

Supplementary Table 1 PRISMA-ScR search methodology details

Source/PRISMA stage	Search strategy/Screening detail	Number of records
Records identified from PubMed	Records identified from PubMed database from January 1, 2000 to June 10, 2026 using search terms: (("Glucagon-Like Peptide-1 Receptor Agonists"[MeSH Terms] OR "glucagon like peptide 1"[MeSH Terms] OR "glucagon like peptide 1"[Title/Abstract] OR "GLP-1"[Title/Abstract] OR "GLP1"[Title/Abstract] OR "glp 1 receptor agonist"[Title/Abstract] OR "semaglutide"[Title/Abstract] OR "liraglutide"[Title/Abstract] OR "dulaglutide"[Title/Abstract] OR "exenatide"[Title/Abstract] OR "lixisenatide"[Title/Abstract] OR "albiglutide"[Title/Abstract] OR "tirzepatide"[Title/Abstract] OR "ozempic"[Title/Abstract] OR "wegovy"[Title/Abstract] OR "rybelsus"[Title/Abstract] OR "saxenda"[Title/Abstract] OR "victoza"[Title/Abstract] OR "trulicity"[Title/Abstract] OR "byetta"[Title/Abstract] OR "bydureon"[Title/Abstract] OR "lyxumia"[Title/Abstract]) AND ("Pregnancy"[MeSH Terms] OR "pregnant"[Title/Abstract] OR "gestation"[Title/Abstract] OR "prenatal"[Title/Abstract] OR "antenatal"[Title/Abstract] OR	140

"periconception*" [Title/ Abstract] OR
 "preconception*" [Title/ Abstract] OR "first
 trimester" [Title/ Abstract] OR "miscarriage" [Title/ Abstract] OR
 "spontaneous abortion" [Title/ Abstract] OR "birth
 defect*" [Title/ Abstract] OR "congenital" [Title/ Abstract] OR
 "malformation*" [Title/ Abstract] OR "fetal" [Title/ Abstract] OR
 "foetal" [Title/ Abstract] OR "neonatal" [Title/ Abstract] OR
 "offspring" [Title/ Abstract]) AND "humans" [MeSH Terms]
 AND 2000/1/1:2026/6/10 [pdat]))

Records identified from Google Scholar	Full query, ranking method, filters, and rationale: Exact Google Scholar query entered on June 10, 2026: ("GLP-1 receptor agonist" OR semaglutide OR liraglutide OR dulaglutide OR exenatide OR lixisenatide OR albiglutide OR tirzepatide) AND (pregnancy OR pregnant OR preconception OR periconception OR "first trimester" OR miscarriage OR congenital OR neonatal OR offspring)	27,700 retrieved; first 500
	manually screened	

Google Scholar default relevance ranking; no date, language, or document-type filters were applied. The first 500 relevance-ranked records from 27,700 records found were manually screened. This threshold was selected because it exceeded the PubMed yield of 140 records by more than threefold, while remaining feasible for manual screening.

Total records identified/screened	640
Duplicate records removed manually	224
Records screened after de-duplication	416
Records excluded after title/abstract screening	359
Full-text reports assessed for eligibility	57
Full-text reports excluded	10
Reports included in evidence map	47
Core evidence synthesis	41

Supplementary Table 2 JBI critical appraisal checklist for included case reports

Case report, Year	D1	D2	D3	D4	D5	D6	D7	D8	Main concern
Williams et al. [4], 2009	Y	Y	Y	Y	U	Y	Y	Y	High-risk T2D background; intervention/exposure details incomplete
Greco et al. [5], 2015	Y	Y	Y	U	Y	Y	U	Y	Short infant follow-up; limited diagnostic detail
Ivanisevic et al. [6], 2018	Y	Y	U	Y	Y	U	NA	Y	Single prolonged exposure case; uncontrolled diabetes confounding
Livadas et al. [7], 2020	Y	U	U	U	Y	U	U	Y	Sparse pregnancy and neonatal outcome detail
Burlina et al. [8], 2023	Y	U	Y	Y	Y	Y	Y	Y	T2D/obesity confounding; limited follow-up
Alghamdi et al. [9], 2023	Y	U	U	Y	Y	Y	Y	Y	Limited fetal follow-up; isolated mild pyelectasis
Skov et al. [10], 2023	Y	U	U	Y	Y	Y	U	Y	Macrosomia/shoulder dystocia confounded by metabolic risk
Wajtryt et al. [11],	Y	U	U	U	U	U	U	Y	Sparse clinical, exposure, and follow-up

