinol*. 2021;19(1):46. doi:[10.1186/s12958-021-00732-1](https://doi.org/10.1186/s12958-021-00732-1)

91. Nguyen CN, Nguyen TTN, Le TT, Vu PB, Le LTP, Pham NB, et al. Effectiveness, safety, and practicality of intramuscular progesterone in restoring pregnancy outcomes during frozen embryo transfers. *Front Endocrinol.* 2025;16:1592956. doi:[10.3389/fendo.2025.1592956](https://doi.org/10.3389/fendo.2025.1592956)
92. White J, Hickey J, Dufton M, Sandila N, Ripley M. Retrospective review of reproductive outcomes comparing vaginal progesterone to intramuscular progesterone as luteal support in frozen embryo transfer cycles. *J Obstet Gynaecol Can.* 2022;44(7):791-797. doi:[10.1016/j.jogc.2022.03.009](https://doi.org/10.1016/j.jogc.2022.03.009)
93. Demirel C, Özcan P, Tülek F, Timur HT, Pasin Ö. Initiating luteal phase support with SC progesterone based on low serum progesterone on the transfer day in true natural cycle frozen embryo transfers. *Front Endocrinol.* 2023;14:1278042. doi:[10.3389/fendo.2023.1278042](https://doi.org/10.3389/fendo.2023.1278042)
94. ESHRE Guideline Group on Ovarian Stimulation. ESHRE guideline: ovarian stimulation for IVF/ICSI. *Hum Reprod Open.* 2020;2020(2):hoaa009. doi:[10.1093/hropen/hoaa009](https://doi.org/10.1093/hropen/hoaa009)
95. Almohammadi A, Raveendran A, Black M, Maheshwari A. The optimal route of progesterone administration for luteal phase support in a frozen embryo transfer: a systematic review. *Arch Gynecol Obstet.* 2023;308(2):341-350. doi:[10.1007/s00404-022-06674-2](https://doi.org/10.1007/s00404-022-06674-2)
96. Elenis E, Joelsson LS, Stavreus-Evers A, Wänggren K. Perinatal outcomes of progesterone in natural frozen-thawed embryo transfer pregnancies: insights from 2 randomized controlled trials. *Fertil Steril.* Published online 2024. doi:[10.1016/j.fertnstert.2024.12.022](https://doi.org/10.1016/j.fertnstert.2024.12.022)
97. Jiang Y, Wang L, Shen H, Wang B, Wu J, Hu K, et al. The effect of progesterone supplementation for luteal phase support in natural cycle frozen embryo transfer: a systematic review and meta-analysis based on randomized controlled trials. *Fertil Steril.* 2023;119(4):597-605. doi:[10.1016/j.fertnstert.2022.12.035](https://doi.org/10.1016/j.fertnstert.2022.12.035)
98. Stavridis K, Balafoutas D, Kalampokas T, Benetou V, Samoli E, Vlahos N, et al. Oral dydrogesterone versus vaginal progesterone for luteal phase support in frozen-thawed embryo transfer cycles: a systematic review and meta-analysis of randomized controlled trials. *J Clin Med.* 2025;14(9):3238. doi:[10.3390/jcm14093238](https://doi.org/10.3390/jcm14093238)
99. Patki A. Role of Dydrogesterone for Luteal Phase Support in Assisted Reproduction. *Reproductive Sciences (Thousand Oaks, Calif).* 2024;31(1):17-29. doi:[10.1007/s43032-023-01302-z](https://doi.org/10.1007/s43032-023-01302-z)
100. Griesinger G, Tournaye H, Macklon N, Petraglia F, Arck P, Blockeel C, et al. Dydrogesterone: Pharmacological Profile and Mechanism of Action as Luteal Phase Support in Assisted Reproduction. *Reproductive Biomedicine Online.* 2019;38(2):249-259. doi:[10.1016/j.rbmo.2018.11.017](https://doi.org/10.1016/j.rbmo.2018.11.017)
