

Applying the Navigation Guide: Case Study #1

The Impact of Developmental Exposure to
Perfluorooctanoic Acid (PFOA) On Fetal Growth

A Systematic Review of the Non-Human Evidence

Final Protocol

PROTOCOL INFORMATION	4
PROTOCOL.....	6
Background	6
The Navigation Guide.....	6
Human Exposure to Perfluorooctanoic Acid	7
Developmental Exposure to Perfluorooctanoic Acid and Human Health.....	8
Objectives.....	9
Methods	10
Criteria for Selecting Studies	10
Search Methods.....	10
Study Selection Criteria.....	12
Data Collection.....	13
Risk of Bias Determination	14
Statistical Analysis.....	14
Quality and Strength of Evidence Ratings.....	16
Next Steps	17
SUPPLEMENTARY INFORMATION.....	18
Appendix I. Navigation Guide Criteria for Evaluating Strength of Evidence from Non-Human Systems of Evidence	18
Appendix II. PubMed Search Terms.....	20
Appendix III. List of Toxicological Websites To Be Searched.....	26
Appendix IV: Exclusion Criteria Screening Forms.....	27
Title and Abstract Screening Form	27
Full-Text Screening Form.....	27
Appendix V: Data Collection Form.....	29
Appendix VI: Risk of Bias Form.....	39
Appendix VII: Instructions for Making Risk of Bias Determinations.....	43
Appendix VIII: Characteristics to Assess Biological Heterogeneity.....	51
ABOUT THE ARTICLE	53
ACKNOWLEDGEMENTS	53

REFERENCES 54

PROTOCOL INFORMATION

Authors

Erica Koustas, PhD

ORISE Post-doctoral Fellow

U.S. EPA Office of Policy, National Center for Environmental Economics

Juleen Lam PhD, MHS, MS

ORISE Postdoctoral Fellow

U.S. EPA, Office of Policy National Center for Environmental Economics

Patrice Sutton, MPH

Research Scientist

Program on Reproductive Health and the Environment

University of California, San Francisco

Paula I. Johnson, PhD, MPH

Postdoctoral Fellow

Program on Reproductive Health and the Environment

University of California, San Francisco

Dylan Atchley

Research Analyst

Program on Reproductive Health and the Environment

University of California, San Francisco

Saunak Sen, PhD

Associate Professor, Epidemiology and Biostatistics

University of California San Francisco

Karen A. Robinson, PhD

Assistant Professor of Medicine, Epidemiology and Health Policy & Management

Johns Hopkins University

Daniel Axelrad, MPP

Environmental Scientist

U.S. EPA, Office of Policy National Center for Environmental Economics

Tracey J. Woodruff, PhD, MPH

Associate Professor and Director

Program on Reproductive Health and the Environment

University of California, San Francisco

Contact Person: Patrice Sutton, UCSF Program on Reproductive Health and the Environment,
1330 Broadway, Suite 1100, Oakland, CA 94612 suttonp@obgyn.ucsf.edu (510) 986-8927

Dates: October 18, 2011 – February 1, 2013

What's New: Applying a systematic and transparent methodology in the field of environmental health science to ascertain the strength of the evidence linking an environmental exposure to an adverse reproductive and developmental health outcome.

History: Demonstration of proof of concept of *Navigation Guide*¹ by USEPA, UCSF and Johns Hopkins University.

Publication: Our intent is to publish the review in a peer-reviewed journal.

PROTOCOL

Background

The Navigation Guide

In the clinical field, weight of evidence reviews of the scientific research have played a transformative role in speeding the incorporation of science into preventative action at the individual and societal level. But while the clinical sciences point the way, these systems are not fully transferable to environmental health science. This is primarily because of differences between clinical and environmental health sciences in: (1) the types of evidence generally available; and (2) how decisions to expose populations and patients are made. In the clinical arena decisions about exposure to an exogenous substance are made based on weighing risks and benefits to patients' health. There is no comprehensive comparable weighing of health benefits and risks in the environmental arena.

To bridge the gap between clinical and environmental health sciences, the Navigation Guide was developed to evaluate the quality and strength of evidence about the relationship between the environment and reproductive and developmental health.¹ The Navigation Guide is a systematic and transparent methodology that proceeds from best practices in the clinical arena but accounts for the differences in evidence and decision context described above. The Navigation Guide methodology outlines four key steps: 1) specify the study question; 2) select the evidence; 3) rate the evidence; 4) grade the strength of the recommendation.

As part of a proof of concept for the Navigation Guide methodology, this systematic review evaluates the non-human evidence for the effects of exposure to the environmental contaminant perfluorooctanoic acid (PFOA) on fetal growth. The human health rationale for this review relates to the pervasiveness of human exposure to PFOA and human evidence of developmental health impacts, as described below.

