National Toxicology Program
U.S. Department of Health and Human Services

## PROTOCOL FOR SCOPING REVIEW OF

# Adverse Health Effects Associated with Prenatal Exposure to Progestogens

July 31, 2020

Office of Health Assessment and Translation
Division of the National Toxicology Program
National Institute of Environmental Health Sciences
National Institutes of Health
U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Date Original Protocol Finalized: September 26, 2018
Date Revised Protocol Published: July 31, 2020
Documentation of Revisions to the Protocol: The principal revisions are detailed in the Protocol History and Revision Table on Page 10 including the reasons for each revision. Text that was deleted is crossed out and inserted text is underlined.

# TABLE OF CONTENTS



# LIST OF TABLES

# BACKGROUND AND SIGNIFICANCE

## Background

Progesterone is a steroid hormone that is involved in supporting pregnancy by inducing cellular differentiation of the uterus to support the embryo in early pregnancy and inhibiting further ovulation and myometrium contractility throughout the pregnancy. Progesterone drugs fall into one of two categories: natural progesterone and synthetic progesterone (also called progestin). Natural progesterone drugs are bioidentical hormones prepared made from plant sources (e.g., wild yams). Synthetic progesterone drugs, also called progestins, can bind to the progesterone receptor, but their chemical structure is modified from the naturally-occurring progesterone in humans. Natural progesterone is used in assisted reproductive technology and, occasionally, is used for cervical ~~widening~~ ripening in the induction of labor in late pregnancy. Progestins are indicated for use in reproductive-aged women as birth control, treatment of menstrual disorders (e.g., amenorrhea, uterine bleeding), treatment of endometriosis, and prevention or treatment of preterm labor (Brucker and Likis 2010). Progestins are also administered as fertility-sparing, hormonal treatment of low grade, early stage uterine cancer in~~drugs for~~ reproductive aged women ~~administered cancer chemotherapy agents for the treatment of cancer~~. The nomenclature for natural and synthetic progesterone is challenging to differentiate because both types of drugs are often used to treat the same condition, and the term "progesterone" is frequently used interchangeably between the natural progesterone (plant-derived) and progestins. For example, some researchers consider any non-endogenous progesterone to be a progestin, while other researchers argue for a clearer distinction between natural progesterone versus progesterone-derivative drugs (progestins, Romero and Stanczyk 2013).

Concern for possible adverse effects of progestogens stems from the observation that alterations in normal steroid hormone exposure during development can cause adverse effects on offspring health and development. For example, congenital adrenal hyperplasia is a genetic disorder causing a~~n elevation of progesterone and androgens.~~ deficiency of the enzymes involved in steroidogenesis (*e.g.*, 21-hydroxylase enzyme), which results in elevated progesterone and androgen levels. Health outcomes associated with classical congenital adrenal hyperplasia are ambiguous genitalia at birth, accelerated development of external genitalia during childhood, menstrual cycle irregularities, difficulties getting pregnant, and altered sexually dimorphic behaviors (Turcu and Auchus 2015, Witchel 2017). Virilization of female infants and hypospadias in male infants have been reported following in utero exposure to different types of progestogens during the first trimester of pregnancy (Wilkins *et al.* 1958, Grumbach *et al.* 1959, Aarskog 1971). Similar effects on reproductive development as well as sexually dimorphic behavior have been reported in animal studies following prenatal exposure to some progestogens.

## Significance

Because of the vulnerability of fetal development to alterations in steroid hormone, the association between prenatal exposure to progestogens and potential adverse health effects was identified as a potential candidate for systematic review. Thus, the National Toxicology Program (NTP) at NIEHS will conduct a scoping review to characterize the extent of evidence published on this topic. This scoping review and associated interactive evidence map is intended to support decision making on this topic regarding whether the database is likely to support hazard characterization conclusions for one or more health effects in a full systematic review or for consideration of future research on prenatal exposure to progestogens and adverse health effects in offspring

# OBJECTIVE AND SPECIFIC AIMS

## Objective

The primary objective of this scoping review is to identify and characterize the literature relevant to progestogen use during pregnancy and offspring health effects, including prematurity-related neonatal outcomes, congenital malformations, neurological effects, cancer, and other health outcomes related to prenatal exposure.

## Specific Aims

- Identify literature reporting on adverse pregnancy outcomes (prematurity-related outcomes, congenital malformations), neurological effects, cancer, and other health outcomes (e.g., reproductive system effects) related to prenatal exposure to progestogens reported in epidemiological, experimental animal, and in vitro model systems studies.
- Summarize data on the health effects with the largest bodies of evidence. If appropriate, extract data on potential health effects identified from relevant studies. (Data extraction files of the included studies will be shared upon release of final report.)
- Summarize and create an evidence map of relevant health effects and mechanistic data by exposure to progestogens overall and by individual formulations (i.e., the extent and types of health effects evidence available by type of drugs).

## PECO Statement

A PECO statement (**P**opulation, **E**xposure(s), **C**omparator(s), and **O**utcome(s)) (**Table 1**) was developed to address and understand the adverse effects of developmental exposure to progestogens reported in humans, animals, and *in* vitro 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 and Green 2011).

**Table 1. Human PECO (Population, Exposure, Comparator, and Outcome) Statement**

| Element | Evidence |
|---|---|
| Population | **Human:** All epidemiological studies<br>**Animal:** Non-human, vertebrate laboratory animal models, including but not limited to mice, monkeys, rats, fish, and amphibians<br>***In vitro*:** *In vitro* models utilizing organs, tissues, cell lines, or cellular components relating to embryonic or fetal exposure |
| Exposure | Natural progesterone (bioidentical, plant-based) or synthetic progesterone-derivative drugs (called progestins) administered during pregnancy; collectively called progestogens |
| Comparators | Both experimental (controlled exposure or treatment) and observational studies (ecological) should be included; experimental studies should include an untreated or vehicle control |
| Outcomes | Pregnancy outcomes (e.g., prematurity-related neonatal outcomes, ~~,~~ congenital malformations), neurological effects, cancer, and other health outcomes related to prenatal exposure |

# METHODS

The systematic review techniques in this scoping review protocol adhere to the framework developed by Office of Health Assessment and Translation (OHAT) (Rooney *et al.* 2014). This protocol is restricted to the first 3 steps of the 7-step OHAT systematic review framework: 1) Problem Formulation, 2) Search and Select Studies for Inclusion, and, as applicable, 3) Data Extraction. The remaining 4 steps of the OHAT systematic review framework are relevant for assessing study quality and synthesizing evidence, which are beyond the scope of the current report.