2023									details
Patel et al. [12], 2024	Y	U	Y	Y	Y	Y	Y	Y	Preeclampsia/polyhydramnios in high-risk metabolic pregnancy
Molteni et al. [13], 2024	Y	Y	Y	Y	Y	Y	NA	Y	Prolonged exposure but single case; metabolic confounding
Okeke et al. [14], 2024	Y	U	U	U	U	U	Y	Y	Single-dose exposure; limited outcome detail; psychiatric/obstetric confounding
Zhu et al. [15], 2024	Y	N	U	Y	Y	Y	Y	Y	Isolated anomaly; history/timeline incomplete
Abushanab et al. [16], 2025	Y	Y	Y	U	Y	U	Y	Y	Large gestational weight gain; developmental follow-up reassuring but single case
Burlina et al. [17], 2026	Y	U	U	Y	Y	U	Y	Y	Prolonged exposure, reassuring, no follow-up
Huang et al. [18], 2025	Y	U	Y	Y	Y	U	Y	Y	Elective termination and severe intercurrent illness prevent fetal inference
Ozdemir et al. [19], 2025	Y	U	Y	Y	Y	Y	Y	Y	Preterm/NICU outcome; limited confounder detail
Jin et al. [20], 2026	Y	Y	Y	Y	Y	Y	Y	Y	Short infant follow-up only

The JBI checklist was used for domain-level contextual appraisal. No numerical score or JBI-defined overall quality grade was assigned, and no report was excluded solely on the basis of appraisal findings. Abbreviations: JBI, Joanna Briggs Institute; T2D, type 2 diabetes; NICU, neonatal intensive care unit; Y, yes; U, unclear; N, no; NA, not applicable.

JBI case report domains: D1, demographic characteristics clearly described; D2, patient history clearly described and presented as a timeline; D3, current clinical condition clearly described; D4, diagnostic tests/methods and results clearly described; D5, intervention/exposure, monitoring process and management clearly described; D6, post-intervention clinical condition clearly described; D7, adverse or unanticipated events identified and described; D8 takeaway lessons provided. (Y = Yes; U = Unclear; N = No; NA = Not applicable)

Supplementary Table 3 JBI critical appraisal checklist for included case series

Case series	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Main concern
Dogan et al. [21], 2024	Y	Y	Y	N	U	Y	Y	Y	Y	NA	Two pregnancies in the same woman; not consecutive case accrual
Banerjee et al. [22], 2026	Y	Y	Y	U	U	Y	Y	U	Y	NA	Very small case series; limited follow-up
Morton et al. [23], 2026	Y	Y	Y	U	U	Y	Y	U	Y	NA	Largest case series but selection/consecutive inclusion unclear

The JBI checklist was used for domain-level contextual appraisal. No numerical score or JBI-defined overall quality grade was assigned, and no report was excluded solely on the basis of appraisal findings. Abbreviations: JBI, Joanna Briggs Institute; Y, yes; U, unclear; N, no; NA, not applicable.

JBI case series domains: D1, clear criteria for inclusion; D2, condition/exposure measured in a standard and reliable way; D3, valid methods used for identification of condition/exposure; D4, consecutive inclusion of participants; D5, complete inclusion of participants; D6, demographics clearly reported; D7, clinical information clearly reported; D8, outcomes and follow-up clearly reported; D9, site/setting clearly reported; D10, appropriate statistical analysis. (Y = Yes; U = Unclear; N = No; NA = Not applicable)

Supplementary Table 4 JBI contextual appraisal checklist for cohort, registry, and controlled human studies

Study, Year	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D1	Main concern
											1	
Chou et al. [26], 2026	Y	Y	U	Y	Y	Y	Y	Y	Y	U	Y	Small exposed sample; GLP-1RA agent not specified
Cesta et al. [27], 2024	Y	Y	U	Y	Y	Y	Y	U	Y	NA	Y	Prescription-based exposure; limited glycemic/BMI granularity
Hanif et al. [28], 2025	Y	Y	U	Y	Y	Y	Y	Y	Y	U	Y	Agent unspecified; residual confounding possible
Lewin et al.	U	Y	U	U	N	Y	Y	NA	NA	NA	Y	Baseline disease-severity