This systematic review employs the Navigation Guide methodology through steps 1-3 for the animal evidence stream. The results of the systematic review will be compared to the criteria in the Navigation Guide for rating the strength of the non-human evidence according to one of the following four statements: 1. Sufficient; 2. Limited; 3. Inadequate; or 4. Evidence of lack of toxicity¹ (see Appendix I). A separate systematic review for evaluating the human evidence stream is also under development. The results will be compared to the criteria in the Navigation Guide for rating the strength of the human evidence according to one of the following four statements: 1. Sufficient; 2. Limited; 3. Inadequate; or 4. Evidence of lack of toxicity. Subsequently, the ratings of the strength of the non-human and human evidence will be combined into a summary statement about the overall strength of the evidence according to the criteria in the Navigation Guide. The final result will be one of the following four statements: 1. Known to be toxic; 2. Probably toxic; 3. Possibly toxic; 4. Probably not toxic; or 5. Not classifiable

Human Exposure to Perfluorooctanoic Acid

Perfluorooctanoic acid (PFOA) (CAS# 335-67-1) has been manufactured since the 1950s, primarily as its ammonium salt, ammonium perfluorooctanoate (APFO), for use in the synthesis of fluoropolymers.² Fluoropolymers are industrial compounds that have applications in waterproofing and protective coatings of clothes, furniture, and other products; and also as constituents of floor polish, adhesives, fire retardant foam, and insulation of electrical wire. A major application of one important fluoropolymer, polytetrafluoroethylene, has been the heat-resistant non-stick coatings used on cooking ware and other protected surfaces. Because of their properties, fluoropolymer products are used in a wide range of industries including aerospace, automotive, building/construction, chemical processing, electrical and electronics, semiconductor, and textiles.³

PFOA is one of many chemicals in the perfluoroalkyl acids (PFAAs) family. Among PFAAs, PFOA has the highest production volume in the U.S.; during the reporting year 2002, manufacturers reported that the production volumes were within the range of 6 to 227 metric tons for PFOA and within the range of 227 to 454 metric tons for APFO.⁴ There are both direct and indirect sources of PFOA emissions to the environment. Direct sources result from the manufacture and use of PFOA, while indirect sources in the environment are those where PFOA are present as chemical reaction impurities or where substances may degrade to form PFOA.^{2, 4} As a result, environmental exposures to PFOA are widespread--PFOA has been detected in the blood of over 95% of the U.S. population.^{5, 6}

Due in part to their chemical properties, some PFAAs, including PFOA, can remain in the environment and bioconcentrate in animals,^{3, 7-12} and may take years to be eliminated from the bodies of exposed humans.¹³⁻¹⁷ PFOA has been detected globally, throughout the environment, from polar bears living in Greenland, to giant pandas in China, to albatrosses on the Midway Atoll in the middle of the Pacific Ocean.¹⁸

In 2006, EPA launched a program to work toward the phase-out of PFOA with eight companies voluntarily agreeing to reduce emissions and product content of PFOA and related chemicals by 95% no later than 2010. The program aims to eliminate emissions and product content of these chemicals by 2015.¹⁹ Half of the participating companies met the program's 2010 goal and all eight companies have informed EPA that they are on track to phase out PFOA by the end of 2015.²⁰ Despite emission reduction, because PFOA can remain in the environment and has a long half-life in humans, the chemical will persist in people for many years to come.^{3, 7-17}

The major known sources of human exposure to PFOA include food, breast milk, indoor dust, and water.⁷ A 2008 study identified food consumption as a primary pathway of exposure to PFOA.²¹ PFOA has been detected in human breast milk; therefore, women's exposure to the compound may result in subsequent exposure to their infants.²²⁻²⁶ In studies measuring related chemicals in house dust, PFOA was found to be present in the majority of dust samples examined.²⁷⁻³⁰ In some areas, such as those near industrial facilities that either make or use PFAAs, PFOA has been found in drinking water, groundwater, and surface water.³¹⁻³⁷

The developing fetus may be exposed to PFOA via transplacental transfer as the chemical has been detected in pregnant women and umbilical cord blood.³⁸⁻⁴² In a 2003-2004 population-based study, PFOA was detected in 99-100% of blood samples collected from a representative

population of pregnant and non-pregnant women in the United States,⁴³ and in a separate study, PFOA was detected in 100% of umbilical cord blood samples collected from newborns in Baltimore.³⁸ A 2012 study of child-mother pairs found that concentrations of PFOA tended to be higher in children than their mothers; this difference persisted until about 12 years of age.⁴⁴ However, in a study conducted in Japan, PFOA was detected in samples taken from the mother but not in corresponding cord blood samples.⁴⁵ Although how maternal and fetal PFOA concentrations correlate has yet to be fully documented, the ubiquitous presence of PFOA in blood of women of childbearing age and its presence in umbilical cord blood indicate that fetal exposure may be widespread.^{38, 43, 46}