## Step 1. Problem Formulation

Prenatal progestin exposure and epigenetic effects was nominated to NTP for possible literature evaluation and/or laboratory research studies in the Spring of 2013. [The nominator used the terms progestin and progesterone interchangeably]. The NTP received further interest in this topic from Drs. Erin Hines (United States Environmental Protection Agency) and John Thorp (University of North Carolina-Chapel Hill) regarding a systematic review of the literature on offspring health outcomes following prenatal progestin exposure with a special interest in neurological outcomes in adulthood following prenatal exposure to 17-[alpha] hydroxyprogesterone [caproate] (Erin Hines, personal communication, 14 May 2013). A preliminary focused literature search of progestins was performed in PubMed on April 15, 2015 yielding 1,483 references; this literature search targeted references reporting on adverse effects or toxicity of these drugs in combination with pregnancy outcomes. The screening of 1,511 references (1,483 references identified by the literature search strategy plus 28 handpicked references) was performed by NTP and ICF staff and Ms. Porscha Walton (NIEHS Scholar Connect Program participant). A total of 313 original research studies were identified as relevant: human (n=197), animal (n=109), and in vitro/mechanistic studies (n=8). In addition, 14 meta-analyses were identified. Most of the human studies evaluated the efficacy of the drug to reduce early preterm labor as the primary outcome, while perinatal outcomes other than preterm birth and miscarriage were either secondary outcomes of these studies or the offspring outcomes were not reported. There were no original research articles identified reporting on progestin-induced fetal germline disruption (i.e., transgenerational effects), while there were collections of studies evaluating the incidence of reproductive tract malformations (e.g., hypospadias) as well as neurological development. Thus, the preliminary targeted literature search was useful in identifying relevant studies that could be used to revise the final literature search strategy for an NTP scoping review to ensure that all relevant articles will be retrieved.

To complete a comprehensive scoping review, the NTP will revise the preliminary targeted literature search strategy to broaden expand the health outcome terms used to identify studies reporting pregnancy outcomes, congenital malformations, developmental effects (e.g., neurological development and disease), and other health outcomes related to prenatal exposure to progestins (e.g., progesterone and progestins).

### *Pharmaceutical drug selection*

A list of progestins to be queried has been developed from a list received from a federal advisor (Erin Hines, personal communication), supplemented by progestin drugs identified in Brucker *et al.* (2010) and Schindler (2008), and in consultation with two obstetricians (Anne Z. Steiner, University of North Carolina-Chapel Hill and John O'Brien, University of Kentucky).

## Step 2. Search and Select Studies for Inclusion

### *Literature Search Strategy*

A literature search strategy was designed using (1) progesterone or progestin as title and abstract search terms, (2) specific progestin MeSH terms, and (3) specific progestin drug names as title and abstract terms, each in combination with health outcomes of interest (e.g., pregnancy outcomes, congenital anomalies, developmental outcomes, and other terminology to identify health outcomes associated with prenatal exposure) (see **Appendix 1**). Only the PubMed and Cochrane Library databases will be searched due to the estimated number of references to be retrieved (approximately 15,000 references). No publication year or language limits will be imposed.

### *Searching Other Resources*

Depending on the number of references available, we will hand-search the reference lists of relevant, authoritative reviews or government-authored (state and federal) technical reports identified during the initial search to identify additional studies that were not identified through the electronic searches. Studies will be evaluated using the same inclusion and exclusion criteria as used for screening records retrieved from the electronic search. Relevant studies identified through these steps will be marked as "provided from other sources" in the study selection flow diagram and added at the full text screening level. If the literature base is relatively large and the number of reviews is plentiful, authoritative reviews or government-authored (state and federal) technical reports will be binned for future reference but will not be hand-searched for references.

### *Screening Process*

Two web-based software systems will be used to screen for literature screening: Swift Active Screener (Sciome, Research Triangle Park, NC) and DistillerSR® (Evidence Partners, Ottawa, CA). Initial title and abstract screening will be performed in Swift Active Screener to identify the potentially relevant references; this software uses machine-learning and text-mining technology to prioritize the unscreened references in order of most relevant to least relevant based on the results of screened references. All references identified as potentially relevant will be further screened by additional title and abstract screening (as necessary) and full text screening in DistillerSR®, a web-based, systematic review software program with structured forms and a searchable database of screening results[1]. 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 prior to uploading the references into Swift Active Screener.

The references identified by the literature search will be screened at both the title/abstract for relevance, and full text levels to confirm relevance and to characterize the original research articles by: language type (e.g., English language, non-English language), publication type, evidence stream, exposure type (e.g., progestin or natural progesterone evaluated), reason for drug use, and category of health outcomes. The relevant studies will be summarized by outcome and drug in text and tables.

In order to be eligible for inclusion in the evidence map of health effects, references must comply with the criteria specified by the PECO statement (**Table 1**) or contain relevant exposure assessment information. In addition to the PECO criteria, the following exclusion criteria will apply to the full text: references that do not contain original data (e.g., reviews, editorials, or commentaries); and references that have not been peer-reviewed (e.g., conference abstracts, technical reports, theses/dissertations,

[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.

working papers from research groups or committees, and white papers). There are no limitations on the language of the publication.

The inclusion and exclusion criteria that will be used to screen references for relevance and eligibility at both the title-and-abstract and full-text screening stages are summarized in **Table 2**.

