PROTOCOL FOR A SYSTEMATIC REVIEW OF

Exposure to Glyphosate and Glyphosate Formulations and Male Reproductive Health

February 17, 2026

Integrative Health Assessments Branch
Division of Translational Toxicology
National Institute of Environmental Health Sciences
National Institutes of Health

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

This revised protocol finalized 02/17/2026.
Signed on behalf of the Evaluation Team ANDREW A. ROONEY

TABLE OF CONTENTS

Table of Contents ..... **ii**
List of Tables ..... **iii**
List of Figures ..... **iii**
Background and Significance **4**
Background 4
Significance 5
Objective and Specific Aims **6**
Objective 6
Specific Aims 6
PECO Statement 6
Methods **7**
Problem Formulation 7
General Considerations 7
Glyphosate Exposure Considerations and Non-Cancer Outcome Selection 8
Search and Select Studies for Inclusion 9
Literature Search Strategy 9
Screening Process 10
Data Extraction and Content Management 13
Quality Assessment of Individual Studies 14
Exposure Characterization 17
Timing of Exposure and Outcome Measurements 18
Other Key Risk-of-Bias Assessment Domains 18
Missing Information for Risk-of-Bias Assessment 19
Assessment of Confidence in the Body of Evidence 19
Evidence Synthesis 20
Confidence Assessment for Glyphosate or Glyphosate Formulations 21
Future Steps 21
Systematic Review Outline **22**
Introduction 22
Objective and Specific Aims 22
Methods 22
Results 22
Discussion 22
References **23**
About the Protocol **27**
Contributors 27
Federal Staff 27
Contract Staff 27
NIEHS Protocol Reviewers 28
Protocol History and Revisions 28
Sources of Support 28
Appendices **29**

Appendix 1. Catalog of Concurrent Similar Systematic Reviews 29
Appendix 2. Literature Search Strategy 34
Appendix 3. Data Extraction Elements for Animal Studies 37
Appendix 4. Risk-of-bias Criteria 39
OHAT Risk of Bias Tool 39
Experimental Animal Studies 40

LIST OF TABLES

Table 1. PECO (Populations, Exposures, Comparators, and Outcomes) Statement 6
Table 2. Detailed Inclusion and Exclusion Criteria to Determine Study Eligibility 11
Table 3. Response Options for Each Risk-of-bias Question 15
Table 4. OHAT Risk-of-bias Questions and Applicability by Study Design 16
Table 5. Factors for Consideration when Combining Data from Animal Studies 21

LIST OF FIGURES

Figure 1. Assessing Confidence in the Body of Evidence 20

BACKGROUND AND SIGNIFICANCE

Background

Glyphosate, or N-(phosphonomethyl) glycine, is a broad-spectrum synthetic herbicide. When applied on a leaf, glyphosate is absorbed into plant tissues and then binds and inhibits 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase (Boocock and Coggins 1983), an essential enzyme in protein production and survival of all plants. Glyphosate is believed to be inactivated by binding to soil (Malik et al. 1989). For general use, typical instructions recommend that it should only be sprayed on the targeted weeds, though glyphosate is also widely used as a pre-harvest crop desiccant to speed up harvest operations (Benbrook 2016). However, when glyphosate is used for genetically modified (GM), glyphosate-tolerant crops (e.g., Roundup Ready soybeans, corn, or cotton), the spray can be applied on the whole field to kill only weeds because the GM crops carry a bacterial gene producing a specific form of EPSP synthetase that does not lose activity when bound by glyphosate (Funke et al. 2006; Rogers et al. 1983).

Most human exposure is to glyphosate formulations, which are commercial products that contain glyphosate as well as surfactants or other ingredients designed to improve adhesion and uptake, as pure glyphosate is rarely used outside occupational settings. Formulations vary widely in their components to achieve the desired effects on different crops in different weather conditions and to meet registration requirements, which also vary by region. Although there have been several patents covering glyphosate or glyphosate salts in the United States, the expiration of one of the patents covering glyphosate use as an herbicide in 2000 (Franz 1983; Moschini et al. 2019) has allowed many companies to make their own glyphosate products since that time. While several salt forms of glyphosate are used by different companies, the same salt form of glyphosate is typically used within a product line. The exact composition of glyphosate formulations/products is confidential business information, but the active ingredients and their percentage by weight in the product are required to be clearly labeled. It is possible the same product name with different formulations might be on the market simultaneously because formulation changes (including changing the active ingredient) do not require a product name change. For instance, beginning in 2024, glyphosate-free residential Roundup brand products (e.g., for lawn and garden use) became available on the market, while residential Roundup products with glyphosate might also remain available because of their relatively long shelf life (3 – 5 years). As of January 2026, professional-grade Roundup formulations (e.g., Aquamaster, Roundup Pro series) continue to be made with glyphosate.

Potential exposure to glyphosate is expected to be highest in the occupational setting, including glyphosate manufacturers, professional applicators, and agricultural workers. Dermal exposure to glyphosate in the workplace is the main contributor to the total body burden, with inhalation and inadvertent ingestion also playing a role (Connolly et al. 2019; Pierce et al. 2020). Food is the main route of exposure for the general population (Mellor et al. 2024; Mohammadi et al. 2025), and other potential sources include breastmilk (Camiccia et al. 2022), drinking or surface water contaminated with glyphosate from agricultural runoff (Li and Kannan 2025; Medalie et al. 2020), and inhalation or dermal exposure to glyphosate in the environment from agricultural and nonagricultural uses [e.g., house dust (Li et al. 2024)]. Using analysis of glyphosate in urine of representative populations, the National Health and Nutrition Examination Survey (NHANES) data from 2013-2014 indicate that over 80% of the U.S. population that are 6 years of age or older have been exposed to glyphosate (Sun et al. 2024; Wu et al. 2025), and the geometric mean concentration was 0.41 ng/mL in urine (Ospina et al. 2022). Human exposures to glyphosate often include co-exposure to glyphosate metabolites, mainly aminomethyl

phosphonic acid (AMPA), N-acetylglyphosate (NAG), and N-acetyl-AMPA (NA-AMPA), which can be generated by various bacteria, fungi, and plants.

Potential health risks of glyphosate exposure have been evaluated by several authoritative bodies including the European Food Safety Authority (EFSA 2023), the U.S. Environmental Protection Agency (USEPA 2017), and the Agency for Toxic Substances and Disease Registry (ATSDR 2020). In the ATSDR review, the authors indicated that developmental, endocrine and reproductive toxicity were potential endpoints of concern regarding exposure to glyphosate formulations to rats and rabbits due to reported higher rates of sperm abnormalities, reductions in testosterone levels in males, and histopathological anomalies and decreased sperm counts in the testes of developmentally exposed animals (ATSDR 2020). A systematic review and meta-analysis by Cai et al. (2017) evaluated the English and Chinese language literature for an association between glyphosate exposure (not stratified by technical grade or formulations) and male reproductive toxicity in mice and rats. The authors identified 8 studies in rats and mice and reported that glyphosate exposure resulted in a significant decrease in sperm concentration that was consistent across many statistical analyses. A more recent systematic review and meta-analysis of in vivo studies in rats evaluated the effects of glyphosate (technical grade and glyphosate formulations) on reproductive hormones; the authors reported that glyphosate exposure (defined broadly) resulted in significant decreases in testosterone, follicle-stimulating hormone, and luteinizing hormone (Mohammadi et al. 2022).

The U.S. Department of Health and Human Services (DHHS) is interested in examining the potential association between exposure to pesticides such as glyphosate and adverse human health outcomes (e.g., MAHA Commission 2025). Given the DHHS interest and the male reproductive toxicity studies identified by ATSDR in 2020, this rapid systematic review was developed to critically evaluate the current scientific evidence on the association between glyphosate exposure and adverse male reproductive health outcomes from studies in non-human mammalian research models. The Division of Translational Toxicology (DTT) will conduct a systematic review of the published studies in non-human mammals to assess the evidence that exposure to glyphosate and glyphosate formulations is associated with male reproductive endpoints as described in this protocol. This systematic review will follow the OHAT framework (Rooney et al. 2014) with any modifications to decrease typical timelines (e.g., narrow the scope) or complete a rapid systematic review noted in the report and the potential impact of those rapid approaches on the confidence of the conclusions noted in the limitations section of the report. Following the problem formulations described in the methods, this review will comprehensively search for published studies of glyphosate exposure in non-human mammalian animal models (that may be further restricted to rodents and rabbits due to time constraints) that address endpoints related to male reproduction, assess risk of bias (internal validity) of the relevant studies, rate confidence in the bodies of evidence (groups of studies on the same or related outcomes), and develop conclusions reflecting those confidence ratings on whether exposure to glyphosate is associated with alterations in male reproduction based on the available animal data. Conclusions will be considered for glyphosate and glyphosate formulations separately, as well as subgroups and the combined evidence for glyphosate and glyphosate formulations overall depending on the availability of sufficient evidence to support specific conclusions. Implications for male reproductive health effects in humans will be described, noting the limitations of the review (including that the human epidemiological studies will not be evaluated for this report).

Significance

As described above, multiple studies have reported that exposure to glyphosate is potentially associated with altered reproductive hormones and lower sperm count and quality in experimental animal models [primarily from rodent data reviewed in ATSDR (2020)]. Therefore, this systematic review was designed

to critically assess experimental evidence from studies in non-human mammalian animal research models to determine whether glyphosate or glyphosate formulations are associated with alterations in male reproductive health, including hormones and sperm parameters. This review will reflect the most current information available, including relevant studies from previous authoritative reviews, in addition to studies published after those reviews were developed, and will be evaluated using a single consistent, transparent, and high-quality method to inform decision making.

OBJECTIVE AND SPECIFIC AIMS

Objective

The main objective of the systematic review outlined in this protocol is to assess the association between glyphosate and glyphosate formulations and alterations in male reproductive health based on an assessment of the experimental evidence from studies in non-human mammalian research models. This systematic review will build upon the identification of relevant literature from authoritative reviews and evaluate the relevant literature to the present time.

Specific Aims

- Identify the published scientific literature that examines the association between glyphosate and glyphosate formulations and effects or outcomes related to male reproductive health in non-human mammalian animal models (as described in problem formulation section, this may be limited to rodents and rabbits given the time constraints of the review).
- Extract data on endpoints or outcomes related to male reproductive health from relevant studies.
- Assess internal validity (i.e., risk of bias) on individual studies.
- For glyphosate and for glyphosate formulations, separately:
    - Synthesize the evidence across studies using a narrative approach, or meta-analysis (if appropriate), and evaluate sources of heterogeneity.
    - Rate the confidence in the body of evidence for effects on male reproductive health outcomes according to one of four statements: *High, Moderate, Low*, or *Very Low/No Evidence* available.
    - Characterize uncertainty, identify key data gaps, and research needs related to limitations of the evidence base and limitations of the systematic review.

PECO Statement

A PECO (**P**opulations, **E**xposures, **C**omparators, and **O**utcomes) statement was developed to address and understand the potential effects of glyphosate and glyphosate formulations on male reproductive health reported in non-human mammalian animal model systems (Table 1). The PECO statement is used to help develop the specific research questions, search terms, and inclusion/exclusion criteria for the systematic review (Higgins et al. 2011).

Table 1. PECO (Populations, Exposures, Comparators, and Outcomes) Statement

PECO Element	Evidence
Populations	Non-human mammalian research models (this may be limited to rodents and rabbits given the time constraints of the review), without restriction as to age at exposure
Exposures	Exposure to glyphosate (including glyphosate metabolites) or glyphosate

Table 1. PECO (Populations, Exposures, Comparators, and Outcomes) Statement	
PECO Element	**Evidence**
	formulations (i.e., commercial herbicides that contain glyphosate as an active ingredient, along with other ingredients) via any exposure route. Exposure metrics in experimental animal models will include external or administered dose, biomarker measurements, and/or modeled exposures.
Comparators	Study groups exposed to vehicle only, unexposed (not exposed to glyphosate, glyphosate metabolites, and/or glyphosate formulations). Note, if a vehicle only or unexposed group was not included, studies may be considered if treatment groups include both higher and lower levels of glyphosate exposure.
Outcomes	Measured or modeled outcomes related to male reproductive health including, but not limited to: • Organ weight, morphology and histopathology (e.g., testes, epididymis, seminal vesicles, prostate) • Hormone levels (e.g., testosterone, follicle stimulating hormone, luteinizing hormone, estrogen, prolactin) • Sperm characteristics (e.g., daily production, count, morphology, motility) • Developmental outcomes (e.g., ano-genital distance [AGD], testis descent, preputial separation)

METHODS

Problem Formulation

Given the time constraints of this review (to be completed in less than 6 months), the evaluation team made several decisions to narrow the scope of the evaluation: 1) the systematic review will be limited to an evaluation of the experimental animal data); 2) this evaluation will be developed as a government report by June, 2026 with a follow up peer-reviewed journal publication that may address additional considerations (e.g., evaluation of human epidemiological studies and mechanistic data) in a longer time frame; 3) authoritative reviews will be used as a guide for considering potential non-cancer health effects for which a targeted systematic review could add value as a high-quality, updated assessment of the evidence; 4) the systematic review will follow the OHAT framework (Rooney et al. 2014) with any modifications to facilitate a rapid review format as noted; and 5) the systematic review will transparently describe the potential impact or uncertainties in the conclusions of this systematic review that might result from narrowing the scope or other techniques to more rapidly complete the review in a shortened time frame (note, this will be included in addition to the normal limitations of this review and limitations of this database sections of the systematic review).

General Considerations

Problem formulation was informed by a review of authoritative documents including the 2020 Agency for Toxic Substances and Disease Registry (ATSDR) Toxicological Profile for Glyphosate (ATSDR 2020). The U.S. EPA's most recent evaluation (EPA, 2017) was focused on cancer outcomes and was therefore

not informative in identifying non-cancer outcomes where a critical review of the experimental animal scientific evidence would have substantive value.

NIEHS team members for this project or the "evaluation team" include experts in public health, epidemiology, reproductive toxicology, systematic review, and chemical mixtures. Throughout the problem formulation process, DTT scientists engaged NIEHS subject matter experts on reproductive and development toxicology and environmental chemicals to provide input and feedback on important considerations in assessing glyphosate exposures and male reproductive outcomes.

Next, searches were conducted in PROSPERO (https://www.crd.york.ac.uk/prospero/) and Epistemonikos (https://www.epistemonikos.org), two of the largest registries of systematic review protocols or published evaluations, to check whether there were any ongoing systematic evidence maps, systematic reviews, and/or meta-analyses on the topic of glyphosate and glyphosate formulations and alterations in male reproductive health. These searches promote awareness of current reviews that are being conducted and ensure potentially relevant studies identified in those evaluations are included. The detailed search strategy and search results for Epistemonikos and PROPERO are provided in Appendix 1.

Glyphosate Exposure Considerations and Non-Cancer Outcome Selection

Glyphosate exposure considerations: Assessment of potential health effects of glyphosate are complicated by the range of potential exposures considered within individual studies. Exposure of the general public is typically to glyphosate formulations, which include different salts, surfactants, and even additional herbicides. Glyphosate formulations differ geographically with multiple formulations used in different products, different chemical ingredients across countries, and changes in product formulations over time. Therefore, in addition to typical consideration of metabolites and how chemical exposure is assessed in people, exposure assessment for glyphosate in human populations needs to consider whether the target of the evaluation is glyphosate alone or glyphosate formulations and the complication of different chemicals included within those formulations. The review team determined that this assessment would review the evidence for health effects associated with glyphosate alone as well as glyphosate formulations. However, given the focus of the review will be on experimental evidence from studies in non-human mammalian research models, the team expect that there will be less variation in formulations tested in animals than would be the case for observational human exposure studies.