Developmental Exposure to Perfluorooctanoic Acid and Human Health

Some human studies have found associations between prenatal exposure to PFOA and restricted fetal growth as measured by low birth weight, decreased head circumference, reduced birth length, and smaller abdominal circumference.^{39, 47, 48} However, other studies did not find an association between prenatal PFOA exposure and reduced fetal growth.^{42, 49, 50} In all of these studies, the participants had PFOA levels comparable to those of the general population. The animal literature also includes a mix of positive and negative findings, but the level of exposures used in animal studies are typically much higher than the levels that humans are normally exposed to. Effects such as reduced birth weight, structural defects, delays in postnatal growth and development, increased neonatal mortality, and pregnancy loss have all been associated with prenatal exposure to PFOA or its salts in rodent studies.⁵¹⁻⁵³

The quality and duration of the gestation period is one of the most important predictors of an infant's health and survival.⁵⁴ Low birth weight infants, born either preterm or after experiencing intrauterine growth restriction, experience greater risk of perinatal mortality and both short and long term infant and childhood morbidity.⁵⁵ Health problems related to low birth weight are a leading cause of infant death in the United States; the infant mortality rate for low birth weight infants is about 25 times that for normal weight babies.⁵⁶ Low birth weight infants are more likely to have underdeveloped lungs and breathing problems; intraventricular hemorrhage (bleeding in the brain); learning or behavioral problems later in life; liver problems; polycythemia or anemia; increased risk of infection; inadequate body fat, leading to trouble maintaining a normal body temperature; and feeding problems.⁵⁷

For a subset of low birth weight infants, the low birth weight is due to growth restriction. This condition may cause the fetus to make adaptations, for example, to preserve brain growth, in order to survive adverse intrauterine conditions. Such adaptations can result in a physiologic trade off which can negatively impact other aspects of development, cardiac and renal function, and adult health.⁵⁸ According to the developmental origins of adult health and disease theory, in utero exposure to certain chemical and physical agents, nutrition, stress, and other environmental exposures can alter the programming of fetal cells in ways that can affect disease risk later in life.⁵⁹ Birth weight and measures of growth restriction are used as indicators of these changes and have been associated with adult diseases, including cardiovascular disease, obesity, metabolic disorders, and cancer.⁶⁰

In summary, ubiquitous and on-going exposure among women of childbearing age to PFOA has been documented. There is also evidence that fetal exposure to PFOA be associated with low birth weight and restricted growth that may influence health across the lifespan of exposed individuals.

Objectives

- To answer the question: “Does fetal developmental exposure to perfluorooctanoic acid (PFOA) or its salts affect fetal growth in animals?”

For the purpose of this review, “PFOA or its salts” is defined as PFOA in its uncharged and anionic form. “Fetal developmental exposure” is defined as maternal exposure to PFOA during preconception and/or pregnancy or direct fetal exposure.

- To rate the strength of the experimental animal evidence according to one of the following four statements: 1. Sufficient; 2. Limited; 3. Inadequate; or 4. Evidence of lack of toxicity.

Methods

Criteria for Selecting Studies

The types of non-human studies selected include: (1) studies where experimental female animals were allocated to exposure or control groups and first generation offspring were analyzed for outcome; and (2) studies where experimental animal embryos were allocated to exposure or control groups and analyzed for outcome.

PECO

“PICO” (Participants, Interventions, Comparator, Outcomes) is an aid used to formulate an answerable question in a systematic review. Because we are evaluating environmental exposures, the acronym was modified to PECO and stands for “Participants”, “Exposure,” “Comparator” and “Outcomes.”

Participants: Animals from non-human species that are studied during reproductive/developmental time period (before and/or during pregnancy for females or during development for embryos).

Exposure: One or more oral, subcutaneous or other treatment(s) of any dosage with perfluorooctanoic acid (PFOA), CAS# 335-67-1, or its salts during the time before pregnancy and/or during pregnancy for females or directly to embryos.

Comparators: Experimental animals receiving different doses of PFOA or vehicle-only treatment.

Outcomes: Changes in fetal weight near term (for example, embryonic day 18 for mice and embryonic day 21 for rat); birth weight; and/or other measures of size at term or birth, such as length.

Search Methods

We will employ a variety of mechanisms to identify relevant data, as outlined below. Our search will not be limited by language or publication date.

Electronic Searches

PubMed

To assist in the development of a list of terms relevant to our PubMed search strategy, we will conduct an analysis of the Medical Subject Headings (MeSH), substance, title, and abstract text terms in a non-random group of seven papers that are known to us, that we judge to be relevant to our study question, and which represent different journals and years of publication. This analysis will produce a list of common and unique terms from these papers that we will incorporate into a search strategy that addresses each of the following components:

- (1) **Substance (PFOA):** We will search the MeSH browser at <http://www.nlm.nih.gov/mesh/MBrowser.html> for all of the substance terms identified in the analysis of the seven relevant papers. We will extract heading and entry terms (including historical terms) listed in the MeSH browser for these substances. We will combine these terms as “OR” statements and terms will be searched for as a text term in the title and abstract [tiab] and/or a MeSH term [mh], if appropriate.
- (2) **Experimental animal data:** We will adapt a search filter developed for animal experimental data retrieval from PubMed by Hooijmans et al.⁶¹ The developers of this filter estimate its sensitivity to be very high and the filter has been demonstrated to retrieve seven percent more records compared to using the PubMed limit “Animals.” We will modify the Hooijmans et al search filter to remove the last term “medline [sb].” By excluding this term we will be able to retrieve any papers that are published but not fully indexed at the time of our search.
- (3) **Reproductive/developmental toxicity:** We will search the MeSH browser at <http://www.nlm.nih.gov/mesh/MBrowser.html> for reproductive/developmental toxicity terms identified in the analysis of the seven relevant papers. We will extract all of the heading and entry terms (including historical terms) listed in the MeSH browser for these terms. We will combine these terms as “OR” statements and terms will be searched for as a text term in the title and abstract [tiab] and/or a MeSH term [mh], if appropriate.

Each of these three components will then be combined into a final search filter (see Appendix II).

Web of Science

To develop a Web of Science search filter, we will modify the PubMed search filter. This will consist of removing all MeSH terms and formatting the search terms for the Web of Science database. Specifically, we will perform a topic search in Web of Science, which will search all keywords, titles, and abstracts for search terms, and turn lemmatization off to maximize control of the search.

Toxic Substance Control Act Docket

We will search for unpublished data from primary manufacturers of PFOA by searching the Toxic Substances Control Act (TSCA) docket (<http://www.ntis.gov/products/ots.aspx>) using the following CAS number and each of the synonymous terms: 335-67-1, perfluorooctanoic acid, perfluorooctanoic acid, perfluoro-n-octanoic acid, pentadecafluorooctanoic acid, APFO, perfluorinated octanoic acid, perfluorooctanoate. TSCA requires that manufacturers report positive toxicological findings to USEPA and these reports are maintained in the TSCA docket.⁶²

Searching Other Resources

We will use other methods to find additional studies that are not identified through electronic sampling and may be in the grey literature i.e., technical reports from government agencies or

scientific research groups, working papers from research groups or committees, white papers, preprints, conference proceedings, personal communications, etc.

These methods include:

- Identifying unpublished studies and other unpublished resources
- Hand searching the reference list of all studies that are included after full text review (prior to study author contact, if applicable) and use Web of Science to search for articles that cite the included studies.
- Personal communication

One author (DA) will search a list of toxicological websites (listed in Appendix III) using each of the following terms: PFOA, 335-67-1, perfluorooctanoic acid, perfluorooctanoic acid, perfluoro-n-octanoic acid, pentadecafluorooctanoic acid, APFO, perfluorinated octanoic acid, and perfluorooctanoate.

The results will be reviewed for relevancy as follows:

Include: Any report, health consultation, abstract, protocol or other publication with original animal data about PFOA

Exclude: Any results: (1) having no original data or references in the report, i.e., a list of chemicals, fact sheet, brochure, national health survey, project description, etc.; and/or (2) having no data about PFOA; and/or (3) having no animal data.

We will also contact one or more experts on the topic of PFOA (e.g., Christopher Lau at USEPA) to identify any additional studies or reports. Any duplicates identified through these methods will be excluded and the remaining references will be evaluated according to the inclusion/exclusion criteria described below.

Study Selection Criteria

All search results will be imported or manually entered into EndNote reference management software. We will eliminate duplicate references before we begin evaluating the eligibility of the studies identified.

Two authors (EK, JL) will independently conduct a title and abstract review of the literature search results to determine whether a reference meets the selection criteria for inclusion. The same two authors will then evaluate the entire text of the remaining articles to identify studies meeting the inclusion criteria. Two other co-authors (PJ/PS) will be brought in to settle any discrepancies between the reviewers resulting from each step of the review process. To ensure quality control, a third reviewer (PJ) will also perform title and abstract screening of five percent of the search results and five percent or 5 papers, whichever is greater, of search results eligible for full text review to ensure quality control. The review of selection criteria for inclusion will be performed using a structured form designed and accessed in DistillerSR (Evidence Partners) an online application designed specifically for the screening and data extraction phases of a

systematic review.¹ Justification for inclusion/exclusion will be recorded within DistillerSR to maximize transparency (see Appendix IV for title and abstract inclusion/exclusion form).

Reports of experimental animal studies in any language or date will be eligible for inclusion. All reports that compare experimental animals exposed to one or more oral (gavage or feed), subcutaneous or other doses of PFOA during reproductive or developmental periods (i.e., maternal preconception and/or prenatal exposure for females or during development for embryos) to untreated control experimental animals will be eligible for inclusion.

Citation titles and abstracts will be screened and excluded if one or more of the following criteria are met:

1. Article does not contain original data
2. The study subjects were not animals
3. PFOA was not administered to study subjects
4. PFOA was not administered during reproductive/developmental time period
5. Other reasons (explanation required)

For citations where the database contains no abstract, authors (EK, JL) will attempt to obtain the abstracts from an internet search. Articles for which an abstract remains unavailable will be screened based on titles and PubMed MeSH headings. Any study not excluded based on above criteria will be tagged as “unclear” and included for full-text review.