**Title and Abstract Screening**

Two members of the evaluation design team will independently conduct title and abstract screening of the literature search results in Swift Active Screener to determine whether a reference meets the inclusion criteria. Screening shall be conducted until Swift Active Screener predicts 95% of relevant references are identified. References will be identified as either "yes, relevant or potentially relevant (e.g. unclear)" or "no, not relevant at the title and abstract level in Swift Active Screener will be screened in a full-text review in Distiller. All relevant reviews will be included at the title and abstract screening. Initially, screeners will be trained using project-specific written instructions in a pilot phase undertaken to improve clarity of the inclusion and exclusion instructions and to improve accuracy and consistency among screeners. 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. Screening conflicts will be resolved by the project lead (Howdeshell). Any articles with unresolved screening conflicts at the title and abstract phase will be included in the full text screening phase.

| Table 2. Detailed inclusion and exclusion criteria to determine study eligibility | | |
|---|---|---|
| **Evidence stream** | **Inclusion Criteria** | **Exclusion Criteria (or blank if none)** |
| ***Participants/Population (Human Studies or Experimental Model Systems)*** | | |
| **Human** | • Reproductive-aged females<br>• No restrictions on country of residence/origin, lifestyle, race/ethnicity, or occupation | • Males<br>• Postmenopausal women<br>• Women treated for infertility |
| **Animal** | • Pregnant adult female non-human animals or non-human animal embryos<br>• Studies in vertebrate laboratory animal models without restriction of species (e.g., rodents, rabbits, fish, and amphibians) | • Non-pregnant non-human animals<br>• Agricultural livestock (e.g., horses, cows, sheet, pigs)<br>• Invertebrates, fungi, plants, bacteria |
| ***In vitro*** | • *In vitro* models utilizing organs, tissues, cell lines, or cellular components | • None |
| ***Exposure*** | | |
| **Human, animal, *In vitro*** | • Exposure to progesterone (natural structure) or progestins (modified structure based on progesterone)<br>• Exposure during pregnancy or just prior to pregnancy<br>• No restriction on route of exposure | • Combination exposure (e.g., combination estrogen/progestogen contraception) |
| ***Comparators*** | | |
| **Human, animal, *In vitro*** | • Individuals exposed to lower levels (or no exposure/exposure below detection levels) of progesterone or progestin<br>• Case reports and case series will be identified, but not categorized by health effect and exposure | • None |
| ***Outcomes*** | | |
| **Human, animal, *In vitro*** | • Adverse pregnancy outcomes (e.g., prematurity-related neonatal outcomes), congenital malformations, neurological effects, cancer and other health outcomes in offspring related to prenatal exposure | • Ectopic pregnancy<br>• Maternal health outcomes only |
| ***Publications (e.g., language restrictions, use of conference abstracts, etc.)*** | | |
| **Human, animal, *In vitro*** | • Reference must contain original data and must be peer-reviewed<br>• References published in a language other than English will be identified, but not categorized by health effect and exposure; references will not be translated. | • Articles with no original data (e.g., editorials, reviews, commentaries)<br>• Non-peer reviewed articles: Conference abstracts or other studies published in abstract form only, grant awards, and theses/dissertations<br>• Retracted articles |
| *Relevant reviews are used as background and will be scanned for relevant references. | | |

**Full-Text Screening**

As needed following the title and abstract screen, full-text will be retrieved[2] for those references where eligibility to meet the inclusion criteria is unclear and/or the title and abstract did not provide sufficient

[2] OHAT will initially attempt to retrieve a full-text copy of the reference using Endnote or an automated program (e.g., QUOSA) when possible, and NIH library services (NIH subscriptions and interlibrary loans). For references not

detail to characterize the reference. Two members of the evaluation design team will independently conduct a full-text screen of the search results to determine whether a reference meets the inclusion criteria. Screening conflicts will be resolved by the project lead (Howdeshell) or, if necessary, through consultation with technical advisors.

*Multiple publications of same data*

Multiple publications with overlapping data for the same relevant study (e.g., publications reporting subgroups, additional outcomes or exposures outside the scope of an evaluation, or longer follow-up) will be identified by examining author affiliations, study designs, cohort name, enrollment criteria, and enrollment dates. If necessary, study authors will be contacted to clarify any uncertainty about the independence of two or more publications. OHAT will include all publications on the study, select one study to use as the primary, and consider the others as secondary publications with annotation as being related to the primary record during data extraction. The primary study will generally be the publication with the longest follow-up, or for studies with equivalent follow-up periods, the study with the largest number of cases or the most recent publication date. OHAT will include relevant data from all publications of the study. If the same outcome is reported in more than one report, OHAT will include only a single instance of the data; thus, excluding duplicate reporting of the data. Although only one publication is identified as the primary study, relevant information will be considered from the secondary publications. For example, when a study refers to a previous publication for additional details of the methods, those citations will be identified and considered with the primary citation for data extraction.

*Tracking study eligibility and reporting the flow of information*

The reason for exclusion at the full-text-review stage will be annotated and reported in a study flow diagram in the final report. A reference will be excluded if it: (1) is a review, commentary, or editorial with no original data; (2) lacks a relevant population; (3) lacks relevant exposure information; (4) lacks a relevant comparator; (5) lacks a relevant health outcome; (6) is a conference abstract, thesis/dissertation, or (7) no PDF is available[2]. Review articles may be characterized by health outcomes and review type (e.g., narrative review, systematic review) to determine when and if specific health outcomes have been evaluated.

## Step 3. Data Extraction, and Content Management

Data extraction, if needed, will be managed with structured forms, and stored in a database format using Health Assessment Workspace Collaborative (HAWC), an open source, web-based interface[3]. Data extraction elements are listed in appendices for human (**Appendix 2**), experimental animal (**Appendix 3**, and *in vitro* studies (**Appendix 4**). Data will be extracted only for health effects that are relevant to the PECO statement. Study information collected during data extraction will be visualized, when appropriate (e.g., when there are data on the same or health effects evaluated across multiple studies) and will be made publicly available upon publication of the finalized report.

The data extraction will be used to summarize the study designs and findings of relevant studies. The content of the data extraction may be revised following the identification of the studies included in the

---

available through NIH, OHAT will search the Internet and/or may attempt to contact the corresponding author. Studies not retrieved through these mechanisms are excluded and notated as "no PDF available."

[3] HAWC (Health Assessment Workspace Collaborative): A Modular Web-based Interface to Facilitate Development of Human Health Assessments of Chemicals (**http://hawcproject.org**).

review. Data extraction will be performed by one member of the evaluation team and checked by a second member of the evaluation team for completeness and accuracy. Data extractors from the evaluation team will be trained using project-specific written instructions in an initial pilot phase using a subset of studies. Any discrepancies in data extraction will be resolved by discussion or consultation with a third member of the evaluation team. Information that is inferred, converted, or estimated during data extraction will be annotated (e.g., using brackets [n=10]).