The literature search for glyphosate, glyphosate formulations, and metabolites was based on the method and search string developed in the ATSDR 2020 authoritative review. Given the time constraints, this search string was tested and quickly deployed for this assessment. The evaluation team also planned for additional testing of the search string and adjustments as needed, recognizing that some potential glyphosate metabolites were not included in the original literature search strategy. While metabolites are more likely to be used to characterize exposure in human studies than in animal studies, the team will develop a comprehensive search string of glyphosate metabolites to test if searches for the metabolites alone will identify additional studies assessing glyphosate exposure and male reproductive outcomes in non-human mammalian models. Similarly, the ATSDR 2020 search string was judged robust but potentially not comprehensive in coverage of glyphosate formulation names. The literature search string developed from ATSDR 2020 is expected to capture all relevant published studies on glyphosate with the current strategy (e.g., most would be identified with the "glyphosate" term) without having all formulation names as part of the search text. An additional follow up assessment will

be conducted to develop and conduct a comprehensive search string of glyphosate formulation names to test if searches for the formulation names alone identified any additional glyphosate studies in non-human animal models with data on male reproductive outcomes. The results of the additional searches for metabolite-alone and formulation-name-alone studies will be included if there is sufficient time to capture these data, assess risk of bias for these studies, and include them in the evidence synthesis or the potential impact of missing these studies would be described in the limitations section.

Outcome selection and the focus on non-human mammalian data: An assessment of male reproductive toxicity was considered for the target outcome because the 2020 ATSDR review indicated that male endocrine and reproductive toxicity were potential endpoints of concern due to reported higher rates of sperm abnormalities, reductions in testosterone levels, and histopathological anomalies and decreased sperm counts in the testes of developmentally exposed animals (ATSDR 2020). Systematic reviews published around the same time also reported an association between glyphosate exposure and decreased sperm count (Cai et al. 2017) and altered reproductive hormones (Mohammadi et al. 2022). The review team also determined that limiting the focus to rodent and rabbit data rather than all experimental mammals could speed up the literature screening process and may have minimal impact on the conclusions given that all of the male reproductive data described in ATSDR were from studies in rodents and rabbits. A separate effort would identify additional non-rodent mammal studies with potentially relevant results that would be included if there are sufficient time to capture these data, assess risk of bias for these studies, and include them in the evidence synthesis.

Documentation of decisions to narrow the scope to meet the time constraints of the review: In addition to targeting the review on non-human mammalian animal data, focusing on the rodent and rabbit data first, and limiting the scope to a single health outcome (reproductive health) in a single sex (males), the evaluation team focused on only English language publications to further narrow the scope of the systematic review. Given the complexity and potential additional time required to identify or translate non-English language studies, the team decided to track potentially relevant non-English studies if they were identified in the search, but to limit the review to English language publications. If time permits, an attempt will be made to translate non-English studies that were identified, and they will be reviewed for relevance and potential inclusion in the evaluation. A clear description of decisions to narrow the scope of the review and describe potential impact of these decisions on conclusions or certainty of conclusions will be described in the limitations section.

Search and Select Studies for Inclusion

Literature Search Strategy

A literature search strategy was developed using relevant exposure terms in combination with male reproductive health outcomes of interest (e.g., testes weight, sperm production, sperm motility). The full details of the search strategies are presented in (Appendix 3). The search string was based on the 2020 ATSDR evaluation of glyphosate combined with male reproductive outcome terms. Although ATSDR used multiple different searches to identify glyphosate-only studies separately from specific glyphosate formulation studies, DTT scientists combined the exposure terms across all the ATSDR searches into a single search string that captures both glyphosate and glyphosate formulations. The search string was then tested in PubMed to confirm it retrieved all the reproductive toxicity studies listed in ATSDR 2020. In addition, evaluation team members developed several follow up literature searches as described in the Problem Formulation section to test the rigor and comprehensiveness of

the exposure search for glyphosate metabolites and formulation names. The ATSDR-based search strategy was combined with follow up testing to ensure that all relevant literature was identified. Potential limitations associated with the literature search will be described in the limitations section of the systematic review. Any modifications implemented as a result of the search strategy testing will be noted in the report with explanation as to why the changes were made and what additional references were identified. An updated protocol will be developed if major modifications are implemented. The exposure terms of the refined literature search will include (1) common names for glyphosate, glyphosate metabolites, and glyphosate formulations (e.g., RoundUp), (2) Chemical Abstract Services Registry Numbers (CASRN), and (3) chemical synonyms from the ChemIDPlus database.

The search will be conducted in three databases considered the most relevant for health effects research on glyphosate and glyphosate formulations – PubMed, Web of Science, and Scopus, with no limits on publication year or language. Given the time constraints of the review, only English language studies will be included, and any potentially relevant non-English studies will be tracked for potential follow-up activities. An informationist will work with DTT to design and optimize the search strategy for the three databases. Literature search update(s) may be conducted to ensure that the systematic review represents current information at the time of its publication.

Screening Process

Prior to the beginning of the literature screening, results of the literature search will be assembled in EndNote software and exact reference duplicates will be removed. The use of machine learning technology during literature screening will depend upon how many unique references are retrieved from the search of all three databases. If the search strategy results exceed 3,500 unique records, the machine learning models for priority ranking and assessment of relevance will be utilized for title-and-abstract screening described below. If equal to or less than 3,500 unique records are retrieved, the team will conduct the title-and-abstract screening manually.

Use of a Validated Model for Priority Order Screening and Cut-Off Decisions

Initial title-and-abstract screening of references will identify potentially relevant references using DistillerSR®[1] (Evidence Partners, Ottawa, CA), a web-based systematic review software program with structured forms and procedures that can be used to screen articles for relevance and eligibility to ensure standardization of the process. Based on initial searches for problem formulation we expect the search results to retrieve fewer than 3,500 studies and to complete title-and-abstract screening manually; however, DistillerSR® has the ability to use priority ranking if the number of potential studies identified by the search string is greater than expected. For title-and-abstract screening, the software program includes the ability to use a statistical model that continually re-orders the remaining studies according to their likelihood for inclusion based on the manual inclusion and exclusion decisions of previous studies. Studies with a higher likelihood of inclusion are prioritized for screening first. In addition, DistillerSR® can employ DistillerSR's Artificial Intelligence System (DAISY), an active machine-learning component, to continually incorporate user feedback during title-and-abstract screening to predict the total number of included studies, thus providing a statistical basis for a decision about when to stop screening (Evidence Partners 2021). The performance of the tool for priority re-ordering and the screening cut off (i.e., predicted likelihood that 95% of relevant studies have been identified) has been

[1] DistillerSR® (http://systematic-review.net/) is a proprietary project management tool for tracking studies through the screening process and storing data extracted from these studies using user-customized forms.

validated (Hamel et al. 2020). The full text of all references identified as potentially relevant during title and abstract screening will be further screened in DistillerSR®.

Selection Criteria for Evidence

Inclusion and exclusion criteria are designed to identify relevant publications that comply with each aspect of the PECO statement. The eligibility of each citation from the literature will be considered based on the criteria outlined in Table 2.

Table 2. Detailed Inclusion and Exclusion Criteria to Determine Study Eligibility

Evidence Stream	Inclusion Criteria	Exclusion Criteria (or blank if none)
Participants/Populations (Human Studies or Experimental Model Systems)		
Animal	• Phase 1: Experimental studies in rodents (e.g., rats, mice, and guinea pigs) and rabbits • Phase 2: If there is time to identify studies, extract data, assess risk of bias, and consider results from studies from other mammals, then the review will include data from all non-human mammalian animal models	• Non-rodent and non-rabbit mammalian experimental animal models (Note: these studies will be tracked during title-and-abstract level screening, so they will be available for later consideration if there is time to add them to the systematic review) • Non-mammalian animal models (e.g., fish, amphibians, birds) • Non-experimental studies (e.g., wildlife studies)
Exposures		
Animal	• Exposure to glyphosate, including: ○ glyphosate (including acid and various salt forms of the chemical), ○ glyphosate metabolite(s), ○ glyphosate formulations (i.e., products containing glyphosate among other ingredients or listing glyphosate as an active ingredient) • Timing: No restrictions on life stage at exposure • Route: No restrictions on route of exposure • Exposure metrics ○ Exposure in experimental animal studies will include external or administered dose, internal dose or biomarker measurements, and modeled exposures	• Studies that only report data on non-glyphosate constituents of glyphosate formulations (i.e., no data on glyphosate exposure)
Comparators		
Animal	• Comparable unexposed animal subjects or animals exposed to vehicle only • Experiments without vehicle-exposed or unexposed animal groups that include lower and higher exposure levels may be considered	• None

Table 2. Detailed Inclusion and Exclusion Criteria to Determine Study Eligibility		
Evidence Stream	**Inclusion Criteria**	**Exclusion Criteria (or blank if none)**
Outcomes		
Animal	• All male reproductive endpoints, including but not limited to: ○ organ weight, morphology, and histopathology of reproductive organs (e.g., testes, epididymis, seminal vesicles, prostate), ○ hormone levels in males (e.g., testosterone, follicle stimulating hormone, luteinizing hormone, estrogen, prolactin), ○ sperm endpoints (production, count, morphology, motility), ○ developmental outcomes in males (e.g., anogenital distance, timing of testis descent, timing of preputial separation)	• Studies only reporting on reproductive endpoints in females
Publications (e.g., language restrictions, use of conference abstracts)		
	• Reference must contain original data in whole or in part relevant to the aims of the systematic review and must be peer-reviewed	• References published in a language other than English will be identified but not included in the systematic review. • Articles with no original data (e.g., editorials, reviews*) • Studies published in abstract form, or without full methods or results (e.g., grant awards, conference abstracts) • Non-peer reviewed articles (e.g., dissertations, unpublished data) • Retracted articles

*Relevant reviews may be used as background or for scanning for additional potential references.

Title-and-Abstract Screening

All screening will be done in duplicate such that each article is reviewed independently by two reviewers from the evaluation team. DTT staff will oversee all levels of screening, and adjustments or clarifications to screening criteria. A pilot title and abstract screening of at least 100 studies will be conducted to train the evaluation design team (consisting of ICF and DTT staff), identify any potential questions related to project-specific written screening instructions, refine the screening criteria, and improve accuracy and consistency among screeners. When results between reviewers disagree, ICF and DTT evaluation team members will review and resolve each disagreement for inclusion and exclusion, and feedback will be provided to the reviewers as necessary to improve consistency. If changes to the inclusion criteria are made based on the pilot phase, they will be documented in a protocol amendment along with the date modifications were made and the logic for the changes.

Following the pilot screening, title-and-abstract screening will be conducted manually if the number of unique references retrieved is equal to or less than 3,500. If the number of references retrieved is greater than 3,500, the team will proceed with title-and-abstract screening with the aid of a machine learning component of the screening platform, and relevant studies evaluating glyphosate effects on male reproduction identified in the ATSDR review will be tagged as included results to train the algorithm. Two members of the evaluation team (at least one of whom is an experienced team member)

will independently conduct title and abstract screening of a subset of the literature search results to determine whether a reference meets the inclusion criteria. References will be identified as either relevant according to PECO (e.g., original research article, relevant review, or relevant non-English study), not relevant, or unclear. Screening will continue until the machine learning re-rank prioritization model predicts that 95% of relevant references have been identified. When results between reviewers disagree during regular title-and-abstract screening, a project lead and/or senior member of the evaluation team will independently review and make a final determination of relevance. Feedback will be provided to reviewers, as needed. Reason for exclusion at the title and abstract stage will be reported in an interactive reference flow diagram (I-REFF) (Walker et al. 2024) as not relevant to PECO or predicted to be not relevant based on the prioritization model. Any articles with unclear PECO relevance at the title-and-abstract screening phase will be included in the full text review.

Full-Text Review

After completion of title-and-abstract screening, full-text articles will be retrieved for those studies that either clearly meet the inclusion criteria or where eligibility to meet the inclusion criteria is unclear. Two members of the evaluation team will independently conduct a full-text screen of each reference identified as relevant in the title-and-abstract step to determine whether a reference meets the inclusion criteria. Screening conflicts will be resolved by a senior member of the evaluation team.

Reporting the flow of information and eligibility: inclusion and exclusion

Reason for exclusion at the full-text review stage will be annotated and reported in an interactive reference flow diagram (I-REFF) (Walker et al. 2024) in the final report [using reporting practices outlined in Moher et al. (2009) and Page et al. (2021)]. Although more than one reason may apply, for simplicity, only one of the following reasons for exclusion will be documented according to the following hierarchy: (1) is a relevant review, editorial, or other publication with no original data; (2) is relevant but not published in English; (3) lacks a relevant exposure; (4) lacks a relevant outcome; (5) not relevant to population; (6) lacks a relevant comparator (e.g., a control group); (7) other (e.g., abstracts, commentaries).

Data Extraction and Content Management

Data will be extracted by two members of the evaluation team per individual study with one member serving as the primary data extractor followed by the second team member conducting a review for confirmation (a quality control step). Data extraction will be managed with structured forms and stored in a database format using Health Assessment Workspace Collaborative (HAWC), a free, open-source, and web-based software application[2]. Data extraction elements are listed in Appendix 3. Data will be extracted only for health effects that are relevant to the PECO statement; however, other health effects may be identified for potential future use and/or use by other groups. Study information collected during data extraction will be visualized, when appropriate (e.g., when there are data on the same health effects evaluated across multiple studies) and made publicly available upon publication of the finalized report.

The extracted data will be used to summarize study designs and findings. The elements of the data extraction may be revised following the identification of the studies included in the review. DTT staff will

[2] Health Assessment Workspace Collaborative (HAWC): A Modular Web-based Interface to Facilitate Development of Human Health Assessments of Chemicals, https://hawcproject.org/portal/.

oversee the data extraction process. Extractors will be trained with an initial pilot phase of at least two example studies undertaken to improve clarity of extraction instructions and to improve consistency among extractors. DTT staff will supervise pilot testing, and adjustments to extraction instructions will occur as directed by DTT. Any discrepancies in data extraction will be resolved by discussion or consultation with a third member of the evaluation team.

Quality Assessment of Individual Studies

Risk of bias (RoB) will be assessed for individual studies using the OHAT RoB tool that takes a parallel approach for evaluating RoB from human and animal studies in order to facilitate consideration of RoB across evidence streams with common terms and categories (NTP 2015; 2019). The RoB tool consists of a set of 11 questions that are answered based on the specific details of individual studies to develop RoB ratings (using the four options in Table 3) for each question. Study design determines the subset of questions used to assess RoB for an individual study (see Table 4). RoB will be assessed at the outcome level because study design or method specifics might increase the RoB for some outcomes and not for others within the same study.

DTT staff will oversee the RoB-assessment process and supervise the pilot testing. Adjustments to the RoB criteria will occur as directed by DTT. If questions arise during the pilot testing phase or subsequent RoB assessment on the appropriateness of a specific method (e.g., outcome assessment, statistical analyses), NIEHS subject matter experts in the relevant area will be contacted. For all studies, two reviewers will independently conduct RoB evaluations and reach consensus on disagreements by discussion and consultation with senior evaluation team members or NIEHS subject matter experts as needed. Assessors will be trained using the criteria from the OHAT RoB tool with an initial pilot phase of 3-5 studies of various study designs undertaken to improve clarity of criteria that distinguish between adjacent ratings and to improve consistency among assessors. All team members involved in the RoB assessment will be trained on the same set of studies and asked to identify potential ambiguities in the criteria used to assign ratings for each question. Any ambiguities and rating conflicts will be discussed relative to opportunities to refine the criteria to distinguish more clearly between adjacent ratings. If major changes to the risk-of-bias criteria are made based on the pilot phase (i.e., those that would likely result in revision of response), they will be documented in a protocol amendment along with the date of modifications and the logic for the changes. Information about confounding, exposure characterization, outcome assessment, and other important issues might be identified during or after data extraction, which can lead to further refinement of the RoB criteria (Sterne et al. 2014). After assessors have independently made risk-of-bias determinations for a study across all risk-of-bias questions, the two assessors will compare their results to identify discrepancies and attempt to resolve them. Any remaining discrepancies will be considered by a senior member of the evaluation team and, if needed, other members of the evaluation design team. The final risk-of-bias rating for each question will be recorded with a statement of its basis.