Citation eligible for full text review will be screened and excluded if one or more of the following criteria are met:

1. Article does not contain original data
2. The study subjects were not animals
3. PFOA was not administered to study subjects
4. PFOA was not administered during reproductive/developmental time period
5. Fetal growth was not measured
6. No fetal growth measurements were taken at birth or near term (approximately gestation day 18 for mouse or day 21 for rat)
7. Control (vehicle-treated group) was not included
8. Other reasons (explanation required)

For articles (including non-English articles) which are not available in the database, three authors (EK, JL, DA) will attempt to obtain articles from a broad internet search (see Appendix IV for full-text inclusion/exclusion form). If resources permit, potentially relevant non-English articles will be translated into English to determine eligibility.

Data Collection

The same two authors (EK, JL) will then independently extract the study characteristics from all of the included articles (see Appendix V for the study characteristics data collection form). These characteristics will be used to evaluate reporting quality, risk of bias and/or to conduct statistical analyses, and were compiled by combining aspects of a variety of checklists and criteria.⁶³⁻⁶⁶ Study

Available at: <http://systematic-review.net/>

researchers will be contacted when additional information is needed. If study researchers do not respond to requests after being contacted through 3 email messages over the course of 1 month or more, review authors will note that attempt to contact study researchers was unsuccessful. The data extraction of study characteristics will be performed using Microsoft Access. Two authors (EK, DA) will independently perform data entry of the raw outcome data using Microsoft Excel.

Risk of Bias Determination

Risk of bias will be assessed using domains from the Cochrane Collaboration's 'Risk of Bias' tool that address selection bias, performance bias, attrition bias, detection bias, and reporting bias.⁶⁷ We have modified terminology and concepts in these domains to make appropriate for animal toxicological studies (see Appendix VI for risk of bias form and Appendix VII for detailed instructions on making a risk of bias determination for review authors).

Informed by empirical data from meta-analyses conducted on pharmacological treatments,⁶⁸ we will also assess funding source and declared conflicts of interest as sources of bias. We will also search for each study in PubMed and note if there has been a retraction of the published article in order to determine if the study may be fraudulent. Two review authors (EK, JL) will independently make risk of bias determinations for each study across all domains and then compare their results. Remaining disputes will be reviewed by 4 co-authors (EK, JL, PS, PJ). If, upon further discussion the 4 co-authors cannot reach agreement, on an appropriate risk of bias determination for a particular domain, the more conservative judgment will be selected (e.g. if one reviewer makes a judgment of 'yes' and the other makes a judgment of 'probably yes', the 'probably yes' judgment will be used). If additional data are acquired from study authors, risk of bias judgments will be modified to reflect the updated study information.

Relevant domains will include:

- Sequence generation
- Allocation concealment
- Blinding of outcome assessment (personnel and outcome assessors)
- Incomplete outcome data
- Selective outcome reporting
- Financial conflict of interest

Publication bias will be addressed by: (1) implementing a comprehensive search of the literature using multiple sources and methods in order to identify all published and unpublished studies that meet the eligibility criteria; and (2) using funnel plot analysis and/or other statistical analyses of the studies included in the systematic review, as appropriate and if possible.

Statistical Analysis

Prior to the analysis, extracted study characteristics from all included articles will be evaluated to assess comparability and determine whether biological heterogeneity is a concern prior to

combination in a meta-analysis. A full list of specified characteristics to be used in this evaluation is outlined in Appendix VIII. Example characteristics include the following:

- Experimental design (randomized or not, multigenerational, etc.)
- Animal model used (species, strain, and genetic background)
- Age of animals (at start of treatment, mating, and/or pregnancy)
- Developmental stage of animals at outcome measurement
- PFOA dose levels, frequency, timing, duration, and exposure route
- Outcome reported (fetal weight, birth weight, or other measurement of fetal growth)

All specified characteristics will be assessed by two reviewers in pairs (EK, JL, PJ) for each included study. These characteristics and their associated heterogeneity concerns will be determined *a priori*, developed through consultation with a panel of experts in the field of PFOA toxicity, toxicological study design, or human/animal toxicity reviews.

For the meta-analysis, we plan on extracting the mean and standard error from each study and utilizing a two-step analysis. In the first step, we will analyze each study separately using a mixed effects model, and obtain an estimate of the dose-response effect (along with standard error). In the second step, we will combine the dose-response effects from each study using a random effects model.

The result will be an estimate of the overall change in outcome for a 1-unit increase in mg/kgBW/day dose, accounting for within- and between-study variability. We will use the programming environment R version 2.13.1 and its standard meta-analysis packages for all statistical analyses.⁶⁹ Consultation with a statistician (SS) will guide the determination of whether this chosen statistical approach is appropriate for the study data available and if further modifications are required prior to performing the meta-analysis.