### Step 4. Study Results and Summaries

The results of all included human, animal, and *in vitro* studies will be summarized by outcome and chemical in text and tables to develop a systematic evidence map of the evidence by health effect and types of evidence and identify data gaps in available research. Visualizations using the extracted data in HAWC will also be generated to summarize the data, if data extraction is deemed necessary.

## SCOPING REVIEW: OUTLINE

The NTP Scoping Review on the adverse developmental effects of progestins will include the following information:

### Introduction

This section will provide a brief background on the topic.

### Objectives and Specific Aims

### Methodology

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, and
- characterization and, as appropriate, data extraction

### Results

This section will include the results from the scoping efforts. Results will be presented in tables or figures as appropriate using Tableau and, as appropriate, HAWC. The results from the included studies will be discussed by outcome. This will include a description of:

- literature search results (description and PRISMA figure),
- Tableau figures to support an evidence map of the key concepts
- summaries of included studies and results in HAWC, and limited visualizations of study result summaries may be generated in HAWC, as appropriate.

# REFERENCES

Aarskog D. 1971. Intersex conditions masquerading as simple hypospadias. *Birth Defects Orig Art Ser* 7(6): 122-130.

Brucker M, Likis F. 2010. Steroid hormones. In: *Pharmacology for Women's Health*. King T, Brucker M, eds. Burlington, MA, Jones & Bartlett Publishers.

Grumbach MM, Ducharme JR, Moloshok RE. 1959. On the fetal masculinizing action of certain oral progestins. *The Journal of clinical endocrinology and metabolism* 19: 1369-1380.

Higgins J, Green S. 2011. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 (updated March 2011). Available: http://handbook.cochrane.org/ [accessed 3 February 2013].

Romero R, Stanczyk FZ. 2013. Progesterone is not the same as 17alpha-hydroxyprogesterone caproate: implications for obstetrical practice. *Am J Obstet Gynecol* 208(6): 421-426.

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.

Schindler AE, Campagnoli C, Druckmann R, Huber J, Pasqualini JR, Schweppe KW, Thijssen JH. 2008. Classification and pharmacology of progestins. *Maturitas* 61(1-2): 171-180.

Turcu AF, Auchus RJ. 2015. The next 150 years of congenital adrenal hyperplasia. *J Steroid Biochem Mol Biol* 153: 63-71.

Wilkins L, Jones HW, Jr., Holman GH, Stempfel RS, Jr. 1958. Masculinization of the female fetus associated with administration of oral and intramuscular progestins during gestation: non-adrenal female pseudohermaphrodism. *The Journal of clinical endocrinology and metabolism* 18(6): 559-585.

Witchel SF. 2017. Congenital Adrenal Hyperplasia. *J Pediatr Adolesc Gynecol* 30(5): 520-534.

# 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. Technical advisors were screened for conflict of interest prior to their service and did not report any conflicts of interest. Epidemiologists and toxicologists on OHAT 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 |
|---|---|
| Kembra Howdeshell | NIEHS/NTP, Project Lead |
| Andrew A Rooney | NIEHS/NTP |
| Mike Shelby | NIEHS/NTP |
| Erin Hines | EPA |

### *Contract Support Staff: Will assist in screening and data extraction*

| Name | Affiliation |
|---|---|
| Ali Goldstone | ICF |
| Susan Goldhaber | ICF |
| Robyn Blain | ICF |

### *Technical Advisors*

Technical advisors are outside experts retained on an as-needed basis to provide individual advice to the NTP for a specific topic. The technical advisors were selected for their ongoing work with, or experience with progestins and offspring health.

| Name | Affiliation |
|---|---|
| Erin Hines, PhD. | U.S. EPA |

## Sources of Support

National Institute of Environmental Health Sciences/Division of the National Toxicology Program

## Protocol History and Revisions

| Date | Activity or revision |
|---|---|
| **September 26, 2018** | Protocol finalized<br>https://ntp.niehs.nih.gov/ntp/ohat/progestogens/progestogens_protocol.pdf |
| **April 14, 2020** | Protocol revised and posted. Protocol revisions included:<br>1. Correcting page numbering error in the protocol beginning on page 6<br>2. Added weblink for this NTP project to Protocol History and Revisions |

| Date | Activity or revision |
|---|---|
| **July 31, 2020** | Protocol revised. Text that was deleted is crossed out and inserted text is underlined. Revisions addressed errors observed in protocol during external peer review, including:<br>1. Background (page 1), changing "cervical widening" to "cervical ripening,"<br>2. Background (page 1), correcting the description of progesterone to "fertility sparing, hormonal treatment for low-grade, early stage uterine cancer in reproductive aged women",<br>3. Background (page 1), removing an extra phrase from the description of congenital adrenal hyperplasia ("congenital adrenal hyperplasia is a genetic disorder causing ~~an elevation of progesterone and androgens.~~ deficiency of the enzymes involved in steroidogenesis", and<br>4. Table 1 PECO Statement, removing an extra comma and space from the last row, second column. |

# APPENDICES

## Appendix 1. Literature Search Strategy

The strategy for this search is broad for the consideration of all offspring health outcomes endpoints and comprehensive for progestins administered during pregnancy as an exposure or treatment in order to ensure inclusion of relevant papers. Only PubMed, Cochrane Library, and DARE databases will be searched.