101. Stute P. Dydrogesterone Indications Beyond Menopausal Hormone Therapy: An Evidence Review and Woman's Journey. *Gynecological Endocrinology: The Official Journal of the International Society of Gynecological Endocrinology.* 2021;37(8):683-688. doi:[10.1080/09513590.2021.1908252](https://doi.org/10.1080/09513590.2021.1908252)
102. Ott J, Egarter C, Aguilera A. Dydrogesterone after 60 years: A glance at the safety profile. *Gynecol Endocrinol.* 2022;38(4):279-287. doi:[10.1080/09513590.2021.2016692](https://doi.org/10.1080/09513590.2021.2016692)
103. Jeschke L, Santamaria CG, Meyer N, Zenclussen AC, Bartley J, Schumacher A, et al. Early-pregnancy dydrogesterone supplementation mimicking luteal-phase support in ART patients did not provoke major reproductive disorders in pregnant mice and their progeny. *Int J Mol Sci.* 2021;22(10):5403. doi:[10.3390/ijms22105403](https://doi.org/10.3390/ijms22105403)
104. Griesinger G, Blockeel C, Kahler E, Pexman-Fieth C, Olofsson JI, Driessen S, et al. Dydrogesterone as an oral alternative to vaginal progesterone for IVF luteal phase support: A systematic review and individual participant data meta-analysis. *PLoS One.* 2020;15(11):e0241044. doi:[10.1371/journal.pone.0241044](https://doi.org/10.1371/journal.pone.0241044)
105. Nadarajah R, Rajesh H, Wong KY, Faisal F, Yu SL. Live Birth Rates and Safety Profile Using Dydrogesterone for Luteal Phase Support in Assisted Reproductive Techniques. *Singapore Medical Journal.* 2017;58(6):294-297. doi:[10.11622/smedj.2016080](https://doi.org/10.11622/smedj.2016080)
106. van Gelder MM, de Jong-van den Berg LT, Roeleveld N. Drugs associated with teratogenic mechanisms. Part II: a literature review of the evidence on human risks. *Hum Reprod.* 2014;29(1):168-183. doi:[10.1093/humrep/det370](https://doi.org/10.1093/humrep/det370)
107. Katalinic A, Shulman LP, Strauss JF, Garcia-Velasco JA, van den Anker JN. A critical appraisal of safety data on dydrogesterone for the support of early pregnancy: a scoping review and meta-analysis. *Reprod Biomed Online.* 2022;45(2):365-375. doi:[10.1016/j.rbmo.2022.03.032](https://doi.org/10.1016/j.rbmo.2022.03.032)
108. Devall AJ, Papadopoulou A, Podesek M, Haas DM, Price MJ, Coomarasamy A, et al. Progesterogens for preventing miscarriage: a network meta-analysis. *Cochrane Database Syst Rev.* 2021;4:CD013792. doi:[10.1002/14651858.CD013792.pub2](https://doi.org/10.1002/14651858.CD013792.pub2)

109. Zhao H, He W, Xia C, Yang Z. Efficacy and Safety of Different Progestogens in Women With First Threatened Miscarriage: A Network Meta-Analysis. *Int J Gynaecol Obstet.* 2025;168(3):944-957. doi:[10.1002/ijgo.15987](https://doi.org/10.1002/ijgo.15987)
110. Chiu YH, Huybrechts KF, Patorno E, Yland JJ, Cesta CE, Bateman BT, et al. Metformin use in the first trimester of pregnancy and risk for nonlive birth and congenital malformations: emulating a target trial using real-world data. *Ann Intern Med.* 2024;177(7):862-870. doi:[10.7326/M23-2038](https://doi.org/10.7326/M23-2038)
111. Malek R, Davis SN. Is metformin safe in pregnancy: a focus on offspring outcomes. *Expert Opin Drug Saf.* 2025;24(1):5-8. doi:[10.1080/14740338.2024.2424410](https://doi.org/10.1080/14740338.2024.2424410)
112. Paschou SA, Shalit A, Gerontiti E, Athanasiadou KI, Kalampokas T, Psaltopoulou T, et al. Efficacy and safety of metformin during pregnancy: an update. *Endocrine.* 2024;83(2):259-269. doi:[10.1007/s12020-023-03550-0](https://doi.org/10.1007/s12020-023-03550-0)