To test statistical heterogeneity across the study estimates, we will estimate the variance component corresponding to between-study variability, and use a likelihood ratio test for the null hypothesis that the between study variability is absent. A p-value of 0.05 or less will be considered statistically significant. Furthermore, to assess the impact of existing study heterogeneity on the meta-analysis, the I^2 test statistic will be calculated and evaluated, by considering the magnitude/direction of effects, strength of evidence for heterogeneity (e.g., p-value from a chi squared test or a confidence intervals for I^2), and the Cochrane's guide to interpretation as follows:

- 0%-40%: might not be important;
- 30%-60%: may represent moderate heterogeneity;
- 50%-90%: may represent substantial heterogeneity;
- 75%-100%: considerable heterogeneity.

Sensitivity analyses will be performed by examining the effects of excluding studies with particular heterogeneous results as well as performing subgroup analyses based on heterogeneous characteristics identified from the review for comparability across studies. Additionally, a funnel plot of the estimated effects will be used to visually assess the possibility of publication bias.

Quality and Strength of Evidence Ratings

Upon completion of the data collection, risk of bias determination, and data analysis, each of the co-authors will compare the results of the systematic review to the criteria in the Navigation Guide for rating the quality and strength of the non-human evidence (Appendix I). All co-authors will be given explicit directions before rating (see “Navigation Guide Protocol for Rating the Quality and Strength of Human and Non-Human Evidence”). **The instructions to co-authors are provided in a separate document, *Navigation Guide Protocol for Rating the Quality and Strength of the Human and Non-Human Evidence*.**

The initial quality level of experimental animal data will be considered as high, comparable to human randomized controlled trials. This assumption is supported by the level of study control exercised in experimental animal studies and the limited heterogeneity within the study population. This is also consistent with GRADE’s consideration of randomization as a key determinant of the default “high”.⁶³ Furthermore, studies suggest that humans are more sensitive to chemical exposures than animals, strengthening the applicability of findings from experimental animal studies to human health outcomes.⁷⁰

Factors that may decrease the quality level of the body of evidence include:

1. Risk of Bias Across Studies: Study limitations – a substantial risk of bias across body of evidence;
2. Indirectness: Evidence was not directly comparable to the question of interest (i.e., population, exposure, comparator, outcome).
3. Inconsistency: Widely different estimates of effect (heterogeneity or variability in results);
4. Imprecision: Studies had few participants and few events (wide confidence intervals); and
5. Publication Bias: Studies missing from body of evidence, resulting in an overestimate or underestimate of true effects from exposure.

The overall strength of the evidence will be based on a combination of considerations, including:

1. Quality of body of evidence
2. Direction of effect
3. Confidence in effect
4. Other compelling attributes of the data that may influence certainty

Possible ratings for quality of evidence are 0 (no change from ‘high’ quality), -1 (1 level downgrade to ‘moderate’ quality) or -2 (2 level downgrade to ‘low’ quality). Authors who decide to rate down quality need to specify the 1 or 2 criteria most responsible for their decision while documenting all factors that contributed to the final decision to rate down quality. After independently evaluating the quality of the evidence, co-authors will compare their evaluations and any discrepancies between the reviewers’ decisions will be resolved through discussion until consensus is reached. The rationale for each decision on each of the five factors will be recorded.

Subsequent to consensus on the quality of the evidence, the co-authors will rate the strength of evidence. The overall strength of the body of human evidence is based on a combination of four criteria: (1) Quality of body of evidence (i.e., the rating from the previous step); (2) Direction of effect; (3) Confidence in effect; and (4) Other compelling attributes of the data that may influence certainty. The results of rating of the strength of the non-human evidence will then be compared to the criteria specified in the Navigation Guide and described according to one of the following four statements: 1. Sufficient; 2. Limited; 3. Inadequate; or 4. Evidence of lack of toxicity (see Appendix 1).¹ Any discrepancies between the reviewers' decisions will be resolved through discussion. The senior author (TW) will be the ultimate arbiter of the discrepancies that cannot be resolved through consensus among the co-authors.

Next Steps

This protocol describes our approach for the systematic and transparent review of the non-human evidence stream. The protocol for the systematic review of the human evidence will be presented in a separate document “The Impact of Developmental Exposure to Perfluorooctanoic Acid (PFOA) On Fetal Growth: A Systematic Review of the Human Evidence.”

At the conclusion of the systematic reviews of the non-human and human evidence, we will combine these ratings according to the criteria specified in the Navigation Guide.¹ The end result will be one of five possible statements: 1. Known to be toxic to reproductive or developmental health; 2. probably toxic; 3. possibly toxic; 4. unclassifiable; or 5. probably not toxic to reproductive or developmental health.

Additional analyses will be performed to explore alternative approaches⁷¹⁻⁷³ to synthesizing information and data from both non-human (including mechanistic data) and human studies. The ultimate goal of these efforts will be to develop quantitative, health-based estimates for humans to support the development of recommendations for prevention.