### Table 3. Pubmed database search terms

| Database | Search Terms |
|---|---|
| Set | Search Strategy |
| Exposure/Progestins | ((Progesterone[mh] OR progesterone-congeners[mh] OR progestins[mh] OR exogenous-progesterone[tiab] OR Algestone[tiab] OR Algestone-acetophenide[tiab] OR Allylestrenol[tiab] OR Chlormadinone-acetate[tiab] OR cyproterone-acetate[tiab] OR desogestrel[tiab] OR dydrogesterone[tiab] OR Ethisterone[tiab] OR ethynodiol-diacetate[tiab] OR Flurogestone-acetate[tiab] OR Gestonorone-caproate[tiab] OR Gestrinone[tiab] OR Levonorgestrel[tiab] OR Lynestrenol[tiab] OR medrogestone[tiab] OR medroxyprogesterone-acetate[tiab] OR Megestrol-acetate[tiab] OR Norethindrone[tiab] OR norethynodrel[tiab] OR Norgestrel[tiab] OR Promegestone[tiab] OR 11-hydroxyprogesterone[tiab] OR 17-alpha-hydroxy-progesterone-caproate[tiab] OR 17-alpha-hydroxyprogesterone-caproate[tiab] OR 17-hydroxyprogesterone-caproate[tiab] OR Demegestone[tiab] OR Dienogest[tiab] OR Drospirenone[tiab] OR Etonogestrel[tiab] OR Gestodene[tiab] OR Nomegestrol-acetate[tiab] OR Norelgestromin[tiab] OR Norethindrone-acetate[tiab] OR Norgestimate[tiab] OR ST-1435[tiab] OR Trimegestone[tiab] OR 11-hydroxyprogesterone[nm] OR 17-alpha-hydroxy-progesterone-caproate[nm] OR Demegestone[nm] OR Dienogest[nm] OR Drospirenone[nm] OR Etonogestrel[nm] OR Gestodene[nm] OR Nomegestrol-acetate[nm] OR Norelgestromin[nm] OR Norethindrone-acetate[nm] OR Norgestimate[nm] OR ST-1435[nm] OR Trimegestone[nm] OR Algestone[mh] OR Algestone-acetophenide[mh] OR Allylestrenol[mh] OR Chlormadinone-acetate[mh] OR cyproterone-acetate[mh] OR desogestrel[mh] OR dydrogesterone[mh] OR Ethisterone[mh] OR ethynodiol-diacetate[mh] OR Flurogestone-acetate[mh] OR Gestonorone-caproate[mh] OR Gestrinone[mh] OR Levonorgestrel[mh] OR Lynestrenol[mh] OR medrogestone[mh] OR medroxyprogesterone-acetate[mh] OR Megestrol-acetate[mh] OR Norethindrone[mh] OR norethynodrel[mh] OR Norgestrel[mh] OR Promegestone[mh] OR progestogens[tiab] OR progestogen[tiab] OR Hydroxyprogesterone-caproate[tiab]) OR ((progesterone[tiab] OR progestins[tiab]) NOT medline[sb])) |
| Window of Exposure | AND<br>(animals,-newborn[mh] OR child[mh] OR infant[mh] OR maternal-exposure[mh] OR Maternal-Fetal Exchange[mh] OR babies[tiab] OR baby[tiab] OR child*[tiab] OR embryo[tiab] OR embryonic[tiab] OR embryos[tiab] OR fetal[tiab] OR fetus[tiab] OR foetal[tiab] OR gestation*[tiab] OR infant*[tiab] OR in-utero[tiab] OR maternal[tiab] OR neonat*[tiab] OR newborn*[tiab] OR offspring[tiab] OR perinat*[tiab] OR postnat*[tiab] OR prenat*[tiab] OR progeny[tiab] OR pups[tiab]) |
| Outcome | AND<br>(pregnancy-outcomes[mh] OR prenatal-exposure-delayed-effects[mh] OR premature-birth[mh] OR premature-birth[tiab] OR abortion,-spontaneous[mh] OR spontaneous-abortion[tiab] OR live-born[tiab] OR still-born[tiab] OR stillborn[tiab] OR stillbirth[tiab] OR still-birth[tiab] OR Congenital-Abnormality[tiab] OR Congenital-Abnormalities[tiab] OR Congenital-Abnormalities[mh] OR Deformities[tiab] OR Deformity[tiab] OR Congenital-defect[tiab] OR Congenital-defects[tiab] OR birth-defects[tiab] OR birth-defect[tiab] OR Congenital-anomalies[tiab] OR Congenital-anomaly[tiab] OR hypospadias[tiab] OR Behavior-and-Behavior-Mechanisms[mh] OR Gene-Expression-Regulation[mh] OR Glucuronosyltransferase[mh] OR Intelligence-tests[mh] OR Malate-Dehydrogenase[mh] OR Mediator-Complex-Subunit-1[mh] OR Mental-disorders[mh] OR Mental- |