113. Toft JH, Økland I. Metformin use in pregnancy: what about long-term effects in offspring? *Acta Obstet Gynecol Scand.* 2024;103(7):1238-1241. doi:[10.1111/aogs.14878](https://doi.org/10.1111/aogs.14878)
114. Tosti G, Barberio A, Tartaglione L, Rizzi A, Di Leo M, Viti L, et al. Lights and shadows on the use of metformin in pregnancy: from the preconception phase to breastfeeding and beyond. *Front Endocrinol.* 2023;14:1176623. doi:[10.3389/fendo.2023.1176623](https://doi.org/10.3389/fendo.2023.1176623)
115. Teede HJ, Tay CT, Laven JJE, Dokras A, Moran LJ, Piltonen TT, et al. Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *J Clin Endocrinol Metab.* 2023;108(10):2447-2469. doi:[10.1210/clinem/dgad463](https://doi.org/10.1210/clinem/dgad463)
116. Biljan MM, Hemmings R, Brassard N. The outcome of 150 babies following the treatment with letrozole or letrozole and gonadotropins. *Fertil Steril.* 2005;84(Suppl 1.1):O-251. doi:[10.1016/j.fertnstert.2005.07.230](https://doi.org/10.1016/j.fertnstert.2005.07.230)
117. Tulandi T, Martin J, Al-Fadhli R, Kabli N, Forman R, Hitkari J, et al. Congenital malformations among 911 newborns conceived after infertility treatment with letrozole or clomiphene citrate. *Fertil Steril.* 2006;85(6):1761-1765. doi:[10.1016/j.fertnstert.2006.03.014](https://doi.org/10.1016/j.fertnstert.2006.03.014)
118. Etrusco A, D'Amato A, Agrifoglio V, Laganà AS, Andrisani A, Cantatore C, et al. Does the use of letrozole in infertility treatment increase the risk of foetal malformations? A systematic review and meta-analysis of randomized controlled trials. *Eur J Obstet Gynecol Reprod Biol.* 2025;312:114531. doi:[10.1016/j.ejogrb.2025.114531](https://doi.org/10.1016/j.ejogrb.2025.114531)
119. Nehard R, Vauzelle C, Beghin D, Latour M, Elefant E, Coulm B, et al. Post-conceptual exposure to clomiphene citrate and congenital malformations: a cohort study. *Drug Saf.* 2024;47(9):883-894. doi:[10.1007/s40264-024-01438-0](https://doi.org/10.1007/s40264-024-01438-0)
120. Chin PY, Chan HY, Kieffer TEC, Prins JR, Russell DL, Davies MJ, et al. Clomiphene citrate administered in periconception phase causes fetal loss and developmental impairment in mice. *Endocrinology.* 2024;165(7):bqae047. doi:[10.1210/endo/bqae047](https://doi.org/10.1210/endo/bqae047)
121. Sood A, Mohiyiddeen G, Ahmad G, Fitzgerald C, Watson A, Mohiyiddeen L, et al. Growth hormone for in vitro fertilisation (IVF). *Cochrane Database Syst Rev.* 2021;11:CD000099. doi:[10.1002/14651858.CD000099.pub4](https://doi.org/10.1002/14651858.CD000099.pub4)
122. Shang Y, Wu M, He R, Ye Y, Sun X. Administration of growth hormone improves endometrial function in women undergoing in vitro fertilization: a systematic review and meta-analysis. *Hum Reprod Update.* 2022;28(6):838-857. doi:[10.1093/humupd/dmac028](https://doi.org/10.1093/humupd/dmac028)
123. Marinaro J, Goldstein M. Microsurgical management of male infertility: compelling evidence that collaboration with qualified male reproductive urologists enhances assisted reproductive technology (ART) outcomes. *J Clin Med.* 2022;11(15):4593. doi:[10.3390/jcm11154593](https://doi.org/10.3390/jcm11154593)