PRIMARY INFORMATION

Investigation Guide Criteria for Evaluating Strength of Evidence from Non-Human Evidence

Available evidence usually includes consistent results from well-designed, well-conducted studies, and the conclusion is unlikely to be strongly affected by the results of future studies. For non-human^a evidence positive information has been established through either multiple positive results or a single appropriate study^b in a single species.

Available evidence is sufficient to determine the effects of the exposure, but confidence in the estimate is maintained by such factors as: the number, size, or quality of individual studies, the confidence in the effect, or consistency of findings across individual studies. As more information becomes available, the observed effect could change, and this change may be large enough to alter the conclusion. For non-human evidence the data suggest an effect, but only in a single study; or there are other important limitations in the quality of the body of evidence as noted.

Findings do not permit a conclusion about a toxic effect. The available evidence is insufficient to assess effects of the exposure. Evidence is insufficient because of: the limited number or size of studies, low quality of individual studies, or inconsistency of findings across individual studies. More information may allow an estimation of effects.

Available evidence includes consistent results from well-designed, well-conducted studies, and the conclusion is unlikely to be strongly affected by the results of future studies. For non-human evidence data on an adequate array of endpoints from more than one study with two species showed no adverse effects at doses that were minimally toxic for inducing an adverse effect. Information on pharmacokinetics, mechanisms, or known properties of the chemical class may also strengthen the evidence.^c Adequate studies in at least two species show that the exposure is not toxic. Conclusion is limited to the species, age at exposure, and/or other conditions and levels of exposure as noted.

Quality and strength of evidence of human and non-human evidence streams separately as 'sufficient', 'limited', 'inadequate' or 'evidence of insufficient data'. Findings are combined to produce one of five possible statements about the overall strength of the evidence of a chemical's toxicity. These definitions are adapted from the criteria used by the International Agency for Research on Cancer (IARC) to categorize the carcinogenicity of agents as noted.

^b IARC's criteria for sufficient evidence of carcinogenicity in animals requires multiple positive results (species, studies, sexes). The Navigation Guide integrates USEPA's minimum criteria for animal data for a reproductive or developmental hazard, i.e., data demonstrating an adverse reproductive effect in a single appropriate, well-executed study in a single test species.⁷⁵ The Navigation Guide also incorporates USEPA's "sufficient evidence category" which includes data that "collectively provide enough information to judge whether or not a reproductive hazard exists within the context of effect as well as dose, duration, timing, and route of exposure. This category may include both human and experimental animal evidence".⁷⁵ The USEPA statement for developmental hazards is slightly different but includes the same relevant information regarding dose, duration, timing, etc.⁷⁶

^c Based on minimum data requirements according to USEPA Guidelines for Reproductive Toxicity.⁷⁵