| Database | Search Terms |
|---|---|
| | processes[mh] OR Monocarboxylic-Acid-Transporters[mh] OR Myelin-Basic-Protein[mh] OR nervous-system[mh] OR nervous-system-diseases[mh] OR nervous-system-physiological-phenomena[mh] OR Neurogranin[mh] OR Oligodendroglia[mh] OR Peroxisome-Proliferator-Activated-Receptors[mh] OR Psychological-Phenomena-and-Processes[mh] OR Academic-performance[tiab] OR auditory[tiab] OR cortical[tiab] OR delayed-development[tiab] OR developmental-impairment[tiab] OR developmental-delay*[tiab] OR developmental-disorder*[tiab] OR euthyroid[tiab] OR gait[tiab] OR glia*[tiab] OR gliogenesis[tiab] OR hyperactiv*[tiab] OR impulse-control[tiab] OR IQ[tiab] OR ischemi*[tiab] OR locomotor[tiab] OR mental-deficiency[tiab] OR mental-development[tiab] OR mental-illness[tiab] OR mental-deficit[tiab] OR mobility[tiab] OR mood[tiab] OR morris-maze[tiab] OR morris-water[tiab] OR motor-abilit*[tiab] OR Motor-activities[tiab] OR motor-performance[tiab] OR nerve[tiab] OR neural[tiab] OR neurobehav*[tiab] OR Neurocognitive-impairment[tiab] OR neurodegenerat*[tiab] OR Neurodevelopment*[tiab] OR neurodisease*[tiab] OR neurologic*[tiab] OR neuromuscular[tiab] OR neuron*[tiab] OR neuropath*[tiab] OR obsessive-compulsive[tiab] OR OCD[tiab] OR olfaction[tiab] OR olfactory[tiab] OR open-field-test[tiab] OR passive-avoidance[tiab] OR plasticity[tiab] OR sociab*[tiab] OR speech*[tiab] OR spelling[tiab] OR stereotypic-movement*[tiab] OR synap*[tiab] OR tauopath*[tiab] OR visual-motor[tiab] OR Visuospatial-processing[tiab] OR water-maze[tiab] OR active-avoidance[tiab] OR ADHD[tiab] OR amygdala[tiab] OR antisocial[tiab] OR anxiety[tiab] OR anxious[tiab] OR asperger*[tiab] OR attention-deficit[tiab] OR autism[tiab] OR autistic[tiab] OR behavioral[tiab] OR behaviors[tiab] OR behavioural[tiab] OR behaviours[tiab] OR bipolar[tiab] OR cerebellum[tiab] OR cognition[tiab] OR cognitive[tiab] OR comprehension[tiab] OR cranial[tiab] OR dementia[tiab] OR dendrit*[tiab] OR dentate-gyrus[tiab] OR depression[tiab] OR down-syndrome[tiab] OR dyslexia[tiab] OR entorhinal-cortex[tiab] OR epilep*[tiab] OR ganglia*[tiab] OR ganglion*[tiab] OR hearing[tiab] OR hippocamp*[tiab] OR human-development[tiab] OR impulsiv*[tiab] OR Intellectual-disability[tiab] OR intelligence[tiab] OR language[tiab] OR learning[tiab] OR lewy-bod*[tiab] OR long-term-potentiation[tiab] OR long-term-synaptic-depression[tiab] OR memory[tiab] OR mental-disorder*[tiab] OR mental-recall[tiab] OR Motor-activity[tiab] OR motor-skill*[tiab] OR myxedema[tiab] OR Nervous-system[tiab] OR nervous-system[tiab] OR neurit*[tiab] OR optic[tiab] OR palsy[tiab] OR panic[tiab] OR parahippocamp*[tiab] OR paranoia[tiab] OR paranoid[tiab] OR perception[tiab] OR perforant*[tiab] OR personality[tiab] OR phobia[tiab] OR problem-solving[tiab] OR proprioception[tiab] OR psychomotor[tiab] OR reflex[tiab] OR risk-taking[tiab] OR schizophrenia[tiab] OR seizure*[tiab] OR sensation*[tiab] OR sleep[tiab] OR smell[tiab] OR spatial-behavior[tiab] OR substantia-nigra[tiab] OR child-development[mh] OR child-development[tiab] OR acanthoma*[tiab] OR acrochord*[tiab] OR acrospiroma*[tiab] OR adamantinoma*[tiab] OR adenoacanthoma*[tiab] OR adenoameloblast*[tiab] OR adenocarcin*[tiab] OR adenofibrom*[tiab] OR adenol*[tiab] OR adenom*[tiab] OR adenosquamous[tiab] OR ameloblast*[tiab] OR androblast*[tiab] OR angiofib*[tiab] OR angiog*[tiab] OR angiok*[tiab] OR angiol*[tiab] OR angiom*[tiab] OR angiomatosis[mh] OR angiosarc*[tiab] OR antibodies, neoplasm[mh] OR antigens, neoplasm[mh] OR apudom*[tiab] OR argentaffin*[tiab] OR arrhenoblast*[tiab] OR astroblast*[tiab] OR astrocytom*[tiab] OR astrogliom*[tiab] OR atypia[tiab] OR baltoma[tiab] OR barrett esophagus[mh] OR blastom*[tiab] OR cancer[tiab] OR cancero*[tiab] OR cancers[tiab] OR carcinog*[tiab] OR carcinogenicity tests[mh] OR carcinogens[mh] OR carcinoid*[tiab] OR carcinom*[tiab] OR carcinos*[tiab] OR cavernom*[tiab] OR cell line, tumor[mh] OR cementom*[tiab] OR cerumin*[tiab] OR chloroma*[tiab] OR cholangio*[tiab] OR chondrob*[tiab] OR chondrom*[tiab] OR chondros*[tiab] OR chord*[tiab] OR chorioa*[tiab] OR choriocarc*[tiab] OR chorioep*[tiab] OR chorionep*[tiab] OR chromaffinom*[tiab] OR collagenom*[tiab] OR comedocarcinom*[tiab] OR condylom*[tiab] OR condylomata acuminata[mh] OR corticotrop*[tiab] OR craniopharyng*[tiab] OR cylindrom*[tiab] OR cystadeno*[tiab] OR cystoma*[tiab] OR cystosa*[tiab] OR dentinom*[tiab] OR dermatofibro*[tiab] OR dermoid[tiab] OR desmoid[tiab] OR desmoplastic*[tiab] OR dictyota[tiab] OR dysgerm*[tiab] OR dyskerat*[tiab] OR dysmyelopoiesis[tiab] OR dysplas*[tiab] OR ectomesenchym*[tiab] OR elastofibr*[tiab] OR enchondrom*[tiab] OR endotheliom*[tiab] OR ependymo*[tiab] OR epidermoid*[tiab] OR epitheliom*[tiab] OR erythrol*[tiab] OR erythropl*[tiab] |