[29], 2026													imbalance; broad exposure window; possible T1D admixture
Bolek et al. [30], 2026	NA	Y	Y	U	N	Y	Y	Y	Y	NA	NA	No comparator; n=20; preconception discontinuation	
Bartelt et al. [33], 2024	N	Y	U	U	U	Y	Y	NA	NA	NA	Y	General population comparator; major obesity confounding	
Guo et al. [31], 2025	Y	Y	Y	U	U	Y	Y	Y	Y	NA	Y	Very small exposed sample; preconception exposure only	
Li et al. [32], 2022	Y	Y	Y	Y	Y	Y	Y	Y	Y	NA	Y	Fertility/preconception study, not designed for fetal safety	
Hviid et al. [39], 2026	Y	Y	U	Y	Y	Y	Y	NA	NA	NA	Y	Mixed indications; residual confounding after matching	
Dao et al. [34], 2024	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Referral/self-selection and elective termination bias	
Kolding et al. [35], 2025	Y	Y	Y	N	N	Y	Y	NA	Y	NA	Y	Very small exposed sample; many exposed women also on insulin	
Imbroane et al. [36], 2025	Y	Y	U	Y	Y	Y	Y	NA	NA	NA	Y	Preconception exposure; healthy-user/weight-loss confounding	
Pondugula et al. [37],	Y	Y	U	Y	Y	Y	Y	NA	NA	NA	Y	Small exposed sample; heterogeneous indications and	

2026												agents
Maya et al. [38], 2025	Y	Y	U	Y	Y	Y	Y	NA	NA	NA	Y	Exposure up to 3 years before conception; weight-regain confounding
Broe et al. [40], 2025	N	Y	U	Y	Y	Y	Y	Y	Y	Y	Y	General population comparator; confounding by indication
Brown et al. [41], 2026	Y	Y	U	Y	Y	Y	Y	NA	NA	NA	Y	Confounding by continuation, pregnancy intention, and clinician counseling

Abbreviations: JBI, Joanna Briggs Institute; GLP-1RA, glucagon-like peptide-1 receptor agonist; T1D, type 1 diabetes; T2D, type 2 diabetes; BMI, body mass index; Y, yes; U, unclear; N, no; NA, not applicable.

JBI cohort domains: D1, groups similar and recruited from same population; D2, exposures measured similarly; D3, exposure measured in a valid and reliable way; D4, confounders identified; D5, strategies to deal with confounders stated; D6, groups free of outcome at start/exposure; D7, outcomes measured in a valid and reliable way; D8, follow-up time reported and sufficient; D9, follow-up complete or loss described; D10, strategies to address incomplete follow-up; D11, appropriate statistical analysis used. (Y = Yes; U = Unclear; N = No; NA = Not applicable).

Supplementary Table 5 Design-adapted appraisal of non-cohort primary evidence

Study, Year	Evidence type	Standard tool fit	Appraisal domains used	Strengths	Key limitations	Design-specific contextual appraisal	Role in synthesis
Zhou et al. [24], 2025	FAERS pharmacovigilance	JBICohort not appropriate	Case definition, exposure coding, signal method, comparator sensitivity, duplicate/reporting bias, causality limits	Large real-world signal detection; multiple GLP-1RAs; ROR and Bayesian IC; pregnancy-specific comparator sensitivity	No denominator; no incidence; no causality; stimulated reporting; confounding; exposure timing incomplete	Moderate for signal detection; low for causal inference	Signal-generating only
Parker et al. [25], 2025	Regulatory clinical-trial inadvertent	JBICohort only partially applicable	Regulatory data source, exposed/comparator pregnancy counts, outcome completeness,	Structured regulatory trial data; class-wide agents; some comparator pregnancies	Pregnancy was exclusion criterion; post hoc extraction; small n; missing outcomes;	Moderate for descriptive evidence; low for causal	Intermediate evidence between case reports and cohorts

pregnanc
y dataset

follow-up,
confounding,
causal
interpretation

adjusted
comparison

inference

Abbreviations: FAERS, Food and Drug Administration Adverse Event Reporting System; GLP-1RA, glucagon-like peptide-1 receptor agonist; JBI, Joanna Briggs Institute; ROR, reporting odds ratio; IC, information component.