Table 3. Response Options for Each Risk-of-bias Question	
++	***Definitely Low*** *risk of bias*: Direct evidence of low risk of bias practices
+	***Probably Low*** *risk of bias*: Indirect evidence of low risk of bias practices **OR** deviations from low risk of bias practices for these criteria during the study are deemed not to bias results appreciably, including consideration of direction and magnitude of bias
– NR	***Probably High*** *risk of bias*: Indirect evidence of high risk of bias practices **OR** insufficient information (e.g., not reported or "NR") is provided about relevant risk of bias practices
––	***Definitely High*** *risk of bias*: Direct evidence of high risk of bias practices

Table 4. OHAT Risk-of-bias Questions and Applicability by Study Design

Risk-of-bias Questions	Experimental Animal[a]	Human Controlled Trials[b]	Cohort[c]	Case-control[d]	Cross-sectional[e]	Case Report/Case Series[f]
1. Was administered dose or exposure level adequately randomized?	X	X				
2. Was allocation to study groups adequately concealed?	X	X				
3. Did selection of study participants result in the appropriate comparison groups?			X	X	X	
4. Did study design or analysis account for important confounding and modifying variables?			X	X	X	X
5. Were experimental conditions identical across study groups?	X					
6. Were research personnel blinded to the study group during the study?	X	X				
7. Were outcome data complete without attrition or exclusion from analysis?	X	X	X	X	X	
8. Can we be confident in the exposure characterization?	X	X	X	X	X	X
9. Can we be confident in the outcome assessment (including blinding of outcome assessors)?	X	X	X	X	X	X
10. Were all measured outcomes reported?	X	X	X	X	X	X
11. Were there no other potential threats to internal validity?	X	X	X	X	X	X

[a]Experimental animal studies are controlled exposure studies. Non-human animal observational studies can be evaluated using the design features of observational human studies such as cross-sectional study design.

[b]Human controlled trials are studies in humans with controlled exposure (e.g., randomized controlled trials, non-randomized experimental studies).

[c]Cohort studies are observational studies in humans that examine a cohort prospectively or retrospectively over time. Although cohort studies may include longitudinal analyses, it is not a prerequisite of the cohort study design.

[d]Case-control studies are observational studies in humans that compare exposures of individuals who have a specific health effect or disease with exposures of controls who do not have the health effect or disease. Controls generally come from the same population from which the cases were derived.

[e]Cross-sectional studies are observational studies in humans that examine the relationship between exposures and outcomes or health effects assessed contemporaneously. Cross-sectional studies include population surveys with individual data (e.g., NHANES) and surveys with aggregate data (i.e., ecological studies).

[f]A case report (or case study) is a descriptive study of a single individual or small group in which the study of an association between an observed effect and a specific environmental exposure is based on clinical evaluations and histories of the individual(s). A case series study in environmental epidemiology is designed to share health-related events on a collection of case reports on subjects with the same or similar health outcome(s) and environmental exposure(s).

Exposure Characterization

Exposure characterization is a key factor in determining study quality when assessing glyphosate hazards. For animal studies, exposure characterization requires consideration of two overarching elements: (1) independent identification and characterization of the test material and exposure conditions (e.g., preparation in the dosing vehicle), and (2) exposure quantification sufficient to confirm expected dose differences across exposure groups, including route and dose.

1. Accurate identification of the test material is essential, as inaccuracies in labeling are not uncommon. From historical experience, approximately 3% of commercially purchased chemicals by the National Toxicology Program were mislabeled with respect to chemical identity; that increased to approximately 10% when purity was considered (unpublished personal communication, Brad Collins, National Toxicology Program chemist). Independent verification of chemical identity and purity is considered best practice, especially for studies of pure glyphosate. Therefore, RoB criteria distinguish between studies evaluating pure glyphosate and those evaluating glyphosate formulations, which differ in composition and exposure characterization needs.

    a. For pure glyphosate studies, those that do not report the source or purity of the test material will be considered “probably high RoB” for exposure characterization. However, if purity information is not reported but can be reasonably inferred from the source (e.g., manufacturer specifications or publicly available product descriptions), a “probably low RoB” rating may be assigned. Studies may also be considered “probably low RoB” if source and/or purity information is not provided, but measured glyphosate concentrations in biological samples (e.g., blood or serum) demonstrate a clear and consistent dose-response gradient across exposure groups.

    b. For glyphosate formulation studies, characterization focuses on the identity of the formulation and the form and concentration of glyphosate present, rather than purity of glyphosate alone. Relevant information includes the complete product name, glyphosate salt form(s), active ingredient concentration, purchase location (e.g., country), and manufacturing date (month and year), which together allow conversion of administered doses to glyphosate acid equivalents and facilitate comparison across studies.

    Characterization of the test material includes the form of glyphosate (acid or specific salt forms), whether the exposure involves pure glyphosate or a glyphosate formulation, the purity (for pure glyphosate studies), methods used to prepare dosing solutions, and the stability of glyphosate in stock solutions, and/or dosing-ready dilutions under the experimental conditions. Glyphosate in herbicide formulations may be present as the acid or various salt forms (e.g., isopropylamine, ammonium, potassium, diammonium, dimethylamine, trimesium) (USEPA 2025). Because glyphosate acid is the biologically active form, exposure levels ideally expressed as glyphosate acid equivalents or molar concentration to facilitate comparisons across studies (Mann and Bidwell 1999; Smith et al. 2019). Differences in salt-to-acid weight ratios mean that identical percentages of active ingredient by weight across different formulations can correspond to different glyphosate acid-equivalent doses.

    Information on active ingredient form and concentration, required for pesticide product registration, combined with a complete product name, purchase location, and manufactured time, allows conversion of treatment concentrations to glyphosate acid-

equivalent doses and accounts for formulation changes over time (Travlos et al. 2017). Glyphosate formulations also contain so-called "inert" ingredients (e.g., surfactants or fungicides) that may have biological activity.

2. Exposure quantification must be sufficient to demonstrate consistent exposure administration and clear differentiation between exposure groups This includes information on exposure timing, concentration, frequency, route. Although some exposure routes (e.g., intravenous injection) are not representative of human exposure, they may still inform toxicity because circulating glyphosate is equivalent to absorbed, unmetabolized glyphosate from other exposure routes.

   For dietary or drinking water studies, reporting of consumption data and/or internal dose metrics is necessary to confirm expected exposure levels. For inhalation, aspiration, or intratracheal instillation studies, detailed methods for administration (e.g., aerosol generation, vaporization, methods, particle size distributions) and/or internal dose metrics are needed to demonstrate exposure differentiation. Where available, internal dose measurements strengthen exposure characterization, particularly when most measurements are above the assay limit of quantitation, assay performance is verified (e.g., using spiked samples or dilution curves), and analytical methods are described or referenced. Finally, exposure must occur during a biologically relevant window for the outcome assessed (e.g., spanning the full spermatogenic cycle for spermatogenesis outcomes).

Timing of Exposure and Outcome Measurements

The timing and duration of exposure and the timing of outcome assessment are important considerations when evaluating the certainty of study results in supporting a link between exposure and outcome. To appropriately address these issues, some aspects of timing are considered under indirectness and applicability in step 5 of the OHAT approach (e.g., whether or not the results address the outcome of interest) and some are considered under RoB (e.g., whether or not the timing of outcome assessment was consistent across all treatment groups). See OHAT RoB tool for detailed discussion of time-related study features relevant to indirectness and RoB.

The following questions on time-related variables in study design and conduct are addressed under indirectness in rating confidence in the body of evidence (Step 5 of the OHAT Approach), not in the risk-of-bias assessment:

- Was the exposure in the appropriate biological window to affect the outcome?
- Was the outcome assessed at an adequate amount of time after the exposure for the development of the outcome?

Other Key Risk-of-Bias Assessment Domains

Other key RoB assessment domains beyond exposure characterization that are considered key include: randomization to treatment group, confidence in outcome assessment (including blinding of the outcome assessors), and controlling for litter effects in developmental studies (NTP 2015; 2019).

A lack of randomization to dose administration or exposure level can bias results away from the null toward larger effect sizes. This effect has been empirically assessed in experimental animals [reviewed in Krauth et al. (2013)]. This element is widely recommended to assess RoB for controlled human trials (Guyatt et al. 2011d; Higgins et al. 2011; IOM 2011; Viswanathan et al. 2012) and is included in most RoB instruments for animal studies [reviewed in Hooijmans et al. (2014) and Krauth et al. (2013)].

A lack of blinding at outcome assessment and larger measures of effect has been reported for experimental animal studies (Bebarta et al. 2003; Sena et al. 2007; Vesterinen et al. 2010). Research specifically evaluating the impact of lack of blinding during allocation to treatment groups or during the course of the study in experimental animal studies is absent or minimal (NTP 2015). In addition, concealment of animal dose information is problematic if exposure results in obvious effects on the normal daily functioning of the animal. Additional steps can be taken in experimental animal studies to reduce the RoB such as counterbalancing critical factors (e.g., sex, observers, apparatus, session, necropsy order) to equally distribute each factor across dose groups for endpoint assessments. Concern for lack of blinding during allocation or the conduct of the study can be attenuated if blinding was implemented at outcome assessment. For these reasons, blinding at outcome assessment is weighed more heavily during RoB assessment than blinding during allocation concealment or during the course of the study.

When developmental exposure regimens are used for experimental animal studies, the preferred analytical design is to consider the litter as the experimental unit for statistical analysis. Failure to adjust for litter, either statistically or experimentally, is an important RoB in an animal study with developmental exposure or for a developmental endpoint. Animals generated from the same litter tend to respond more similarly than animals from different litters. The direction of the bias is away from the null toward a larger effect size (Haseman et al. 2001) if the individual pup is considered as the statistical unit rather than the dam or litter from which the pup is derived. This can be due to inflation of the sample size or biological influence of the dam and litter. For non-developmental exposure studies, the statistical analysis is based on the individual and litter of origin is not considered.

Missing Information for Risk-of-Bias Assessment

The evaluation team will attempt to contact authors of included studies by email to obtain missing information considered critical for evaluating RoB that cannot be inferred from the study. If additional information or data are received from study authors, risk-of-bias judgments will be modified to reflect the updated study information. If the evaluation team does not receive a response from the authors by one month of the contact attempt, a RoB response of “NR” for “not reported; probably high RoB” will be used and a note made in the data extraction files that an attempt to contact the authors was unsuccessful. Note that we may have to limit response time with authors to 2 weeks rather than the preferred one month due to the limited timeframe of this review.

Assessment of Confidence in the Body of Evidence

The quality of evidence for each outcome will be evaluated using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system for rating the confidence in the body of evidence (Guyatt et al. 2011a; Guyatt et al. 2011b; Guyatt et al. 2011c; Guyatt et al. 2011d; NTP 2015; 2019). To approach evidence assessment, the framework provides guidance for determining overall certainty in the evidence as “high,” “moderate,” “low,” or “very low” based on five factors for downgrading (e.g., RoB across studies, indirectness, unexplained inconsistency, imprecision, publication bias) and five for upgrading (e.g., large magnitude of the effect, dose response, residual confounding, consistency of findings across studies, and other)(Figure 1). More detailed guidance is available in the OHAT Handbook for conducting systematic review (NTP 2015; 2019).

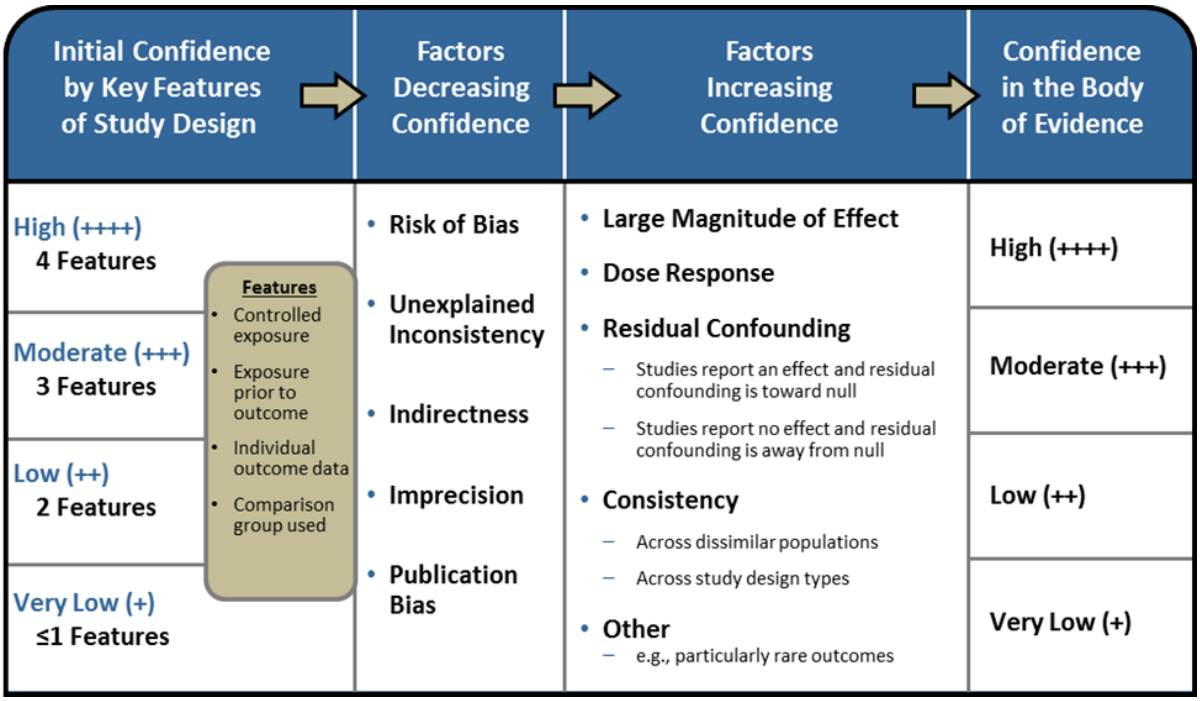

Figure 1. Assessing Confidence in the Body of Evidence

Evidence Synthesis

Considerations for Pursuing a Narrative or Quantitative Evidence Synthesis

Heterogeneity within the available evidence will determine the type of evidence integration that is appropriate: either a quantitative synthesis (meta-analysis) or narrative approach for evidence integration. Where appropriate, a meta-analysis will be conducted to summarize the findings. Summaries of main study design characteristics for each included study will be compiled to determine comparability between studies, identify data transformations necessary to ensure comparability, and determine whether study heterogeneity is a concern. Including a study might not be appropriate when (1) data on exposure or outcome are too different to be combined, (2) concerns about high RoB are present, (3) endpoints or measurement scales are not sufficiently similar, or (4) other circumstances indicate that averaging study results would not produce meaningful results. Topic-specific experts will be consulted to help assess whether studies are too heterogeneous for meta-analysis to be appropriate. When quantitatively combining results is inappropriate or infeasible, findings will be narratively described or visually presented.

The main characteristics considered when determining whether to combine studies quantitatively include the following for human studies and animal studies are presented in Table 5.