| Database | Search Terms |
|---|---|
| | OR esthesioneuro*[tiab] OR etiolog*[tiab] OR fibroaden*[tiab] OR fibrochond*[tiab] OR fibroe*[tiab] OR fibrofol*[tiab] OR fibroid*[tiab] OR fibrolip*[tiab] OR fibrom*[tiab] OR fibroodontom*[tiab] OR fibrosarcom*[tiab] OR fibrothecom*[tiab] OR fibroxantho*[tiab] OR ganglioblast*[tiab] OR gangliocytom*[tiab] OR gangliogliom*[tiab] OR ganglioneuro*[tiab] OR gastrinom*[tiab] OR genes, neoplasm[mh] OR germinom*[tiab] OR glioblast*[tiab] OR gliom*[tiab] OR glomangio*[tiab] OR glucagonom*[tiab] OR gonadoblastom*[tiab] OR gonocytom*[tiab] OR gynandroblastom*[tiab] OR haemangio*[tiab] OR hamartom*[tiab] OR hemangio*[tiab] OR hepatoblastom*[tiab] OR hepatom*[tiab] OR hibernom*[tiab] OR hidradenom*[tiab] OR hidrocy*[tiab] OR hodgkin*[tiab] OR hydatidiform*[tiab] OR hydradenom*[tiab] OR hypernephrom*[tiab] OR IARC[tiab] OR immunocytom*[tiab] OR insulinom*[tiab] OR leiomyo*[tiab] OR lesion*[tiab] OR leukaemia*[tiab] OR leukemia*[tiab] OR leukoplak*[tiab] OR leukostas*[tiab] OR leukostasis[mh] OR lipoadenom*[tiab] OR lipoblastom*[tiab] OR lipom*[tiab] OR liposarcom*[tiab] OR luteinom*[tiab] OR luteom*[tiab] OR lymphangio*[tiab] OR lymphoepitheliom*[tiab] OR lymphom*[tiab] OR lymphoscintigraph*[tiab] OR macroglobulinem*[tiab] OR macroprolactinom*[tiab] OR malignan*[tiab] OR maltom*[tiab] OR masculinovoblastom*[tiab] OR mastocyto*[tiab] OR mcf-7[tiab] OR medullo*[tiab] OR meigs syndrome[tiab] OR melanoa*[tiab] OR melanocytom*[tiab] OR melanom*[tiab] OR meningio*[tiab] OR mesenchymom*[tiab] OR mesonephrom*[tiab] OR mesotheliom*[tiab] OR metaplas*[tiab] OR metaplasia[mh] OR metastas*[tiab] OR metastat*[tiab] OR microgliom*[tiab] OR micrometastas*[tiab] OR mucositis[mh] OR mycosis fungoides*[tiab] OR myelodysplas*[tiab] OR myelodysplastic syndromes[mh] OR myelodysplastic-myeloproliferative diseases[mh] OR myelofibrosis[tiab] OR myelol*[tiab] OR myeloma*[tiab] OR myeloproliferat*[tiab] OR myeloproliferative disorders[mh] OR myelosuppression*[tiab] OR myoblastom*[tiab] OR myoepitheliom*[tiab] OR myofibro*[tiab] OR myolipom*[tiab] OR myoma*[tiab] OR myosarcom*[tiab] OR myxof*[tiab] OR myxom*[tiab] OR naevus[tiab] OR neoplas*[tiab] OR neoplasm proteins[mh] OR neoplasms[mh] OR neoplastic stem cells[mh] OR nephroblastom*[tiab] OR neurilem*[tiab] OR neurinom*[tiab] OR neuroblastom*[tiab] OR neurocytom*[tiab] OR neuroepitheliom*[tiab] OR neurofibro*[tiab] OR neurolipocytom*[tiab] OR neuroma*[tiab] OR neuronevus[tiab] OR neurothekeom*[tiab] OR nevus[tiab] OR non coding RNA[tiab] OR nonseminom*[tiab] OR odontoam*[tiab] OR odontom*[tiab] OR oligoastrocytom*[tiab] OR oligodendrogliom*[tiab] OR oncocytom*[tiab] OR oncogen*[tiab] OR oncogene fusion[mh] OR oncogene proteins[mh] OR oncogenic viruses[mh] OR oncolog*[tiab] OR oncolytic viruses[mh] OR oncoprotein*[tiab] OR opsoclonus-myoclonus[tiab] OR orchioblastom*[tiab] OR osteoblastom*[tiab] OR osteoch*[tiab] OR osteofibrosarcom*[tiab] OR osteom*[tiab] OR osteosarcom*[tiab] OR pancreatoblastom*[tiab] OR papillom*[tiab] OR parachordom*[tiab] OR paragangliom*[tiab] OR paraneoplas*[tiab] OR perineuriom*[tiab] OR phaeochromocytom*[tiab] OR pheochromo*[tiab] OR pilomatri*[tiab] OR plasmacytom*[tiab] OR pneumoblast*[tiab] OR pneumocytom*[tiab] OR polyembryom*[tiab] OR polyhistiom*[tiab] OR polyp[tiab] OR polyps[mh] OR porocarcinom*[tiab] OR porom*[tiab] OR pre-cancer*[tiab] OR precancer*[tiab] OR preleukaem*[tiab] OR preleukem*[tiab] OR prelymphom*[tiab] OR pre-lymphom*[tiab] OR pre-malign*[tiab] OR premalignan*[tiab] OR preneoplas*[tiab] OR pre-neoplas*[tiab] OR prolactinom*[tiab] OR protooncogen*[tiab] OR pseudotum*[tiab] OR reninom*[tiab] OR retinoblastom*[tiab] OR rhabdo*[tiab] OR RNA, neoplasm[mh] OR sarcoma*[tiab] OR schwannom*[tiab] OR SEER program[mh] OR seminom*[tiab] OR sentinel lymph node[tiab] OR sentinel lymph node biopsy[mh] OR sertoli-leydig cell tumor[tiab] OR sezary syndrome[tiab] OR somatostatinom*[tiab] OR somatotropinom*[tiab] OR spermatocytom*[tiab] OR spiradenom*[tiab] OR spongioblastom*[tiab] OR subependymom*[tiab] OR thecom*[tiab] OR thymom*[tiab] OR trichilemmom*[tiab] OR trichoadenom*[tiab] OR trichoblastom*[tiab] OR trichodiscom*[tiab] OR trichoepitheliom*[tiab] OR trichofolliculom*[tiab] OR tricholemm*[tiab] OR tumor[tiab] OR tumor markers, biological[mh] OR tumorgen*[tiab] OR tumorig*[tiab] OR tumor-inhibit*[tiab] OR tumorog*[tiab] OR tumors[tiab] OR tumors[tiab] OR tumour[tiab] OR up-regulat*[tiab] OR vipom*[tiab] OR waldenstrom*[tiab] OR xanthoma*[tiab]) |

## Table 4. Cochrane and DARE databse search terms

| Set | Search Strategy |
|---|---|
| Exposure/Progestins | (exogenous-progesterone OR Algestone OR Algestone-acetophenide OR Allylestrenol OR Chlormadinone-acetate OR cyproterone-acetate OR desogestrel OR dydrogesterone OR Ethisterone OR ethynodiol-diacetate OR Flurogestone-acetate OR Gestonorone-caproate OR Gestrinone OR Levonorgestrel OR Lynestrenol OR medrogestone OR medroxyprogesterone-acetate OR Megestrol-acetate OR Norethindrone OR norethynodrel OR Norgestrel OR Promegestone OR 11-hydroxyprogesterone OR 17-alpha-hydroxy-progesterone-caproate OR 17-alpha-hydroxyprogesterone-caproate OR 17-hydroxyprogesterone-caproate OR Demegestone OR Dienogest OR Drospirenone OR Etonogestrel OR Gestodene OR Nomegestrol-acetate OR Norelgestromin OR Norethindrone-acetate OR Norgestimate OR ST-1435 OR Trimegestone OR progesterone OR progestins OR progestogens OR progestogen OR Hydroxyprogesterone-caproate) |
| Window of Exposure | AND<br>(babies OR baby OR child* OR embryo OR embryonic OR embryos OR fetal OR fetus OR foetal OR gestation* OR infant* OR in-utero OR maternal OR neonat* OR newborn* OR offspring OR perinat* OR postnat* OR prenat* OR progeny OR pups) |