Supplementary Table 6 AMSTAR-2 critical-domain appraisal of systematic reviews

7 Critical domains of AMSTAR-2	Definition
C1 / Item 2	Protocol registered or predefined before conduct of the review
C2 / Item 4	Comprehensive literature search strategy
C3 / Item 7	List of excluded studies with justification
C4 / Item 9	Risk-of-bias (RoB) assessment of included studies
C5 / Item 11	Appropriate meta-analytic methods, if meta-analysis was performed
C6 / Item 13	Risk of bias (RoB) considered when interpreting/discussing results
C7 / Item 15	Publication bias or small-study bias assessed when quantitative synthesis was performed

AMSTAR-2 contains 16 items; this table presents the seven critical domains to improve readability. Non-critical limitations are summarized narratively. Reviews without meta-analysis were appraised for contextual value only and

were not used as primary evidence for risk estimation. (Y = Yes; PY = Partial yes; N = No; NR = not reported/unclear from available text; NA = not applicable)

Supplementary Table 7 AMSTAR-2 methodological qualities of systematic reviews

Review, Year	Type	C1 Protoc ol	C2 Searc h	C3 Exclud ed list	C4 assessme nt	RoB	C5 MA metho d	C6 interpretati on	RoB in	C7 Publicatio n bias	Key critical concerns	non-	Context ual confide nce	Use review	in
Kodali et al. [42], 2026	SRMA	No	PY	No	No		No	Yes		No	Only observational studies; limited search; clearly reported protocol, duplicate processes, excluded-study list, or detailed	4	Criticall y low	Cautious secondar y synthesis	

									RoB table; text-figure inconsistency for adverse fetal outcome.		
Armstrong et al. [43], 2025	SRMA	Yes	PY	PY	Yes	No	Yes	No	Heterogeneous evidence pooled, including Parker regulatory-trial pregnancy data and preconception/fertility evidence; internal inconsistency in the number of included	Low	Useful, but cautious interpretation

										studies.		
Hakim et al. [44], 2026	SRMA	Yes	Yes	PY	Yes	Yes	Yes	Yes	Yes	Observational overlap; comparator heterogeneity; exposure misclassification; renal signal largely driven by one large cohort.	Moderate	Stronger among current SRMA
Muller et al. [45], 2023	SR without MA	NR	PY	NR	NR	No	No	No	No	Animal-dominant review; sparse human pregnancy data; no	Critically low	Background only

										quantitative synthesis.		
Minis et al. [46], 2023	SR without MA	NR	PY	NR	NR	No	No	No	Mixed evidence including case reports, case series, preconception RCTs and follow-up data; no risk estimation.	Criticall y low	Context only	
Kuitune n et al. [47], 2024	SR without MA	NR	PY	NR	NR	No	No	No	Very few human studies; maternal outcomes not meaningfully assessed; no meta-analysis.	Criticall y low	Context only	

Drumm ond et al. [48], 2025	SR without MA	NR	PY	NR	NR	No	No	No	Mixed reports early surveillance/co hort data; washout recommendatio n largely pharmacokinet ic rather than outcome- derived.	case and y low	Criticall	Counsell ng context
Mandal et al. [49], 2026	SR without MA; semaglut ide- focused	NR	PY	NR	NR	No	No	No	Semaglutide- specific, but based on overlapping case reports, cohorts and pharmacovigila nce; exposure	but on	Criticall y low	Semaglut ide context

									windows heterogeneous.		
Ozbek et al. [50], 2026	Broad SR/scoping-style review without MA	NR	PY	NR	NR	No	PY	NA	Broad evidence map but highly heterogeneous: PCOS, obesity, diabetes, case reports, cohorts, pharmacovigilance, continuation-through-pregnancy and lactation data.	Low	Contextual mapping

All 16 AMSTAR-2 items were considered in determining contextual confidence; this table displays the seven critical domains, while relevant non-critical weaknesses are summarized narratively. Abbreviations: AMSTAR-2, A Measurement Tool to Assess Systematic Reviews-2; MA, meta-analysis; RoB, risk of bias; SR, systematic review; SRMA, systematic review and meta-analysis; Y, yes; PY, partial yes; N, no; NR, not reported or insufficiently reported; NA, not applicable; PCOS, polycystic ovary syndrome.