Table 5. Factors for Consideration when Combining Data from Animal Studies
Population: Animal model used (species, strain, sex, genetic background)
Population: Age of animals (at start of treatment and outcome assessment)
Exposure: Dose levels, frequency of treatment, timing of exposure, duration, and exposure route
Outcome: Type of reproductive health outcome assessed in males and timing of outcome assessment
Outcome: Type of data (e.g., continuous, dichotomous), statistics presented in paper, ability to access raw or additional data
Risk of bias: Variation in degree of RoB at individual study level or very serious concern for RoB across studies

Confidence Assessment for Glyphosate or Glyphosate Formulations

- Data extraction and summary of the results
  - Results will be presented by glyphosate exposure (e.g., pure glyphosate, glyphosate metabolites, and glyphosate formulations).
  - Results will also be presented for specific glyphosate chemical form (e.g., salt from formulation) and may be further stratified by specific formulation as deemed necessary.
  - Although results for chemicals with few studies or little data will be available in tables, the HAWC database, or supplemental files, results on each chemical may be featured in the systematic review.
- Quality of evidence and corresponding level-of-evidence conclusion rating
  - Consistency of evidence will be evaluated for glyphosate and glyphosate formulations including, but not limited to, consistency in the endpoints, direction of effect, and mechanisms of effect where data are available to support evaluations.
  - Summaries will indicate where data are sufficient to evaluate the consistency of different chemical forms of glyphosate or different glyphosate formulations (e.g., formulations all have animal studies in more than one species showing the same or similar effects on male reproductive health-related endpoints); where data are available to evaluate consistency with some uncertainty (e.g., glyphosate or glyphosate formulations with consistent findings in only animal species for the same or similar effects); where data are available to evaluate consistency with high uncertainty (e.g., two studies with inconsistent findings on similar endpoints); or data are insufficient (no data or data on different endpoints across the body of evidence).

Future Steps

The primary deliverable, the government report, will include only confidence ratings for the evidence for an association between exposure to glyphosate or glyphosate formulations and male reproductive health outcomes based on the critical evaluation of the available animal data described in this protocol. The report will not include hazard conclusions because studies with human or mechanistic data are outside of the current scope. Therefore, the protocol does not outline procedures for developing level of evidence or hazard conclusions at this time. After the report is developed, DTT scientists plan to consider the broader evidence base in an update and expansion of the systematic review for peer review and potential journal publication. The update is likely to include and assessment of available human and mechanistic data to increase the human relevance of the review. If the evidence supports the development of level of evidence and hazard conclusions, that evaluation will follow the remaining

steps (Preparation of Level of Evidence Statement and Integrate Evidence to Develop Hazard Identification Conclusions) described in the OHAT Handbook. The specific methods for identifying and critically assessing (i.e., risk of bias) human and mechanistic studies protocol tailored to the assessment of glyphosate and glyphosate formulations would be outlined. Similarly, level of evidence, evidence integration, and developing hazard conclusions will also be described with any updates provided as an addendum to this protocol.

SYSTEMATIC REVIEW OUTLINE

The DTT Systematic Review of Exposure to Exposure to Glyphosate and Glyphosate Formulations and Male Reproductive Health will include the following information:

Introduction

This section will provide a brief background on the topic.

Objective and Specific Aims

This section will identify the objective and goals of the systematic review.

Methods

This section will provide a brief overview of the methodologies used in the review process, including:

- problem formulation and protocol development,
- PECO statement,
- literature search,
- study selection,
- quality assessment of included studies,
- data extraction methods,
- assessment in confidence in the bodies of evidence.

Results

This section will include the results from the systematic review. Results will be presented in tables or figures as appropriate using Tableau®. The results from the studies included will be discussed by outcome. This will include a description of:

- Literature search results (description and interactive reference flow diagram),
- Quality of the studies
- Data extraction and summary of the results by class (i.e., groups or subgroups) of chemicals
- Quality of evidence rating (confidence in the bodies of evidence).

Discussion

This section will include a discussion of results of the systematic review, including evaluation of whether the data support consistency of effects from glyphosate exposure alone, glyphosate formulations (to the extent data are available), limitations of the database, and the limitations of the systematic review including potential limitations of any methods adaptations to facilitate a rapid approach.

REFERENCES

Agency for Toxic Substances and Disease Registry (ATSDR). 2020. Toxicological profile for glyphosate. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry. [Accessed: December 31, 2025]. https://www.atsdr.cdc.gov/toxprofiles/tp214.pdf

Bebarta V, Luyten D, Heard K. 2003. Emergency medicine animal research: Does use of randomization and blinding affect the results? Acad Emerg Med. 10(6):684-687. https://doi.org/10.1111/j.1553-2712.2003.tb00056.x

Boocock MR, Coggins JR. 1983. Kinetics of 5-enolpyruvylshikimate-3-phosphate synthase inhibition by glyphosate. FEBS Lett. 154(1):127-133. https://doi.org/10.1016/0014-5793(83)80888-6

Cai W, Ji Y, Song X, Guo H, Han L, Zhang F, Liu X, Zhang H, Zhu B, Xu M. 2017. Effects of glyphosate exposure on sperm concentration in rodents: A systematic review and meta-analysis. Environ Toxicol Pharmacol. 55:148-155. https://doi.org/10.1016/j.etap.2017.07.015

Camiccia M, Candiotto LZP, Gaboardi SC, Panis C, Kottiwitz LBM. 2022. Determination of glyphosate in breast milk of lactating women in a rural area from Paraná state, Brazil. Braz J Med Biol Res. 55:e12194. https://doi.org/10.1590/1414-431X2022e12194

Connolly A, Coggins MA, Galea KS, Jones K, Kenny L, McGowan P, Basinas I. 2019. Evaluating glyphosate exposure routes and their contribution to total body burden: A study among amenity horticulturalists. Ann Work Expo Health. 63(2):133-147. https://doi.org/10.1093/annweh/wxy104

European Food Safety Authority (EFSA), Álvarez F, Arena M, Auteri D, Binaglia M, Castoldi AF, Chiusolo A, Crivellente F, Egsmose M, Fait G, et al. 2023. Peer review of the pesticide risk assessment of the active substance glyphosate. EFSA J. 21(7):e08164. https://doi.org/10.2903/j.efsa.2023.8164

Evidence Partners. 2021. DistillerSR version 2.35. Ottawa, Ontario, Canada: Evidence Partners.

Franz JE, inventor; Monsanto Company, assignee. 1983 Sep 20. Salts of N-phosphonomethylglycine. United States patent US 4,405,531.

Funke T, Han H, Healy-Fried ML, Fischer M, Schönbrunn E. 2006. Molecular basis for the herbicide resistance of Roundup Ready crops. Proc Natl Acad Sci U S A. 103(35):13010-13015. https://doi.org/10.1073/pnas.0603638103

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, deBeer H, et al. 2011a. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 64(4):383-394. https://doi.org/10.1016/j.jclinepi.2010.04.026

Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G, et al. 2011b. GRADE guidelines 6. Rating the quality of evidence—imprecision. J Clin Epidemiol. 64(12):1283-1293. https://doi.org/10.1016/j.jclinepi.2011.01.012

Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, et al. 2011c. GRADE guidelines: 5. Rating the quality of evidence—publication bias. J Clin Epidemiol. 64(12):1277-1282. https://doi.org/10.1016/j.jclinepi.2011.01.011

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, et al. 2011d. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). J Clin Epidemiol. 64(4):407-415. https://doi.org/10.1016/j.jclinepi.2010.07.017

Hamel C, Kelly SE, Thavorn K, Rice DB, Wells GA, Hutton B. 2020. An evaluation of DistillerSR's machine learning-based prioritization tool for title/abstract screening – Impact on reviewer-relevant outcomes. BMC Med Res Methodol. 20(1):256. https://doi.org/10.1186/s12874-020-01129-1

Hartzler B. 2001. Glyphosate - A review. Ames, IA: Iowa State University, Extension and Outreach. [Accessed: December 29, 2025]. https://crops.extension.iastate.edu/encyclopedia/glyphosate-review

Haseman JK, Bailer AJ, Kodell RL, Morris R, Portier K. 2001. Statistical issues in the analysis of low-dose endocrine disruptor data. Toxicol Sci. 61(2):201-210. https://doi.org/10.1093/toxsci/61.2.201

Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JA, et al. 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 343:d5928. https://doi.org/10.1136/bmj.d5928

Hooijmans CR, Rovers MM, de Vries RBM, Leenaars M, Ritskes-Hoitinga M, Langendam MW. 2014. SYRCLE's risk of bias tool for animal studies. BMC Med Res Methodol. 14:43. https://doi.org/10.1186/1471-2288-14-43

Institute of Medicine (IOM). 2011. Finding what works in health care: Standards for systematic reviews. Washington, DC: The National Academies Press. https://doi.org/10.17226/13059

Krauth D, Woodruff TJ, Bero L. 2013. Instruments for assessing risk of bias and other methodological criteria of published animal studies: A systematic review. Environ Health Perspect. 121(9):985-992.

Li ZM, Jeong H, Kannan K. 2024. Widespread occurrence of glyphosate and aminomethylphosphonic acid in indoor dust from urban homes across the United States and its contribution to human exposure. Environ Int. 192:109005. https://doi.org/10.1016/j.envint.2024.109005

Li ZM, Kannan K. 2025. Occurrence and distribution of glyphosate and aminomethylphosphonic acid in surface water, drinking water, wastewater and other water types from New York State, USA. Water Res. 286:124232. https://doi.org/10.1016/j.watres.2025.124232

MAHA Commission. 2025. The MAHA report: Make our Children Healthy Again. Washington, DC: The White House. [Accessed: December 29, 2025]. https://www.whitehouse.gov/wp-content/uploads/2025/05/MAHA-Report-The-White-House.pdf

Malik J, Barry G, Kishore G. 1989. The herbicide glyphosate. BioFactors. 2(1):17-25.

Mann RM, Bidwell JR. 1999. The toxicity of glyphosate and several glyphosate formulations to four species of southwestern Australian frogs. Arch Environ Contam Toxicol. 36(2):193-199. https://doi.org/10.1007/s002449900460

Medalie L, Baker NT, Shoda ME, Stone WW, Meyer MT, Stets EG, Wilson M. 2020. Influence of land use and region on glyphosate and aminomethylphosphonic acid in streams in the USA. Sci Total Environ. 707:136008. https://doi.org/10.1016/j.scitotenv.2019.136008

Mellor E, Trasande L, Albergamo V, Kannan K, Li Z, Ghassabian A, Afanasyeva Y, Liu M, Cowell W. 2024. Sociodemographic and dietary determinants of glyphosate exposure in a NYC-based pregnancy cohort. Environ Pollut. 363(Pt 1):125083. https://doi.org/10.1016/j.envpol.2024.125083

Mohammadi K, Sani MA, Safaei P, Rahmani J, Molaee-Aghaee E, Jafari SM. 2022. A systematic review and meta-analysis of the impacts of glyphosate on the reproductive hormones. Environ Sci Pollut Res Int. 29(41):62030-62041. https://doi.org/10.1007/s11356-021-16145-x

Mohammadi Z, Rahaie M, Moradifar F. 2025. A novel approach for colorimetric detection of glyphosate in food based on a split aptamer nanostructure and DNAzyme activity. J Fluoresc. 35(8):6515-6527. https://doi.org/10.1007/s10895-024-03998-x

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moschini G, Perry E, Hennessy DA. 2019. Patent expiration, product concentration, and glyphosate use: A tale of unexpected consequences. Ames, IA: Iowa State University, Center for Agricultural and Rural Development. [Accessed: December 29, 2025]. https://agpolicyreview.card.iastate.edu/spring-2019/patent-expiration-product-concentration-and-glyphosate-use-tale-unexpected-consequences

National Toxicology Program (NTP). 2015. OHAT risk of bias rating tool for human and animal studies. Research Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service, National Institute of Environmental Health Sciences, Division of the National Toxicology Program, Office of Health Assessment and Translation. [Accessed: December 29, 2025]. https://ntp.niehs.nih.gov/sites/default/files/ntp/ohat/pubs/riskofbiastool_508.pdf

National Toxicology Program (NTP). 2019. Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration. Research Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service, National Institute of Environmental Health Sciences, Division of the National Toxicology Program, Office of Health Assessment and Translation. [Accessed: December 29, 2025]. https://ntp.niehs.nih.gov/sites/default/files/ntp/ohat/pubs/handbookmarch2019_508.pdf

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, et al. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 372:n71. https://doi.org/10.1136/bmj.n71

Pierce JS, Roberts B, Kougias DG, Comerford CE, Riordan AS, Keeton KA, Reamer HA, Jacobs NFB, Lotter JT. 2020. Pilot study evaluating inhalation and dermal glyphosate exposure resulting from simulated heavy residential consumer application of Roundup®. Inhal Toxicol. 32(8):354-367. https://doi.org/10.1080/08958378.2020.1814457

Rogers SG, Brand LA, Holder SB, Sharps ES, Brackin MJ. 1983. Amplification of the aroA gene from Escherichia coli results in tolerance to the herbicide glyphosate. Appl Environ Microbiol. 46(1):37-43. https://doi.org/10.1128/aem.46.1.37-43.1983

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014. Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect. 122(7):711-718.

Sena E, van der Worp HB, Howells D, Macleod M. 2007. How can we improve the pre-clinical development of drugs for stroke? Trends Neurosci. 30(9):433-439. https://doi.org/10.1016/j.tins.2007.06.009

Smith CM, Vera MKM, Bhandari RK. 2019. Developmental and epigenetic effects of Roundup and glyphosate exposure on Japanese medaka (Oryzias latipes). Aquat Toxicol. 210:215-226. https://doi.org/10.1016/j.aquatox.2019.03.005

Sterne JAC, Higgins JPT, Reeves BC, on behalf of the development group for ACROBAT-NRSI. 2014. A Cochrane risk of bias assessment tool: For non-randomized studies of interventions (ACROBAT-NRSI), version 1.0.0. [Accessed: January 6, 2026]. https://www.bristol.ac.uk/population-health-sciences/centres/beam-centre/barr/riskofbias/robins-i/acrobat-nrsi/

Sun X, Zhang H, Huang X, Yang D, Wu C, Liu H, Zhang L. 2024. Associations of glyphosate exposure and serum sex steroid hormones among 6–19-year-old children and adolescents. Ecotoxicol Environ Saf. 275:116266. https://doi.org/10.1016/j.ecoenv.2024.116266

Travlos I, Cheimona N, Bilalis D. 2017. Glyphosate efficacy of different salt formulations and adjuvant additives on various weeds. Agronomy. 7(3):60. https://doi.org/10.3390/agronomy7030060

U.S. Environmental Protection Agency (USEPA). 2017. [Memorandum re: Transmission of meeting minutes and final report of the December 13-16, 2016 FIFRA SAP meeting held to consider and review scientific issues associated with EPA's evaluation of the carcinogenic potential of glyphosate]. Washington, DC: U.S. Environmental Protection Agency. [Accessed: December 31, 2025]. https://www.epa.gov/sites/default/files/2017-03/documents/december_13-16_2016_final_report_03162017.pdf

U.S. Environmental Protection Agency (USEPA). 2025. Active Pesticide Project Registration Informational Listing (APPRIL). Washington, DC: U.S. Environmental Protection Agency, Office of Pesticide Programs. [Accessed: December 29, 2025]. https://ordspub.epa.gov/ords/pesticides/f?p=APPRIL_PUBLIC:2

Vesterinen HM, Sena ES, ffrench-Constant C, Williams A, Chandran S, Macleod MR. 2010. Improving the translational hit of experimental treatments in multiple sclerosis. Mult Scler. 16(9):1044-1055. https://doi.org/10.1177/1352458510379612

Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters M, Santaguida PL, Shamliyan T, Singh K, Tsertsvadze A, et al. 2012. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 12-EHC047-EF. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Healthcare Research and Quality. [Accessed: December 29, 2025]. https://effectivehealthcare.ahrq.gov/products/methods-guidance-bias-individual-studies/methods

Walker VR, Lemeris CR, Magnuson K, Sibrizzi CA, Shipkowski KA, Taylor KW, Rooney AA. 2024. I-REFF diagrams: Enhancing transparency in systematic review through interactive reference flow diagrams. Syst Rev. 13(1):33. https://doi.org/10.1186/s13643-023-02420-0

Wu WY, Wang DS, Lin HC, Wang C, Lin CY. 2025. Urinary glyphosate concentrations and serum sex hormones in a nationally representative U.S. sample: NHANES 2017-2018. Life (Basel). 15(7):1024. https://doi.org/10.3390/life15071024

ABOUT THE PROTOCOL

Contributors

Evaluation teams are composed of federal staff and contractor staff. Contractor staff members are screened for potential conflicts of interest. Federal staff members should do a self-evaluation. Epidemiologists and toxicologists on IHAB evaluation teams should have at least three years' experience and/or training in reviewing studies, including summarizing studies and critical review (e.g., assessing study quality and interpreting findings). Team members should have at least a master's degree or equivalent experience in epidemiology, toxicology, environmental health sciences, or a related field.