# Appendix 2. Data Extraction Elements for Human Studies

| **HUMAN** | |
|---|---|
| ***Funding*** | Funding source(s) |
| | Reporting of conflict of interest (COI) by authors (*reporting bias) |
| ***Subjects*** | Study population name/description |
| | Dates of study and sampling timeframe |
| | Geography (country, region, state, etc.) |
| | Demographics (sex, race/ethnicity, age or lifestage and exposure and outcome assessment) |
| | Number of subjects (target, enrolled, n per group in analysis, and participation/follow-up rates) (*missing data bias) |
| | Inclusion/exclusion criteria/recruitment strategy (*selection bias) |
| | Description of reference group (*selection bias) |
| ***Methods*** | Study design (e.g., prospective or retrospective cohort, nested case-control study, cross-sectional, population-based case-control study, intervention, case report, etc.) |
| | Length of follow-up (*information bias) |
| | Health outcome category, e.g., neurodevelopment |
| | Health outcome, e.g., memory (*reporting bias) |
| | Diagnostic or methods used to measure health outcome (*information bias) |
| | Confounders or modifying factors and how considered in analysis (e.g., included in final model, considered for inclusion but determined not needed (*confounding bias) |
| | Substance name and CAS number |
| | Exposure assessment (e.g., blood, urine, hair, air, drinking water, job classification, residence, administered treatment in controlled study, etc.) (*information bias) |
| | Methodological details for exposure assessment (e.g., HPLC-MS/MS, limit of detection) (*information bias) |
| | Statistical methods (*information bias) |
| ***Results*** | Exposure levels (e.g., mean, median, measures of variance as presented in paper, such as SD, SEM, 75th/90th/95th percentile, minimum/maximum); range of exposure levels, number of exposed cases |
| | Statistical findings (e.g., adjusted β, standardized mean difference, adjusted odds ratio, standardized mortality ratio, relative risk, etc.) or description of qualitative results. When possible, OHAT will convert measures of effect to a common metric with associated 95% confidence intervals (CI). Most often, measures of effect for continuous data are expressed as mean difference, standardized mean difference, and percent control response. Categorical data are typically expressed as odds ratio, relative risk (RR, also called risk ratio), or β values, depending on what metric is most commonly reported in the included studies and on OHAT's ability to obtain information for effect conversions from the study or through author query. |
| | If not presented in the study, statistical power can be assessed during data extraction using an approach that can detect a 10% to 20% change from response by control or referent group for continuous data, or a relative risk or odds ratio of 1.5 to 2 for categorical data, using the prevalence of exposure or prevalence of outcome in the control or referent group to determine sample size. For categorical data where the sample sizes of exposed and control or referent groups differ, the sample size of the exposed group will be used to determine the relative power category. 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 as "appears to be adequately powered" (sample size for 80% power met), somewhat underpowered (sample size is 75% to < 100% of number required for 80% power), "underpowered" (sample size is 50% to < 75% of number required for 80% power), or "severely underpowered" (sample size is < 50% of number required for 80% power). |

| **HUMAN** | |
|---|---|
| | Observations on dose response (e.g., trend analysis, description of whether dose-response shape appears to be monotonic, non-monotonic) |
| ***Other*** | Documentation of author queries, use of digital rulers to estimate data values from figures, exposure unit, and statistical result conversions, etc. |

# Appendix 3. Data Extraction Elements for Animal Studies

| **ANIMAL** | |
|---|---|
| ***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) |
| | Randomization procedure, allocation concealment, blinding during outcome assessment (*selection 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) |
| | Endpoint (e.g., infertility) |
| | Diagnostic or method to measure endpoint (*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, OHAT 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, 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. |

| **ANIMAL** | |
|---|---|
| | 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.<br>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 unit, statistical result conversions, etc. |

# Appendix 4. Data Extraction Elements for *In Vitro* Studies

| ***In vitro*** | |
|---|---|
| ***Funding*** | Funding source(s) |
| | Reporting of COI by authors and/or translators (*reporting bias) |
| ***Cell/Tissue Model*** | Cell line, cell type, or tissue |
| | Source of cells/tissues (and validation of identity) |
| | Sex of human/animal origin |
| | Species |
| | Strain |
| ***Treatment*** | Chemical name and CAS number |
| | Concentration levels (as presented and converted to µM when possible) |
| | Source of chemical |
| | Purity of chemical (*information bias) |
| | Vehicle used for experimental/control conditions |
| | Duration and frequency of dosing (e.g., hours, days, weeks when administration was ended, days per week) |
| ***Methods*** | 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) |
| | Randomization procedure, allocation concealment, blinding during outcome assessment (*selection bias) |
| | Number of replicates per group (*information bias) |
| | Percent serum/plasma in medium |
| | Use of negative controls and whether controls were untreated, vehicle-treated, or both |
| | Report on data from positive controls – was expected response observed? (*information bias) |
| | Endpoint health category (e.g. neurological and thyroid) |
| | Endpoint or assay target (e.g., T3, T4, TSH levels). |
| | Name and source of assay kit |
| | Diagnostic or method to measure endpoint (e.g., reporter gene)(*information bias) |
| | Statistical methods (*information bias) |
| ***Results*** | Measures of effect at each dose or concentration level (e.g., mean, median, frequency, and measures of precision or variance) or description of qualitative results. When possible, OHAT 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 mean difference, standardized mean difference, and percent control response. Categorical data will be expressed as relative risk (RR, also called risk ratio). |
| | No Observed Effect Concentration (NOEC), Lowest Observed Effect Concentration (LOEC), statistical significance of other concentration levels, AC50, or other estimates of effect presented in paper. Note: The NOEC and LOEC are highly influenced by study design, do not give any quantitative information about the relationship between dose and response, and can be subject to author's interpretation (e.g., a statistically significant effect may not be considered biologically important). Also, a NOEC does not necessarily mean zero response. |
| | Observations on dose response (e.g., trend analysis, description of whether dose-response shape appears to be monotonic, non-monotonic) |
| ***Other*** | Documentation of author queries, use of digital rulers to estimate data values from figures, exposure unit, statistical result conversions, etc. |