124. Jwa SC, Jwa J, Kuwahara A, Irahara M, Ishihara O, Saito H. Male subfertility and the risk of major birth defects in children born after in vitro fertilization and intracytoplasmic sperm injection: a retrospective cohort study. *BMC Pregnancy Childbirth.* 2019;19(1):192. doi:[10.1186/s12884-019-2322-7](https://doi.org/10.1186/s12884-019-2322-7)
125. Ahmadi H, Aghebati-Maleki L, Rashidani S, Csabai T, Nnaemeka OB, Szekeres-Bartho J. Long-term effects of ART on the health of the offspring. *Int J Mol Sci.* 2023;24(17):13564. doi:[10.3390/ijms241713564](https://doi.org/10.3390/ijms241713564)
126. Foresta C, Garolla A, Bartoloni L, Bettella A, Ferlin A. Genetic abnormalities among severely oligospermic men who are candidates for intracytoplasmic sperm injection. *J Clin Endocrinol Metab.* 2005;90:152-156. doi:[10.1210/jc.2004-1469](https://doi.org/10.1210/jc.2004-1469)
127. Banica AM, Popescu SD, Vlădăreanu S. Eye anomalies in children born through ART. *Rom J Ophthalmol.* 2021;65(4):310-314. doi:[10.22336/rjo.2021.65](https://doi.org/10.22336/rjo.2021.65)
128. Hattori H, Hiura H, Kitamura A, Miyauchi N, Kobayashi N, Takahashi S, et al. Association of four imprinting disorders and ART. *Clin Epigenet.* 2019;11:21. doi:[10.1186/s13148-019-0623-3](https://doi.org/10.1186/s13148-019-0623-3)

129. Choux C, Petazzi P, Sánchez-Delgado M, Hernandez Mora JR, Monteagudo A, Sagot P, et al. The hypomethylation of imprinted genes in IVF/ICSI placenta samples is associated with concomitant changes in histone modifications. *Epigenetics*. 2020;15(6-7):635-651. doi:[10.1080/15592294.2020.1783168](https://doi.org/10.1080/15592294.2020.1783168)
130. van Montfoort APA, Hanssen LLP, de Sutter P, Viville S, Geraedts JPM, de Boer P. Assisted reproduction treatment and epigenetic inheritance. *Hum Reprod Update*. 2012;18(2):171-197. doi:[10.1093/humupd/dmr047](https://doi.org/10.1093/humupd/dmr047)
131. Vermeiden JP, Bernardus RE. Are imprinting disorders more prevalent after human in vitro fertilization or intracytoplasmic sperm injection? *Fertil Steril*. 2013;99(3):642-651. doi:[10.1016/j.fertnstert.2013.01.125](https://doi.org/10.1016/j.fertnstert.2013.01.125)
132. DeBaun MR, Niemitz EL, Feinberg AP. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. *Am J Hum Genet*. 2003;72(1):156-160. doi:[10.1086/346031](https://doi.org/10.1086/346031)
133. Esteves SC, Roque M, Bedoschi G, Haahr T, Humaidan P. Intracytoplasmic sperm injection for male infertility and consequences for offspring. *Nat Rev Urol*. 2018;15(9):535-562. doi:[10.1038/s41585-018-0051-8](https://doi.org/10.1038/s41585-018-0051-8)
134. Mandourah S, Badeghiesh A, Baghlaf H, Dahan MH. How does the need for IVF affect pregnancy complications among multiple gestations? The study of a large American population database including almost 100,000 multiple gestations. *Eur J Obstet Gynecol Reprod Biol X*. 2023;21:100270. doi:[10.1016/j.eurox.2023.100270](https://doi.org/10.1016/j.eurox.2023.100270)
135. Farhangniya M, Rabori ED, Kermani RM, Haghdoost AA, Bahrapour A, Bagheri P, et al. Comparison of Congenital Abnormalities of Infants Conceived by Assisted Reproductive Techniques versus Infants with Natural Conception in Tehran. *Int J Fertil Steril*. 2013;7(2):82-86.