Federal Staff

Name	Affiliation, Principal Role
Brandiese E.J. Beverly	NIEHS/DTT, Internal Review Lead
Kembra L. Howdeshell	NIEHS/DTT, Project Co-Lead
Andrew A. Rooney	NIEHS/DTT
Kelly Shipkowski	NIEHS/DTT, Contracting Officer's Representative (COR)
Anisha Singh	NIEHS/DTT
Kyla W. Taylor	NIEHS/DTT, Project Co-Lead
Vickie Walker	NIEHS/DTT, Search, Extraction, and Workflow Lead
Amy Wang	NIEHS/DTT, Exposure Lead

Contract Staff

ICF contract staff to NIEHS/DTT will work on multiple aspects of the evaluation in partnership with the DTT team under the direction of the COR and technical guidance of the DTT evaluation team. The list of ICF team members provided below is a potential list likely to be added to for the systematic review. Similarly, a list of tasks for the ICF staff includes but is not limited to assisting in literature search strategy development, screening, data extraction, and data visualization.

Name	Affiliation
Robyn Blain	ICF
Dave Burch	ICF
Meredith Clemons	ICF
Anastasia Freedman	ICF
Ryan Gan	ICF
Anthony Hannani	ICF
Pamela Hartman	ICF
Kate Helmick	ICF
Jeanne Luh	ICF
Kristen Magnuson	ICF

Name	Affiliation
Andrew Maresca	ICF
Danielle Moore	ICF
Lisa Prince	ICF
Courtney Rosenthal	ICF
Amanda Ross	ICF
Christopher Sibrizzi	ICF
Hannah Starnes	ICF
Nicole Vetter	ICF
Jessica Wignall	ICF
Angelina Winnett	ICF
Farha Zindah	ICF

NIEHS Protocol Reviewers

The protocol reviewers were selected for their ongoing work or experience with public health, toxicology, systematic review methods, glyphosate, and male reproductive toxicology.

Name	Affiliation
Abee Boyles, Ph.D.	NIEHS/Division of Extramural Research and Training (DERT)
Chad Blystone, Ph.D., DABT	NIEHS/DTT
Stephanie Smith-Roe, Ph.D.	NIEHS/DTT

Protocol History and Revisions

Date	Activity or Revision
January 1, 2026	Draft protocol for review by NIEHS subject matter experts
February 17, 2026	Revised protocol finalized based on reviewer comments

Sources of Support

This research will be funded by the Intramural Research Program (ES103379-01) at the National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health, and will be performed for NIEHS under contract 75N96025C00003. This work will be funded, in part, by the Division of Intramural Research at the NIH, National Institute of Environmental Health Sciences (Z1AES103325 (CLJ)).

APPENDICES

Appendix 1. Catalog of Concurrent Similar Systematic Reviews

We conducted a search to identify whether there were any similar systematic evidence maps, systematic reviews, or meta-analyses on the topic of exposure to glyphosate and male reproductive outcomes. The search strategy for Epistemonikos, the largest source of systematic reviews relevant for health decision-making (https://www.epistemonikos.org), is presented in Table 1-1 and the results are presented in Table 1-2. We similarly conducted a search of PROSPERO, an international prospective register of systematic reviews, to observe whether there were any ongoing systematic reviews and/or meta-analyses evaluating exposure to exposure to glyphosate and male reproductive outcomes (https://www.crd.york.ac.uk/prospero/). The PROSPERO search strategy is presented in Table 1-3 and the search results are presented in Table 1-4. We identified published systematic review in the Epistemonikos and all were identified with the literature search strategy for the systematic review; two of the systematic reviews were relevant with literature updated to year 2020 (Table 1-2). Eight protocols were identified in PROSPERO and none were directly relevant to the current systematic review effort (Table 1-4).

Table 1-1. Epistemonikos Database Search Terms

Set	Search Strategy
Exposures	"glyphosate" OR "1071-83-6" OR "(Carboxymethylamino)methylphosphonic acid" OR "Carboxymethylaminomethanephosphinic acid" OR "C-K Yuyos FAV" OR "CP 67573" OR "Folusen" OR "Forsat" OR "Glialka" OR "Glifoglex" OR "Glifosan 747" OR "gliphosate" OR "Gliz" OR "Glyfos" OR "GlyGran" OR "Glyphodin A" OR "Glyphomax" OR "Glyphosate" OR "Glyphosphate" OR "Ground Bio" OR "Herbatop" OR "HM 2028" OR "Kickdown" OR "Lancer herbicide" OR "MON 2139" OR "MON 3539" OR "MON 6000" OR "N-(Phosphonomethyl)glycine" OR "N-(phosphonomethyl)-Glycine" OR "N-Phosphomethylglycine" OR "NPhosphonomethylglycine" OR "Phorsat" OR "Phosphonomethylglycine" OR "Phosphonomethyliminoacetic acid" OR "Pondmaster" OR "Rebel Garden" OR "Roundup Max" OR "Safal" OR "Scout herbicide" OR "Silglif" OR "yerbimat" OR "Roundup" OR "34494-03-6" OR "MON 0459" OR "40465-66-5" OR "MON 14420" OR "MON 8750" OR "Roundup Hi-Load" OR "Roundup PRODry" OR "70393-85-0" OR "MON 8000" OR "Monsanto 8000" OR "Polado" OR "Trisodium hydrogen bis(N-(phosphonatomethyl)aminoacetate" OR "39600-42-5" OR "Glyphosate potassium" OR "Glyphosate monopotassium salt" OR "Glyphosate potassium" OR "Glyphosate-potassium" OR "Monopotassium glyphosate" OR "Roundup Attack" OR "Roundup Energy" OR "Roundup Maxload" OR "Roundup Original Max" OR "Roundup Power Max" OR "Roundup Ultramax II" OR "Roundup Weathermax" OR "Touchdown Forte HiTech" OR "Transorb R" OR "Weathermax" OR "Zapp Qi" OR "70901-12-1" OR "Glyphosate-potassium" OR "Potassium glyphosate" OR "Potassium N-(phosphonomethyl)glycine" OR "Uragan Forte" OR "VisionMAX" OR "N-(phosphonomethyl)glycine potassium salt" OR "114370-14-8" OR "Glyphosate ammonium" OR "N-(phosphonomethyl)glycine ammonium salt" OR "69254-40-6" OR "Glyphosate-diammonium" OR "Diammonium N-(phosphonomethyl)glycine" OR "N-(phosphonomethyl)glycine diammonium salt")) OR (("glyphosate, isopropyl amine salt" OR "N-(phosphonomethyl)glycine trimethylsulfonium salt" OR "38641-94-0" OR "Glyphosate-isopropylammonium" OR "Glyphosate isopropylamine salt" OR "Azural AT" OR "CP 70139" OR "Fosulen" OR "Glifosato estrella" OR "Glycel"

Table 1-1. Epistemonikos Database Search Terms	
Set	**Search Strategy**
	OR "Glycine, N-(phosphonomethyl)-, cmpd with 2- propanamine (1:1)" OR "Glyfos AU" OR "Glyfos BIO" OR "Glyphosate isopropylamine salt" OR "Glyphosate mono(isopropylamine) salt" OR "Glyphosateisopropylammonium" OR "Glyphosate-mono(isopropylammonium)" OR "Landmaster" OR "MON 139" OR "MON 39" OR "N-(Phosphonomethyl)glycine isopropylamine salt" OR "N-(Phosphonomethyl)glycine isopropylammonium salt" OR "N-(Phosphonomethyl)glycine monoisopropylamine salt" OR "Nitosorg" OR "Rondo" OR "Utal" OR "Utal (herbicide)" OR "Vision (herbicide)" OR "2- Propanamine, compd, with N-(phosphonomethyl)glycine (1:1)" OR "Glycine, N-(phosphonomethyl)-, compd. with 2-propanamine (1:1)" OR "N-(Phosphonomethyl)glycine, compound with 2-propylamine (1:1)" OR "Isopropylamine glyphosate" OR "81591-81- 3" OR "Glyphosate-trimesium" OR "Glyphosphate-trimesium" OR "Avans 330" OR "Glyphosate mono(trimethylsulfonium) salt" OR "Glyphosate trimethylsulfonium salt" OR "Glyphosate-trimesium" OR "Medallon" OR "Ouragan" OR "R 50224" OR "SC 0224" OR "Sulfosate" OR "Sulphosate" OR "Touchdown herbicide" OR "Trimethylsulfonium carboxymethylamino-methylphosphonate" OR "Trimethylsulfonium glyphosate" OR "Glycine, N-(phosphonomethyl)-, ion(1-), trimethylsulfonium" OR "Sulfosate"
Male reproductive health outcomes	("Reproduction" OR reproducti* OR sperm* OR "spermatozoa" OR "Testis" OR testicular OR seminiferous OR seminal OR "seminal vesical" OR epididymis OR sertoli OR hormone OR prostate OR testosterone OR "Blood-Testis Barrier" OR mullerian OR "Luteinizing Hormone" OR LH OR estrogen OR progesterone OR "follicle-stimulating hormone" OR FSH OR "ano-genital distance" OR "AGD" OR "testis decent" OR "preputial separation" OR gonad* OR scrot* OR semin* OR spermat* OR testes OR testi*)
Reviews	Systematic Reviews

Table 1-2. Epistemonikos Database Search Results of Systematic Reviews of Glyphosate and Male Reproductive Health	
Reference and Epistemonikos number	**Comparison to current review**
Cai W, Ji Y, Song X, Guo H, Han L, Zhang F, Liu X, Zhang H, Zhu B, Xu M. 2017. Effects of glyphosate exposure on sperm concentration in rodents: A systematic review and meta-analysis. Environmental toxicology and pharmacology. 55:148-155. 10.1016/j.etap.2017.07.015. af6c43f88057ce72bb305688bb9a02fca1491c79.	Similar scope, but lacks systematic review procedures to minimize bias and enhance transparency (e.g., no protocol, no risk of bias assessment, etc.); includes literature through November 2016
Cardona Maya WD. 2020. Glyphosate Negatively Affects Human Sperm: In vitro evidence. Urol Colomb. 29(2):96-98. 10.1055/s-0039-1696699. 6f8b090b7ae155029bbc2ec1995d36a1741be32f.	Population not relevant (includes in vitro studies only)
de Araujo JS, Delgado IF, Paumgartten FJ. 2016. Glyphosate and adverse pregnancy outcomes, a systematic review of observational studies. BMC public health. 16(1):472. 10.1186/s12889-016-3153-3. 80cd19efad6b81a4ef3ecacf97fd7489b83d0dfc	Outcomes not relevant review (includes pregnancy outcomes only)

Table 1-2. Epistemonikos Database Search Results of Systematic Reviews of Glyphosate and Male Reproductive Health	
Reference and Epistemonikos number	**Comparison to current review**
Kaboli Kafshgiri S, Farkhondeh T, Miri-Moghaddam E. 2021. Glyphosate effects on the female reproductive systems: a systematic review. Reviews on environmental health. 10.1515/reveh-2021-0029. 616c9fcc9cd6f0db4c07fdf943131b12f065c802	Population not relevant (includes females only)
Mink PJ, Mandel JS, Lundin JI, Sceurman BK. 2011. Epidemiologic studies of glyphosate and non-cancer health outcomes: a review. Regulatory toxicology and pharmacology. 61(2):172-184. 10.1016/j.yrtph.2011.07.006. 93575f19e4dd9f2b0aa17d57d23c2506e2e8dc55	Population not relevant (includes human evidence only)
Mohammadi K, Sani MA, Safaei P, Rahmani J, Molaee-Aghaee E, Jafari SM. 2021. A systematic review and meta-analysis of the impacts of glyphosate on the reproductive hormones. Environmental science and pollution research international. 10.1007/s11356-021-16145-x	Similar scope, but lacks systematic review procedures to minimize bias and enhance transparency (e.g., no protocol, no risk of bias assessment, etc.); includes literature through November 2020
Nagy K, Duca RC, Lovas S, Creta M, Scheepers PTJ, Godderis L, Ádám B. 2020. Systematic review of comparative studies assessing the toxicity of pesticide active ingredients and their product formulations. Environmental research. 181:108926. 10.1016/j.envres.2019.108926. a0dcf034e36a21cd9634d1abba37af66b2f612e0	Population and Exposure not relevant (Not specific to glyphosate and includes both animal and in vitro evidence streams)
Rana I, Nguyen PK, Rigutto G, Louie A, Lee J, Smith MT, Zhang L. 2023. Mapping the key characteristics of carcinogens for glyphosate and its formulations: A systematic review. Chemosphere.139572. 10.1016/j.chemosphere.2023.139572. 253972e47d117ec5acab0f875e19fda0416295b0	Outcome not relevant (focused on carcinogenesis)
Tosi S, Sfeir C, Carnesecchi E, van Engelsdorp D, Chauzat MP. 2022. Lethal, sublethal, and combined effects of pesticides on bees: A meta-analysis and new risk assessment tools. The Science of the total environment.156857. 10.1016/j.scitotenv.2022.156857. 96e2330ce4bacd481371b1e3759e97c5db13fb85	Population not relevant (non-mammalian (bees) only)
Tyler C, Pullin AS, Stewart GB. 2006. Effectiveness of management interventions to control invasion by Rhododendron ponticum. Environmental management. 37(4):513-522. 10.1007/s00267-005-0127-0. fee8ddb8425622e537d2f4e357b678bc40702c43	Outcome not relevant (focused on weed control)
Zilnik G, Bergeron PE, Chuang A, Diepenbrock L, Hanel A, Middleton E, Moretti E, Schmidt-Jeffris R. 2023. Meta-Analysis of Herbicide Non-Target Effects on Pest Natural Enemies. Insects. 14(10). 10.3390/insects14100787. 7bd550f8af86f73158c09f7032d862e8b7ac54a7.	Population not relevant (insects and glyphosate's non-targeted effects on them)

Date Searched: 1/12/2026
Limits: None
Publication Year: No date limit

Table 1-3. Prospero Search Terms	
Set	**Search Strategy**
Exposures	"glyphosate" OR "1071-83-6" OR "(Carboxymethylamino)methylphosphonic acid" OR "Carboxymethylaminomethanephosphinic acid" OR "C-K Yuyos FAV" OR "CP 67573" OR "Folusen" OR "Forsat" OR "Glialka" OR "Glifoglex" OR "Glifosan 747" OR "gliphosate" OR "Gliz" OR "Glyfos" OR "GlyGran" OR "Glyphodin A" OR "Glyphomax" OR "Glyphosate" OR "Glyphosphate" OR "Ground Bio" OR "Herbatop" OR "HM 2028" OR "Kickdown" OR "Lancer herbicide" OR "MON 2139" OR "MON 3539" OR "MON 6000" OR "N-(Phosphonomethyl)glycine" OR "N-(phosphonomethyl)-Glycine" OR "N-Phosphomethylglycine" OR "NPhosphonomethylglycine" OR "Phorsat" OR "Phosphonomethylglycine" OR "Phosphonomethyliminoacetic acid" OR "Pondmaster" OR "Rebel Garden" OR "Roundup Max" OR "Safal" OR "Scout herbicide" OR "Silglif" OR "yerbimat" OR "Roundup" OR "34494-03-6" OR "MON 0459" OR "40465-66-5" OR "MON 14420" OR "MON 8750" OR "Roundup Hi-Load" OR "Roundup PRODry" OR "70393-85-0" OR "MON 8000" OR "Monsanto 8000" OR "Polado" OR "Trisodium hydrogen bis(N-(phosphonatomethyl)aminoacetate" OR "39600-42-5" OR "Glyphosate potassium" OR "Glyphosate monopotassium salt" OR "Glyphosate potassium" OR "Glyphosate-potassium" OR "Monopotassium glyphosate" OR "Roundup Attack" OR "Roundup Energy" OR "Roundup Maxload" OR "Roundup Original Max" OR "Roundup Power Max" OR "Roundup Ultramax II" OR "Roundup Weathermax" OR "Touchdown Forte HiTech" OR "Transorb R" OR "Weathermax" OR "Zapp Qi" OR "70901-12-1" OR "Glyphosate-potassium" OR "Potassium glyphosate" OR "Potassium N-(phosphonomethyl)glycine" OR "Uragan Forte" OR "VisionMAX" OR "N-(phosphonomethyl)glycine potassium salt" OR "114370-14-8" OR "Glyphosate ammonium" OR "N-(phosphonomethyl)glycine ammonium salt" OR "69254-40-6" OR "Glyphosate-diammonium" OR "Diammonium N-(phosphonomethyl)glycine" OR "N-(phosphonomethyl)glycine diammonium salt")) OR (("glyphosate, isopropyl amine salt" OR "N-(phosphonomethyl)glycine trimethylsulfonium salt" OR "38641-94-0" OR "Glyphosate-isopropylammonium" OR "Glyphosate isopropylamine salt" OR "Azural AT" OR "CP 70139" OR "Fosulen" OR "Glifosato estrella" OR "Glycel" OR "Glycine, N-(phosphonomethyl)-, cmpd with 2- propanamine (1:1)" OR "Glyfos AU" OR "Glyfos BIO" OR "Glyphosate isopropylamine salt" OR "Glyphosate mono(isopropylamine) salt" OR "Glyphosateisopropylammonium" OR "Glyphosate-mono(isopropylammonium)" OR "Landmaster" OR "MON 139" OR "MON 39" OR "N-(Phosphonomethyl)glycine isopropylamine salt" OR "N-(Phosphonomethyl)glycine isopropylammonium salt" OR "N-(Phosphonomethyl)glycine monoisopropylamine salt" OR "Nitosorg" OR "Rondo" OR "Utal" OR "Utal (herbicide)" OR "Vision (herbicide)" OR "2- Propanamine, compd, with N-(phosphonomethyl)glycine (1:1)" OR "Glycine, N-(phosphonomethyl)-, compd. with 2-propanamine (1:1)" OR "N-(Phosphonomethyl)glycine, compound with 2-propylamine (1:1)" OR "Isopropylamine glyphosate" OR "81591-81- 3" OR "Glyphosate-trimesium" OR "Glyphosphate-trimesium" OR "Avans 330" OR "Glyphosate mono(trimethylsulfonium) salt" OR "Glyphosate trimethylsulfonium salt" OR "Glyphosate-trimesium" OR "Medallon" OR "Ouragan" OR "R 50224" OR "SC 0224" OR "Sulfosate" OR "Sulphosate" OR "Touchdown herbicide" OR "Trimethylsulfonium carboxymethylamino-methylphosphonate" OR "Trimethylsulfonium glyphosate" OR "Glycine, N-(phosphonomethyl)-, ion(1-), trimethylsulfonium" OR "Sulfosate"

Table 1-3. Prospero Search Terms	
Set	**Search Strategy**
Male reproductive health outcomes	("Reproduction" OR reproducti* OR sperm* OR "spermatozoa" OR "Testis" OR testicular OR seminiferous OR seminal OR "seminal vesical" OR epididymis OR sertoli OR hormone OR prostate OR testosterone OR "Blood-Testis Barrier" OR mullerian OR "Luteinizing Hormone" OR LH OR estrogen OR progesterone OR "follicle-stimulating hormone" OR FSH OR "ano-genital distance" OR "AGD" OR "testis decent" OR "preputial separation" OR gonad* OR scrot* OR semin* OR spermat* OR testes OR testi*)

Filters: None

Table 1-4. PROSPERO Database Search Results of Systematic Reviews of Glyphosate and Male Reproductive Health	
Reference and PROSPERO number	**Comparison to current review**
Anaïs G, Cédric V, Inge H. Systematic review on diet-related endocrine disrupting chemicals in relation to cancer risk. CRD42016042277.	Outcome not relevant
Audencio V, Ana Raquel Manuel G, Sancho Pedro X, Melsequisete Daniel V, Nelson Domingos C, Shelsia Victor R, Vanilda Alves M, Patrícia HCR, Alexandre Dias Porto Chiavegatto F, Akuze J, et al. Deep Learning approaches for COVID-19 Detection on Chest X-Ray Imaging: A Systematic review. CRD42025642875.	Exposure and Outcome not relevant
Joanna B, Cinthia de Fátima M, Manuela C. A systematic review of agricultural pesticides exposure factors for adverse pregnancy outcomes. CRD42021234375.	Outcome not relevant
Juan Pablo Alzate G, Orlando S, Daniel R. Effects of glyphosate on human health: a systematic review of the literature. CRD42019128975.	Population not relevant (human evidence only)
Leena P, Nishtha M, Priyanka W. Endocrine-Disrupting Chemical Excretion through Sweat: A Systematic Review. CRD42024592618.	Outcome not relevant
Maria Laura Marconi F, Milena M, Isabella N, Fernando A, José Mendes A, Adriana de Azevedo P, Liania Alves L, Patrícia Helen de Carvalho R, Patrícia Borges Botelho G, Guilherme da Silva F, et al. Vitamin D Supplementation and Cognitive Performance in Women Over 45: A Systematic Review. CRD42025631689.	Exposure and Outcome not relevant
Saanu S, Hitha Veeramani N. Health Impacts of Household Pesticides: A Systematic Review of Short-term and Long-term Effects in Humans. CRD42024588062.	Population not relevant (human evidence only)
Sarah Barbosa Castelo B, Susana Isabel S, Sandra Carla L, Pedro Filipe P. Pesticide exposure during pregnancy and the risk of adverse maternal and fetal outcomes: a systematic review. CRD420251172677.	Population (human evidence only) and Outcome not relevant

https://www.crd.york.ac.uk/prospero
Date Searched: 1/12/26
Publication Year: No date limit
Results: 8

Appendix 2. Literature Search Strategy

The strategy for this search is based on the combined search strategies in the 2020 ATSDR Toxicological Profile of Glyphosate. The exposure terms are designed to capture, glyphosate, glyphosate metabolites, and glyphosate formulations including formulation names. The outcome terms are designed to capture male reproductive effects or reproductive toxicity studies. Only PubMed, Web of Science, and Scopus databases will be searched with separate syntax for each bibliographic database.

Table 2-1. PubMed Database Search Terms

Set	Search Strategy
Exposure	(("glyphosate"[nm] OR "1071-83-6"[tw] OR "(Carboxymethylamino)methylphosphonic acid"[tw] OR "Carboxymethylaminomethanephosphinic acid"[tw] OR "C-K Yuyos FAV"[tw] OR "CP 67573"[tw] OR "Folusen"[tw] OR "Forsat"[tw] OR "Glialka"[tw] OR "Glifoglex"[tw] OR "Glifosan 747"[tw] OR "gliphosate"[tw] OR "Gliz"[tw] OR "Glyfos"[tw] OR "GlyGran"[tw] OR "Glyphodin A"[tw] OR "Glyphomax"[tw] OR "Glyphosate"[tw] OR "Glyphosphate"[tw] OR "Ground Bio"[tw] OR "Herbatop"[tw] OR "HM 2028"[tw] OR "Kickdown"[tw] OR "Lancer herbicide"[tw] OR "MON 2139"[tw] OR "MON 3539"[tw] OR "MON 6000"[tw] OR "N-(Phosphonomethyl)glycine"[tw] OR "N-(phosphonomethyl)-Glycine"[tw] OR "N-Phosphomethylglycine"[tw] OR "NPhosphonomethylglycine"[tw] OR "Phorsat"[tw] OR "Phosphonomethylglycine"[tw] OR "Phosphonomethyliminoacetic acid"[tw] OR "Pondmaster"[tw] OR "Rebel Garden"[tw] OR "Roundup Max"[tw] OR "Safal"[tw] OR "Scout herbicide"[tw] OR "Silglif"[tw] OR "yerbimat"[tw] OR "Roundup"[tw] OR "34494-03-6"[tw] OR "MON 0459"[tw] OR "40465-66-5"[tw] OR "MON 14420"[tw] OR "MON 8750"[tw] OR "Roundup Hi-Load"[tw] OR "Roundup PRODry"[tw] OR "70393-85-0"[tw] OR "MON 8000"[tw] OR "Monsanto 8000"[tw] OR "Polado"[tw] OR "Trisodium hydrogen bis(N-(phosphonatomethyl)aminoacetate"[tw] OR "39600-42-5"[tw] OR "Glyphosate potassium"[tw] OR "Glyphosate monopotassium salt"[tw] OR "Glyphosate potassium"[tw] OR "Glyphosate-potassium"[tw] OR "Monopotassium glyphosate"[tw] OR "Roundup Attack"[tw] OR "Roundup Energy"[tw] OR "Roundup Maxload"[tw] OR "Roundup Original Max"[tw] OR "Roundup Power Max"[tw] OR "Roundup Ultramax II"[tw] OR "Roundup Weathermax"[tw] OR "Touchdown Forte HiTech"[tw] OR "Transorb R"[tw] OR "Weathermax"[tw] OR "Zapp Qi"[tw] OR "70901-12-1"[tw] OR "Glyphosate-potassium"[tw] OR "Potassium glyphosate"[tw] OR "Potassium N-(phosphonomethyl)glycine"[tw] OR "Uragan Forte"[tw] OR "VisionMAX"[tw] OR "N-(phosphonomethyl)glycine potassium salt"[tw] OR "114370-14-8"[tw] OR "Glyphosate ammonium"[tw] OR "N-(phosphonomethyl)glycine ammonium salt"[tw] OR "69254-40-6"[tw] OR "Glyphosate-diammonium"[tw] OR "Diammonium N-(phosphonomethyl)glycine"[tw] OR "N-(phosphonomethyl)glycine diammonium salt"[tw])) OR (("glyphosate, isopropyl amine salt"[nm] OR "N-(phosphonomethyl)glycine trimethylsulfonium salt"[nm] OR "38641-94-0"[tw] OR "Glyphosate-isopropylammonium"[tw] OR "Glyphosate isopropylamine salt"[tw] OR "Azural AT"[tw] OR "CP 70139"[tw] OR "Fosulen"[tw] OR "Glifosato estrella"[tw] OR "Glycel"[tw] OR "Glycine, N-(phosphonomethyl)-, cmpd with 2- propanamine (1:1)"[tw] OR "Glyfos AU"[tw] OR "Glyfos BIO"[tw] OR "Glyphosate isopropylamine salt"[tw] OR "Glyphosate mono(isopropylamine) salt"[tw] OR "Glyphosateisopropylammonium"[tw] OR "Glyphosate-mono(isopropylammonium)"[tw] OR "Landmaster"[tw] OR "MON 139"[tw] OR "MON 39"[tw] OR "N-(Phosphonomethyl)glycine isopropylamine salt"[tw] OR "N-(Phosphonomethyl)glycine isopropylammonium salt"[tw] OR "N-(Phosphonomethyl)glycine monoisopropylamine salt"[tw] OR "Nitosorg"[tw] OR "Rondo"[tw] OR "Utal"[tw] OR "Utal (herbicide)"[tw] OR "Vision (herbicide)"[tw] OR "2- Propanamine, compd, with N-(phosphonomethyl)glycine (1:1)"[tw] OR "Glycine, N- (phosphonomethyl)-, compd. with 2-propanamine (1:1)"[tw] OR "N-(Phosphonomethyl)glycine, compound with 2-propylamine (1:1)"[tw] OR "Isopropylamine glyphosate"[tw] OR "81591-81- 3"[tw] OR "Glyphosate-trimesium"[tw] OR "Glyphosphate-trimesium"[tw] OR "Avans 330"[tw] OR "Glyphosate mono(trimethylsulfonium) salt"[tw] OR "Glyphosate trimethylsulfonium salt"[tw] OR "Glyphosate-trimesium"[tw] OR "Medallon"[tw] OR "Ouragan"[tw] OR "R 50224"[tw] OR "SC 0224"[tw] OR "Sulfosate"[tw] OR "Sulphosate"[tw] OR "Touchdown herbicide"[tw] OR "Trimethylsulfonium carboxymethylamino-methylphosphonate"[tw] OR "Trimethylsulfonium glyphosate"[tw] OR "Glycine, N-(phosphonomethyl)-, ion(1-), trimethylsulfonium"[tw] OR "Sulfosate"[tw]))

Table 2-1. PubMed Database Search Terms	
Set	**Search Strategy**
Outcome	"Reproduction"[Mesh] OR reproducti*[tiab] OR sperm*[tiab] OR "spermatozoa"[Mesh] OR "Testis"[Mesh] OR testicular[tiab] OR seminiferous[tiab] OR seminal[tiab] OR “seminal vesical”[tiab] OR epididymis[tiab] OR sertoli[tiab] OR hormone[tiab] OR prostate[tiab] OR testosterone[tiab] OR "Blood-Testis Barrier"[Mesh] OR mullerian[tiab] OR "Luteinizing Hormone"[Mesh] OR LH OR estrogen[tiab] OR progesterone[tiab] OR “follicle-stimulating hormone”[tiab] OR FSH[tiab] OR “ano-genital distance”[tiab] OR “AGD”[tiab] OR “testis decent”[tiab] OR “preputial separation”[tiab] OR gonad*[tiab] OR scrot*[tiab] OR semin*[tiab] OR spermat*[tiab] OR testes[tiab] OR testi*[tiab]