136. Zhu J, Zhu Q, Wang Y, Wang B, Lyu Q, Kuang Y. Comparative study on risk for birth defects among infants after in vitro fertilization and intracytoplasmic sperm injection. *Syst Biol Reprod Med*. 2019;65(1):54-60. doi:[10.1080/19396368.2018.1554012](https://doi.org/10.1080/19396368.2018.1554012)
137. Canfield MA, Honein MA, Yuskiv N, Xing J, Mai CT, Collins JS, et al. National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999-2001. *Birth Defects Res A Clin Mol Teratol*. 2006;76(11):747-756. doi:[10.1002/bdra.20294](https://doi.org/10.1002/bdra.20294)
138. Corsello G, Giuffre M. Congenital malformations. *J Matern Fetal Neonatal Med*. 2012;25 Suppl 1:25-29. doi:[10.3109/14767058.2012.664943](https://doi.org/10.3109/14767058.2012.664943)
139. Wang M, Huang R, Liang X, Mao Y, Shi W, Li Q, et al. Recombinant LH supplementation improves cumulative live birth rates in the GnRH antagonist protocol: a multicenter retrospective study using a propensity score-matching analysis. *Reprod Biol Endocrinol*. 2022;20(1):114. doi:[10.1186/s12958-022-00985-4](https://doi.org/10.1186/s12958-022-00985-4)
140. American Diabetes Association Professional Practice Committee. Management of Diabetes in Pregnancy: Standards of Care in Diabetes—2025. *Diabetes Care*. 2025;48(1 Suppl 1):S306-S320. doi:[10.2337/dc25-S015](https://doi.org/10.2337/dc25-S015)
141. Alsbjerg B, Humaidan P. What to expect from a “standard vaginal progesterone regimen” in hormone replacement therapy frozen embryo transfer (HRT-FET) -- a PRISMA review and meta-analysis. *Reprod Biomed Online*. 2025;50(5):104736. doi:[10.1016/j.rbmo.2024.104736](https://doi.org/10.1016/j.rbmo.2024.104736)
142. Rinaldi L, Crescenzi F, Selman H. Oral dydrogesterone along with vaginal micronized progesterone supplementation for luteal phase support in IVF patients, and its impact on pregnancy and live birth rates: a prospective randomized trial. *BMC Pregnancy Childbirth*. 2024;24(1):845. doi:[10.1186/s12884-024-07069-8](https://doi.org/10.1186/s12884-024-07069-8)
143. Etwel F, Koren G. Bias against the null hypothesis in retrospective registries of gestational drug exposure. *J Obstet Gynaecol Can*. 2016;38(12):1120-1123.e1. doi:[10.1016/j.jogc.2016.09.077](https://doi.org/10.1016/j.jogc.2016.09.077)
144. Elizur SE, Tulandi T. Drugs in infertility and fetal safety. *Fertil Steril*. 2008;89(6):1595-1602. doi:[10.1016/j.fertnstert.2008.02.092](https://doi.org/10.1016/j.fertnstert.2008.02.092)
145. Jia Y, Wang J, Liu C, Zhao P, Ren Y, Xiong Y, et al. The methodological quality of observational studies examining the risk of pregnancy drug use on congenital malformations needs substantial improvement: a cross-sectional survey. *Drug Saf*. 2024;47(11):1171-1188. doi:[10.1007/s40264-024-01465-x](https://doi.org/10.1007/s40264-024-01465-x)

SUPPLEMENTARY MATERIALS

Supplementary Material S1: Detailed Search Strategies

Download: <https://jivfww.scholasticahq.com/article/147524-safety-of-pharmacological-agents-used-in-assisted-reproductive-technology-a-comprehensive-systematic-review-of-fetal-malformation-risk/attachment/325152.pdf>