Table 2-2. Web of Science Database Search Terms	
Set	**Search Strategy**
Exposure	TOPIC ("glyphosate" OR "1071-83-6" OR "(Carboxymethylamino)methylphosphonic acid" OR "Carboxymethylaminomethanephosphinic acid" OR "C-K Yuyos FAV" OR "CP 67573" OR "Folusen" OR "Forsat" OR "Glialka" OR "Glifoglex" OR "Glifosan 747" OR "gliphosate" OR "Gliz" OR "Glyfos" OR "GlyGran" OR "Glyphodin A" OR "Glyphomax" OR "Glyphosate" OR "Glyphosphate" OR "Ground Bio" OR "Herbatop" OR "HM 2028" OR "Kickdown" OR "Lancer herbicide" OR "MON 2139" OR "MON 3539" OR "MON 6000" OR "N-(Phosphonomethyl)glycine" OR "N-(phosphonomethyl)-Glycine" OR "N-Phosphomethylglycine" OR "NPhosphonomethylglycine" OR "Phorsat" OR "Phosphonomethylglycine" OR "Phosphonomethyliminoacetic acid" OR "Pondmaster" OR "Rebel Garden" OR "Roundup Max" OR "Safal" OR "Scout herbicide" OR "Silglif" OR "yerbimat" OR "Roundup" OR "34494-03-6" OR "MON 0459" OR "40465-66-5" OR "MON 14420" OR "MON 8750" OR "Roundup Hi-Load" OR "Roundup PRODry" OR "70393-85-0" OR "MON 8000" OR "Monsanto 8000" OR "Polado" OR "Trisodium hydrogen bis(N-(phosphonatomethyl)aminoacetate" OR "39600-42-5" OR "Glyphosate potassium" OR "Glyphosate monopotassium salt" OR "Glyphosate potassium" OR "Glyphosate-potassium" OR "Monopotassium glyphosate" OR "Roundup Attack" OR "Roundup Energy" OR "Roundup Maxload" OR "Roundup Original Max" OR "Roundup Power Max" OR "Roundup Ultramax II" OR "Roundup Weathermax" OR "Touchdown Forte HiTech" OR "Transorb R" OR "Weathermax" OR "Zapp Qi" OR "70901-12-1" OR "Glyphosate-potassium" OR "Potassium glyphosate" OR "Potassium N-(phosphonomethyl)glycine" OR "Uragan Forte" OR "VisionMAX" OR "N-(phosphonomethyl)glycine potassium salt" OR "114370-14-8" OR "Glyphosate ammonium" OR "N-(phosphonomethyl)glycine ammonium salt" OR "69254-40-6" OR "Glyphosate-diammonium" OR "Diammonium N-(phosphonomethyl)glycine" OR "N-(phosphonomethyl)glycine diammonium salt")) OR (("glyphosate, isopropyl amine salt" OR "N-(phosphonomethyl)glycine trimethylsulfonium salt" OR "38641-94-0" OR "Glyphosate-isopropylammonium" OR "Glyphosate isopropylamine salt" OR "Azural AT" OR "CP 70139" OR "Fosulen" OR "Glifosato estrella" OR "Glycel" OR "Glycine, N-(phosphonomethyl)-, cmpd with 2-propanamine (1:1)" OR "Glyfos AU" OR "Glyfos BIO" OR "Glyphosate isopropylamine salt" OR "Glyphosate mono(isopropylamine) salt" OR ("Glyphosateisopropylammonium" OR "Glyphosate-mono(isopropylammonium)" OR "Landmaster" OR "MON 139" OR "MON 39" OR "N-(Phosphonomethyl)glycine isopropylamine salt" OR "N-(Phosphonomethyl)glycine isopropylammonium salt" OR "N-(Phosphonomethyl)glycine monoisopropylamine salt" OR "Nitosorg" OR "Rondo" OR "Utal" OR "Utal (herbicide)" OR "Vision (herbicide)" OR "2- Propanamine, compd, with N-(phosphonomethyl)glycine (1:1)" OR "Glycine, N- (phosphonomethyl)-, compd. with 2-propanamine (1:1)" OR "N-(Phosphonomethyl)glycine, compound with 2-propylamine (1:1)" OR "Isopropylamine glyphosate" OR "81591-81- 3" OR "Glyphosate-trimesium" OR "Glyphosphate-trimesium" OR "Avans 330" OR "Glyphosate mono(trimethylsulfonium) salt" OR "Glyphosate trimethylsulfonium salt" OR "Glyphosate-trimesium" OR "Medallon" OR "Ouragan" OR "R 50224" OR "SC 0224" OR "Sulfosate" OR "Sulphosate" OR "Touchdown herbicide" OR "Trimethylsulfonium carboxymethylamino-methylphosphonate" OR "Trimethylsulfonium glyphosate" OR "Glycine, N-(phosphonomethyl)-, ion(1-), trimethylsulfonium" OR "Sulfosate")
Outcome	TOPIC ("Reproduction” OR reproducti* OR sperm* OR "spermatozoa" OR "Testis" OR testicular OR seminiferous OR seminal OR “seminal vesical” OR epididymis OR sertoli OR hormone OR prostate OR testosterone OR "Blood-Testis Barrier" OR mullerian OR "Luteinizing Hormone" OR LH OR estrogen OR

Table 2-2. Web of Science Database Search Terms	
Set	**Search Strategy**
	progesterone OR “follicle-stimulating hormone” OR FSH OR “ano-genital distance” OR “AGD” OR “testis decent” OR “preputial separation” OR gonad* OR scrot* OR semin* OR spermat* OR testes OR testi*)

Table 2-3. Scopus Database Search Terms	
Set	**Search Strategy**
Exposure	TITLE-ABS ("glyphosate" OR "1071-83-6" OR "(Carboxymethylamino)methylphosphonic acid" OR "Carboxymethylaminomethanephosphinic acid" OR "C-K Yuyos FAV" OR "CP 67573" OR "Folusen" OR "Forsat" OR "Glialka" OR "Glifoglex" OR "Glifosan 747" OR "gliphosate" OR "Gliz" OR "Glyfos" OR "GlyGran" OR "Glyphodin A" OR "Glyphomax" OR "Glyphosate" OR "Glyphosphate" OR "Ground Bio" OR "Herbatop" OR "HM 2028" OR "Kickdown" OR "Lancer herbicide" OR "MON 2139" OR "MON 3539" OR "MON 6000" OR "N-(Phosphonomethyl)glycine" OR "N-(phosphonomethyl)-Glycine" OR "N-Phosphomethylglycine" OR "NPhosphonomethylglycine" OR "Phorsat" OR "Phosphonomethylglycine" OR "Phosphonomethyliminoacetic acid" OR "Pondmaster" OR "Rebel Garden" OR "Roundup Max" OR "Safal" OR "Scout herbicide" OR "Silglif" OR "yerbimat" OR "Roundup" OR "34494-03-6" OR "MON 0459" OR "40465-66-5" OR "MON 14420" OR "MON 8750" OR "Roundup Hi-Load" OR "Roundup PRODry" OR "70393-85-0" OR "MON 8000" OR "Monsanto 8000" OR "Polado" OR "Trisodium hydrogen bis(N-(phosphonatomethyl)aminoacetate" OR "39600-42-5" OR "Glyphosate potassium" OR "Glyphosate monopotassium salt" OR "Glyphosate potassium" OR "Glyphosate-potassium" OR "Monopotassium glyphosate" OR "Roundup Attack" OR "Roundup Energy" OR "Roundup Maxload" OR "Roundup Original Max" OR "Roundup Power Max" OR "Roundup Ultramax II" OR "Roundup Weathermax" OR "Touchdown Forte HiTech" OR "Transorb R" OR "Weathermax" OR "Zapp Qi" OR "70901-12-1" OR "Glyphosate-potassium" OR "Potassium glyphosate" OR "Potassium N-(phosphonomethyl)glycine" OR "Uragan Forte" OR "VisionMAX" OR "N-(phosphonomethyl)glycine potassium salt" OR "114370-14-8" OR "Glyphosate ammonium" OR "N-(phosphonomethyl)glycine ammonium salt" OR "69254-40-6" OR "Glyphosate-diammonium" OR "Diammonium N-(phosphonomethyl)glycine" OR "N-(phosphonomethyl)glycine diammonium salt")) OR (("glyphosate, isopropyl amine salt" OR "N-(phosphonomethyl)glycine trimethylsulfonium salt" OR "38641-94-0" OR "Glyphosate-isopropylammonium" OR "Glyphosate isopropylamine salt" OR "Azural AT" OR "CP 70139" OR "Fosulen" OR "Glifosato estrella" OR "Glycel" OR "Glycine, N-(phosphonomethyl)-, cmpd with 2-propanamine (1:1)" OR "Glyfos AU" OR "Glyfos BIO" OR "Glyphosate isopropylamine salt" OR "Glyphosate mono(isopropylamine) salt" OR "Glyphosateisopropylammonium" OR "Glyphosate-mono(isopropylammonium)" OR "Landmaster" OR "MON 139" OR "MON 39" OR "N-(Phosphonomethyl)glycine isopropylamine salt" OR "N-(Phosphonomethyl)glycine isopropylammonium salt" OR "N-(Phosphonomethyl)glycine monoisopropylamine salt" OR "Nitosorg" OR "Rondo" OR "Utal" OR "Utal (herbicide)" OR "Vision (herbicide)" OR "2- Propanamine, compd, with N-(phosphonomethyl)glycine (1:1)" OR "Glycine, N- (phosphonomethyl)-, compd. with 2-propanamine (1:1)" OR "N-(Phosphonomethyl)glycine, compound with 2-propylamine (1:1)" OR "Isopropylamine glyphosate" OR "81591-81- 3" OR "Glyphosate-trimesium" OR "Glyphosphate-trimesium" OR "Avans 330" OR "Glyphosate mono(trimethylsulfonium) salt" OR "Glyphosate trimethylsulfonium salt" OR "Glyphosate-trimesium" OR "Medallon" OR "Ouragan" OR "R 50224" OR "SC 0224" OR "Sulfosate" OR "Sulphosate" OR "Touchdown herbicide" OR "Trimethylsulfonium carboxymethylamino-methylphosphonate" OR "Trimethylsulfonium glyphosate" OR "Glycine, N-(phosphonomethyl)-, ion(1-), trimethylsulfonium" OR "Sulfosate")
Outcome	TITLE-ABS ("Reproduction” OR reproducti* OR sperm* OR "spermatozoa" OR "Testis" OR testicular OR seminiferous OR seminal OR “seminal vesical” OR epididymis OR sertoli OR hormone OR prostate OR testosterone OR "Blood-Testis Barrier" OR mullerian OR "Luteinizing Hormone" OR LH OR estrogen OR progesterone OR “follicle-stimulating hormone” OR FSH OR “ano-genita* OR scrot* OR semin* OR spermat* OR testes OR testi*)
Limits	NOT Medline

Appendix 3. Data Extraction Elements for Animal Studies

Table 3-1. Data Extraction Elements for Animal Studies	
Funding	Funding source(s)
	Reporting of COI by authors and/or translators (*reporting bias)
Animal Model	Sex
	Species
	Strain
Treatment	Chemical name and CAS number
	Source of chemical
	Purity of chemical (*information bias)
	Dose levels or concentration (as presented and converted to mg/kg bw/d when possible)
	Other dose-related details, such as whether administered dose level was verified by measurement, information on internal dosimetry (*information bias)
	Vehicle used for exposed animals
	Route of administration (e.g., oral, inhalation, dermal, injection)
	Age or lifestage at start of dosing and at health outcome assessment
	Duration and frequency of dosing (e.g., hours, days, weeks when administration was ended, days per week)
Methods	Study design (e.g., single treatment, acute, subchronic (e.g., 90 days in a rodent), chronic, multigenerational, developmental, other)
	Guideline compliance (i.e., use of EPA, OECD, NTP or another guideline for study design, conducted under GLP guideline conditions, non-GLP but consistent with guideline study, non-guideline peer-reviewed publication)
	Number of animals per group (and dams per group in developmental studies) (*missing data bias)
	Method to control for litter effects in developmental studies (*information bias)
	Use of negative controls and whether controls were untreated, vehicle-treated, or both
	Endpoint health category (e.g., reproductive health)
	Endpoint (e.g., sperm count)
	Diagnostic or method to measure endpoint (*information bias)
	Timing of outcome assessment (e.g., gestational, day of birth, early postnatal) (*information bias)
	Statistical methods (*information bias)
Results	Measures of effect at each dose or concentration level (e.g., mean, median, frequency, measures of precision or variance) or description of qualitative results. When possible, the evaluation team will convert measures of effect to a common metric with associated 95% confidence intervals (CI). Most often, measures of effect for continuous data will be expressed as percent control response, mean difference,

Table 3-1. Data Extraction Elements for Animal Studies	
	or standardized mean difference. Categorical data will be expressed as relative risk (RR, also called risk ratio).
	No observed effect level (NOEL), lowest observed effect level (LOEL), benchmark dose (BMD) analysis, statistical significance of other dose levels, or other estimates of effect presented in paper. **Note:** The NOEL and LOEL are highly influenced by study design, give no quantitative information about the relationship between dose and response, and can be subject to author's interpretation (e.g., a statistically significant effect might not be considered biologically important). Also, a NOEL does not necessarily mean zero response. Ideally, the response rate or effect size at specific dose levels is used as the primary measure to characterize the response.
	If not presented in the study, statistical power can be assessed during data extraction using an approach that assesses the ability to detect a 10% to 20% change from control group's response for continuous data, or a relative risk or odds ratio of 1.5–2 for categorical data, using the outcome frequency in the control group to determine sample size. Recommended sample sizes to achieve 80% power for a given effect size, i.e., 10% or 20% change from control, will be compared to sample sizes used in the study to categorize statistical power. Studies will be considered adequately powered when sample size for 80% power is met.
	Observations on dose response (e.g., trend analysis, description of whether dose-response shape appears to be monotonic, nonmonotonic)
	Data on internal concentration, toxicokinetics, or toxicodynamics (when reported)
Other	Documentation of author queries, use of digital rulers to estimate data values from figures, exposure units, statistical result conversions, etc.

Appendix 4. Risk-of-bias Criteria

OHAT Risk of Bias Tool

The OHAT RoB tool for human and animal studies (version date January 2015 and available at https://ntp.niehs.nih.gov/go/38673) reflects the DTT Integrative Health Assessments Branch's current best practices and provides the detailed discussion and instructions for the RoB practices used in this evaluation. The OHAT tool uses a single set of questions (also called "elements" or "domains") to assess RoB across various study types to facilitate consideration of conceptually similar potential sources of bias across the human and animal evidence streams with a common terminology. Individual RoB questions are designated as only applicable to certain study designs (e.g., cohort studies or experimental animal studies), and a subset of the questions apply to each study design.

- The specific criteria used to assess RoB for this evaluation are outlined below for experimental animal studies.
- In general, in order to be considered as having *definitely low* or *definitely high* RoB, studies need to show direct evidence that bias was or was not addressed. Direct evidence is typically quantitative information or details about method validation. However, qualitative statements can provide indirect evidence that there was not (*probably low*) or was (*probably high*) potential for bias. When information provided is insufficient to determine potential bias and information cannot be ascertained from the authors, this is interpreted as indirect evidence of potential bias and considered "probably high RoB."

Experimental Animal Studies

1. Was administration of dose or exposure level adequately randomized?

Q1 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that animals were allocated to all study groups, including concurrent controls, using a method with a random component, **AND** there is direct evidence that the study used a concurrent control group as an indication that randomization covered all study groups, **Note:** Acceptable methods of randomization include: referring to a random number table, using a computer random number generator, coin tossing, or shuffling cards (Higgins et al. 2011). **Note:** Restricted randomization (e.g., blocked randomization) to ensure particular allocation ratios will be considered low bias. Similarly, stratified randomization approaches that attempt to minimize imbalance between groups on important prognostic factors (e.g., body weight) will be considered acceptable.

Q1 Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that animals were allocated to all study groups including concurrent controls using a method with a random component (i.e., authors state random allocation, without description of method), **AND** evidence that the study used a concurrent control group as an indication that randomization covered all study groups, **OR** it is deemed that allocation without a clearly random component would not appreciably bias results.

Q1 Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that animals were allocated to all study groups using a method with a non-random component, **OR** direct evidence that there was a lack of a concurrent control group, **OR** there is insufficient information provided about how animals were allocated to study groups (record "NR" as basis for answer).

Q1 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that animals were allocated to all study groups using a non-random method including judgment of the investigator, the results of a laboratory test, or a series of tests, **OR** direct evidence that there was a lack of a concurrent control group.

2. Was allocation to study groups adequately concealed?

Q2 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that at the time of assigning study groups the research personnel did not know what group animals were allocated to, and it is unlikely that they could have broken the blinding of allocation until after assignment was complete and irrevocable. **Note:** Acceptable methods used to ensure allocation concealment include sequentially numbered treatment containers of identical appearance or equivalent methods.

Q2 Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that at the time of assigning study groups the research personnel did not know what group animals were allocated to and it is unlikely that they could have broken the blinding of allocation until after assignment was complete and irrevocable, **OR** it is deemed that lack of adequate allocation concealment would not appreciably bias results.

Q2 Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that at the time of assigning study groups it was possible for the research personnel to know what group animals were allocated to, or it is likely that they could have broken the blinding of allocation before assignment was complete and irrevocable, **OR** there is insufficient information provided about allocation to study groups (record "NR" as basis for answer).

Q2 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that at the time of assigning study groups it was possible for the research personnel to know what group animals were allocated to, or it is likely that they could have broken the blinding of allocation before assignment was complete and irrevocable.

3. Did selection of study participants result in the appropriate comparison groups? [NA]

4. Did study design or analysis account for important confounding and modifying variables? [NA]

5. Were experimental conditions identical across study groups?

Q5 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that same vehicle was used in control and experimental animals, **AND** direct evidence that non-exposure-related experimental conditions were identical across study groups (i.e., the study explicitly states that animals were all in the same room or provides other details to indicate that the conditions were identical).

Q5 Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that the same vehicle was used in control and experimental animals, **OR** it is deemed that the vehicle used would not appreciably bias results, **AND** the authors do not explicitly state that non-exposure-related experimental conditions were identical (e.g., experimental conditions were provided, but there is no statement or demonstration that conditions were the same across study groups)

Q5 Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that the vehicle differed between control and experimental animals, **OR** authors did not report the vehicle used (record "NR" as basis for answer), **OR** indirect evidence that non-exposure-related experimental conditions were not comparable between study groups.

Q5 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence from the study report that control animals were not exposed, or administered a different vehicle than experimental animals, **OR** direct evidence that non-exposure-related experimental conditions were not comparable between study groups.

6. Were the research personnel blinded to the study group during the study?

Q6 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that the research personnel were adequately blinded to study group, and it is unlikely that they could have broken the blinding during the study.

Q6 Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that the research personnel were adequately blinded to study group, and it is unlikely that they could have broken the blinding during the study, **OR** it is deemed that lack of adequate blinding during the study would not appreciably bias results. This would include cases where blinding was not possible, but research personnel took steps to minimize potential bias, such as restricting the knowledge of study group to veterinary or supervisory personnel monitoring for overt toxicity, or randomized husbandry or handling practices (e.g., placement in the animal room, necropsy order, etc.).

Q6 Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that the research personnel were not adequately blinded to study group, **OR** there is insufficient information provided about blinding to study group during the study (record "NR" as basis for answer).

Q6 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that the research personnel were not adequately blinded to study group.

7. Were outcome data complete without attrition or exclusion from analysis?

Q7 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that loss of animals was adequately addressed and reasons were documented when animals were removed from a study. **OR** direct evidence that missing data have been imputed using appropriate methods (ensuring that characteristics of animals are not significantly different from animals retained in the analysis). **Note**: Acceptable handling of attrition includes: very little missing outcome data; reasons for missing animals unlikely to be related to outcome (or for survival data, censoring unlikely to be introducing bias); missing outcome data balanced in numbers across study groups, with similar reasons for missing data across groups; missing outcomes is not enough to impact the effect estimate.

Q7 Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that loss of animals was adequately addressed and reasons were documented when animals were removed from a study, **OR** it is deemed that the proportion lost would not appreciably bias results. This would include reports of no statistical differences in characteristics of animals removed from the study from those remaining in the study.

Q7 Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that loss of animals was unacceptably large and not adequately addressed, **OR** there is insufficient information provided about loss of animals (record "NR" as basis for answer).

Q7 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that loss of animals was unacceptably large and not adequately addressed. **Note**: Unacceptable handling of attrition or exclusion includes: reason for loss is likely to be related to true outcome, with either imbalance in numbers or reasons for loss across study groups.

8A. Can we be confident in the exposure characterization (glyphosate only studies)?

Q8A Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that exposure to glyphosate was independently characterized, and chemical purity confirmed generally as ≥98% (i.e., independently from the supplier of the chemical), (and compliance with the treatment, if applicable), **AND** exposure was consistently administered (i.e., with the same method and time frame) across exposure groups, **AND** for dietary or drinking water studies information is provided on consumption or internal dose metrics to confirm expected exposure levels sufficiently to allow discrimination between exposure groups, **AND** for inhalation, aspiration, or intratracheal instillation studies that details are provided on how the chemical is administered including vaporization methods or internal dose metrics confirm expected exposure levels sufficiently to allow discrimination between exposure groups, **AND** if internal dose metrics are available, there is direct evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished, **AND** if internal dose metrics are available, the study used spiked samples or a dilution curve to confirm assay performance, **AND** the analytical methods used to independently characterize internal measurements are described or referenced. **Note:** Exposure should cover the relevant window to impact the reproductive outcomes that are measured in the study. For example, when assessing potential effects on spermatogenesis in mouse, rat, and rabbit models, exposure should cover the full spermatogenesis window.

Q8A Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that glyphosate exposure was appropriately characterized, and purity confirmed generally as ≥98% (i.e., the supplier of the chemical provides documentation of the purity of the chemical) **OR** direct evidence that purity was independently confirmed as ≥95% and it is deemed that impurities of up to 5% would not appreciably bias results, **AND** exposure was consistently administered (i.e., with the same method and time frame) across treatment groups, **AND** for dietary or drinking water studies, no information is provided on consumption or internal dose metrics, **AND** if internal dose metrics are available, there is indirect evidence that most of the exposure data measurements are above the lowest limit of quantitation for the assay such that different exposure groups can be distinguished. **Note**: Studies without purity, stability, or consumption information can still be considered to be probably low risk of bias if there are internal measurements (e.g., blood levels of glyphosate or glyphosate metabolites) that indicate there is low concern for bias or if purity information can be reasonably inferred for the source (e.g., manufacturer specifications or publicly available product descriptions).

Q8A Experimental Animal - Probably Low Risk of bias (+)
Note: When assessing potential effects on spermatogenesis the exposure does not cover the full spermatogenesis window

Q8A Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that the glyphosate exposure (including purity of the test substance and compliance with the treatment, if applicable) was assessed using poorly validated methods, **OR** there is insufficient information provided about the validity of the exposure assessment method, but no evidence for concern (record "NR" as basis for answer), **AND** if internal dose metrics are available, there is indirect evidence that most of the exposure data measurements are below the limit of quantitation for the assay such that different exposure groups cannot be distinguished.

Q8A Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that the glyphosate exposure (including purity of the test substance and compliance with the treatment, if applicable) was assessed using poorly validated methods.

8B. Can we be confident in the exposure characterization (glyphosate formulation studies)?

Q8B Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that exposure to the glyphosate formulation was independently characterized, glyphosate formulation was provided with the complete formulation name (i.e., commercial product name in full), and concentration of glyphosate in the formulation was confirmed independently from the supplier of the formulation, (and compliance with the treatment, if applicable), **AND** exposure was consistently administered (i.e., with the same method and time frame) across exposure groups, **AND** for dietary or drinking water studies information is provided on consumption or internal dose metrics to confirm expected exposure levels sufficiently to allow discrimination between exposure groups, **AND** for inhalation, aspiration, or intratracheal instillation studies that details are provided on how the chemical is administered including vaporization methods or internal dose metrics confirm expected exposure levels sufficiently to allow discrimination between exposure groups, **AND** if internal dose metrics are available, there is direct evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished, **AND** if internal dose metrics are available, the study used spiked samples or a dilution curve to confirm assay performance, **AND** the analytical methods used to independently characterize internal measurements are described or referenced. **Note:** Exposure should cover the relevant window to impact the reproductive outcomes that are measured in the study. For example, when assessing potential effects on spermatogenesis in mouse, rat, and rabbit models, exposure should cover the full spermatogenesis window.

Q8B Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that glyphosate formulation exposure was appropriately characterized, glyphosate formulation was provided with sufficient information to calculate concentration of glyphosate acid equivalent (e.g., commercial product name in full, purchase location/country, and ideally also the purchase year), but the concentration of glyphosate was not independently confirmed, **AND** exposure was consistently administered (i.e., with the same method and time frame) across treatment groups, **AND** for dietary or drinking water studies, no information is provided on consumption or internal dose metrics, **AND** if internal dose metrics are available, there is indirect evidence that most of the exposure data measurements are above the limit of quantitation for the assay such that different exposure groups can be distinguished. **Note**: Studies without indirect information on the concentration of glyphosate may be considered probably low risk of bias if there are internal measurements (e.g., blood levels of glyphosate or glyphosate metabolites) that indicate there is low concern for bias. **Note:** When assessing potential effects on spermatogenesis the exposure does not cover the full spermatogenesis window.

Q8B Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that the glyphosate formulation exposure (including concentration of the test substance and compliance with the treatment, if applicable) was assessed using poorly validated methods, **OR** indirect evidence of exposure to a glyphosate formulation without sufficient information to calculate the concentration of glyphosate acid equivalent in the formulation, **OR** there is insufficient information provided about the validity of the exposure assessment method, but no evidence for concern (record "NR" as basis for answer), **AND** if internal dose metrics are available, there is indirect evidence that most of the exposure data measurements are below the limit of quantitation for the assay such that different exposure groups cannot be distinguished. **Note:** Studies that report exposure to a glyphosate formulation without details on what the formulation contains (e.g., glyphosate chemical form and concentration) will be considered probably high risk of bias, and the authors will be contacted for more information.

Q8B Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that the glyphosate formulation exposure (including concentration of the test substance and compliance with the treatment, if applicable) was assessed using poorly validated methods.

9. Can we be confident in the outcome assessment?

Q9 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that the outcome was assessed using well-established methods and assessed at the same length of time after initial exposure in all study groups, **AND** direct evidence that the outcome assessors were adequately blinded to the study group, and it is unlikely that they could have broken the blinding prior to reporting outcomes. **Note**: Outcomes that were assessed with a fully automated method. **Note**: Experimental design and outcome assessment should be appropriate for the reproductive outcomes that are measured in the study.

Q9 Experimental Animal - Definitely Low Risk of bias (++)

Note: The age of animals at outcome assessment should be appropriate for full reproductive maturity; assessment in younger animals (e.g., before PND 90 in rats) may introduce variability and bias results.

Note: For measures of pubertal timing (preputial gland separation (PPS)), specific step-by-step methods do not need to be stated or cited; this is considered a well-established method.

Note: For hormone measurements, the gold standard method is liquid chromatography-mass spectrometry (LC-MS). LC-MS methods should report: upper and lower limits of quantification, calibration curve, quality controls, and internal standards or recovery compounds, and samples should be run in duplicate.

Note: For measuring hormones, LC-MS is the preferred method (Faupel-Badger 2010), especially for ages or sexes with low estrogen levels.

Q9 Experimental Animal - Probably Low Risk of bias (+)

Indirect evidence that the outcome was assessed using acceptable methods (i.e., deemed valid and reliable but not the gold standard) **AND** indirect evidence that the outcome was assessed at the same length of time after initial exposure in all study groups,

OR it is deemed that the outcome assessment methods used would not appreciably bias results,

AND indirect evidence that the outcome assessors were adequately blinded to the study group, and it is unlikely that they could have broken the blinding prior to reporting outcomes,

OR it is deemed that lack of adequate blinding of outcome assessors would not appreciably bias results, which is more likely to apply to objective outcome measures.

Note: For some outcomes, particularly histopathology assessment, outcome assessors are not blind to study group as they require comparison to the control to appropriately judge the outcome, but additional measures such as multiple levels of independent review by trained pathologists can minimize potential bias.

Note: Although the age of animals at outcome assessment below full reproductive maturity (e.g., before PND 90 in rats) may introduce variability, it may be deemed that the age at assessment would not appreciably bias results.

Note: For hormone measurements, radioimmunoassay (RIA) and enzyme-linked immunosorbent assay (ELISA) are the next best method for quantifying hormones; either method is considered reliable if it provides some evidence of validating the results to either LC-MS or to a standard curve.

Note: For hormone measurements, methods should report at the lower limit of quantification, quality controls, a standard curve, and samples should be run in duplicate.

Q9 Experimental Animal - Probably High Risk of bias (-) or (NR)

Indirect evidence that the outcome assessment method is an insensitive instrument,

OR indirect evidence that the length of time after initial exposure differed by study group,

OR indirect evidence that it was possible for outcome assessors to infer the study group prior to reporting outcomes without sufficient quality control measures,

OR there is insufficient information provided about blinding of outcome assessors (record "NR" as basis for answer).

Note: For hormone measurements, the RIA or ELISA method without validation or evidence of running a standard curve may be prone to inaccuracy or lack of precision thus is rated Probably High risk-of-bias.

Q9 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that the outcome assessment method is an insensitive, or internally or externally invalid instrument, **OR** direct evidence that the length of time after initial exposure differed by study group, **OR** direct evidence for lack of adequate blinding of outcome assessors, including no blinding or incomplete blinding without quality control measures.

10. Were all measured outcomes reported?

Q10 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported. This would include outcomes reported with sufficient detail to be included in meta-analysis or fully tabulated during data extraction and analyses had been planned in advance.

Q10 Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported, **OR** indirect evidence that analyses that had not been planned in advance (i.e., retrospective unplanned subgroup analyses) are clearly indicated as such and deemed that unplanned analyses were appropriate and selective reporting would not appreciably bias results (e.g., appropriate analyses of an unexpected effect). This would include outcomes reported with insufficient detail such as only reporting that results were statistically significant (or not).

Q10 Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported, **OR** indirect evidence that unplanned analyses were included that may appreciably bias results, **OR** there is insufficient information provided about selective outcome reporting (record "NR" as answer basis).

Q10 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that all of the study's measured outcomes (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have not been reported. In addition to not reporting outcomes, this would include reporting outcomes based on composite score without individual outcome components or outcomes reported using measurements, analysis methods or subsets of the data (e.g., subscales) that were not pre-specified or reporting outcomes not pre-specified, or that unplanned analyses were included that would appreciably bias results.

11. Were there no other potential threats to internal validity?

Internal validity refers to whether the methods and modes of analysis used in a study can be interpreted to reflect a potential causal relationship between specific factor(s) and observed outcomes. It can be interpreted to mean, generally, whether a study has been done "right" so that the results are "valid" in the specific study setting and is free of bias and confounding. For this evaluation, question 11 will be used to a) examine individual studies for appropriate statistical methods and b) address whether

developmental exposure studies used litter as the unit of analysis following the recommended practices. Statistical considerations will include confirmation of homogeneity of variance for ANOVA and other statistical tests that require normally distributed data). It will also be used for RoB considerations that do not fit under the other questions.

Q11 Experimental Animal - Definitely Low Risk of bias (++)
Direct evidence that homogeneity of variance was tested for any statistical test that requires normally distributed data (e.g., t-test, ANOVA). **AND** Direct evidence that repeated measures statistical analyses were used for any experiments that repeatedly measured outcomes in the same animals, **AND** Direct evidence that the litter was considered the basic unit of analysis for any study that used littermates in an experiment.

Q11 Experimental Animal - Probably Low Risk of bias (+)
Indirect evidence that the litter was considered the basic unit of analysis for any study that used littermates in an experiment. **AND** indirect evidence that repeated measures statistical analyses were used for any experiments that repeatedly measured outcomes in the same animals.

Q11 Experimental Animal - Probably High Risk of bias (-) or (NR)
Indirect evidence that homogeneity of variance was not tested for any statistical test that requires normally distributed data (e.g., t-test, ANOVA), **OR** indirect evidence that repeated measures statistical analyses were not used for any experiments that repeatedly measured outcomes in the same animals, **OR** indirect evidence that the litter was not considered the basic unit of analysis for any study that used littermates in an experiment. **OR** there is insufficient information provided about statistical methods including if litter was used as the basic unit (record "NR" as basis for answer).

Q11 Experimental Animal - Definitely High Risk of bias (--)
Direct evidence that homogeneity of variance was not tested for any statistical test that requires normally distributed data (e.g., t-test, ANOVA), **OR** direct evidence that repeated measures statistical analyses were not used for any experiments that repeatedly measured outcomes in the same animals, **OR** direct evidence that the litter was not considered the basic unit of analysis. **Note**: Includes studies that considered each littermate an independent observation and the individual pup as the experimental unit.