Appendix II. PubMed Search Terms

Search	PubMed
<p>#1</p> <p>Substance terms</p>	<p>335-67-1 [rn] OR perfluorooctanoic acid [nm] OR (perfluorooctanoic acid [tiab] OR perfluorooctanoic acids [tiab]) OR (perfluorooctanoic acid [tiab] OR perfluorooctanoic acids [tiab]) OR (perfluoro-n-octanoic acid [tiab] OR perfluoro-n-octanoic acids [tiab]) OR (pentadecafluorooctanoic acid [tiab] OR pentadecafluorooctanoic acids [tiab]) OR APFO [tiab] OR (perfluorinated [tiab] AND octanoic acid [tiab]) OR (perfluorinated [tiab] AND octanoic acids [tiab]) OR (perfluorooctanoate [tiab] OR perfluorooctanoates [tiab]) OR perfluorooctanoyl chloride [tiab] OR PFOA [tiab] OR (fluorinated telomer alcohol [tiab] OR fluorinated telomer alcohols [tiab]) OR (fluoro-telomer alcohol [tiab] OR fluoro-telomer alcohols [tiab]) OR (fluorocarbon emulsion [tiab] OR fluorocarbon emulsions [tiab]) OR (perfluorocarbon [tiab] OR perfluorocarbons [tiab]) OR (fluorocarbon polymer [tiab] OR fluorocarbon polymers [tiab]) OR (fluorinated polymer [tiab] OR fluorinated polymers [tiab]) OR octanoic acids [mh] OR (octanoic acid [tiab] OR octanoic acids [tiab]) OR caprylates [mh] OR (caprylate [tiab] OR caprylates [tiab]) OR (polyfluoroalkyl [tiab] OR polyfluoroalkyls [tiab] OR polyfluoroalkylated [tiab]) OR PFAA [tiab] OR (perfluoroalkyl chemical [tiab] OR perfluoroalkyl chemicals [tiab]) OR (c8 [tiab] AND perfluorinated [tiab]) OR (fluoropolymer [tiab] OR fluoropolymers [tiab] OR fluoropolymeric [tiab]) OR (fluorosurfactant [tiab] OR fluorosurfactants [tiab]) OR (perfluorochemical [tiab] OR perfluorochemicals [tiab]) OR PFCs [tiab] OR (perfluoroalkyl carboxylate [tiab] OR perfluoroalkyl carboxylates [tiab]) OR (perfluorocarboxylate [tiab] OR perfluorocarboxylates [tiab]) OR PFCA [tiab] OR (perfluorinated carboxylic acid [tiab] OR perfluorinated carboxylic acids [tiab]) OR FC 143 [tiab] OR (pentadecafluorooctanoate [tiab] OR pentadecafluorooctanoates [tiab])</p>
<p>#2</p> <p>Experimental animal terms (modified from Hooijmans, et al., 2010)⁶¹</p>	<p>("animal experimentation"[MeSH Terms] OR "models, animal"[MeSH Terms] OR "invertebrates"[MeSH Terms] OR "Animals"[Mesh:noexp] OR "animal population groups"[MeSH Terms] OR "chordata"[MeSH Terms:noexp] OR "chordata, nonvertebrate"[MeSH Terms] OR "vertebrates"[MeSH Terms:noexp] OR "amphibians"[MeSH Terms] OR "birds"[MeSH Terms] OR "fishes"[MeSH Terms] OR "reptiles"[MeSH Terms] OR "mammals"[MeSH Terms:noexp] OR "primates"[MeSH Terms:noexp] OR "artiodactyla"[MeSH Terms] OR "carnivora"[MeSH Terms] OR "cetacea"[MeSH Terms] OR "chiroptera"[MeSH Terms] OR "elephants"[MeSH Terms] OR "hyraxes"[MeSH Terms] OR "insectivora"[MeSH Terms] OR "lagomorpha"[MeSH Terms] OR "marsupialia"[MeSH Terms] OR "monotremata"[MeSH Terms] OR "perissodactyla"[MeSH Terms] OR "rodentia"[MeSH Terms] OR "scandentia"[MeSH Terms] OR "sirenia"[MeSH Terms] OR "xenarthra"[MeSH Terms] OR "haplorhini"[MeSH Terms:noexp] OR "strepsirhini"[MeSH Terms] OR "platyrrhini"[MeSH Terms] OR "tarsii"[MeSH Terms] OR "catarrhini"[MeSH Terms:noexp] OR "cercopithecidae"[MeSH Terms] OR "hylobatidae"[MeSH Terms] OR "hominidae"[MeSH Terms:noexp] OR "gorilla gorilla"[MeSH Terms] OR</p>

"pan paniscus"[MeSH Terms] OR "pan troglodytes"[MeSH Terms] OR "pongo pygmaeus"[MeSH Terms]) OR (animals[tiab] OR animal[tiab] OR mice[tiab] OR mus[tiab] OR mouse[tiab] OR murine[tiab] OR woodmouse[tiab] OR rats[tiab] OR rat[tiab] OR murinae[tiab] OR muridae[tiab] OR cottonrat[tiab] OR cottonrats[tiab] OR hamster[tiab] OR hamsters[tiab] OR cricetinae[tiab] OR rodentia[tiab] OR rodent[tiab] OR rodents[tiab] OR pigs[tiab] OR pig[tiab] OR swine[tiab] OR swines[tiab] OR piglets[tiab] OR piglet[tiab] OR boar[tiab] OR boars[tiab] OR "sus scrofa"[tiab] OR ferrets[tiab] OR ferret[tiab] OR polecat[tiab] OR polecats[tiab] OR "mustela putorius"[tiab] OR "guinea pigs"[tiab] OR "guinea pig"[tiab] OR cavia[tiab] OR callithrix[tiab] OR marmoset[tiab] OR marmosets[tiab] OR cebuella[tiab] OR hapale[tiab] OR octodon[tiab] OR chinchilla[tiab] OR chinchillas[tiab] OR gerbillinae[tiab] OR gerbil[tiab] OR gerbils[tiab] OR jird[tiab] OR jirds[tiab] OR merione[tiab] OR meriones[tiab] OR rabbits[tiab] OR rabbit[tiab] OR hares[tiab] OR hare[tiab] OR diptera[tiab] OR flies[tiab] OR fly[tiab] OR dipteral[tiab] OR drosophila[tiab] OR drosophilidae[tiab] OR cats[tiab] OR cat[tiab] OR carus[tiab] OR felis[tiab] OR nematoda[tiab] OR nematode[tiab] OR nematodes[tiab] OR sipunculida[tiab] OR dogs[tiab] OR dog[tiab] OR canine[tiab] OR canines[tiab] OR canis[tiab] OR sheep[tiab] OR sheeps[tiab] OR mouflon[tiab] OR mouflons[tiab] OR ovis[tiab] OR goats[tiab] OR goat[tiab] OR capra[tiab] OR capras[tiab] OR rupicapra[tiab] OR chamois[tiab] OR haplorhini[tiab] OR monkey[tiab] OR monkeys[tiab] OR anthropoidea[tiab] OR anthropoids[tiab] OR saguinus[tiab] OR tamarin[tiab] OR tamarins[tiab] OR leontopithecus[tiab] OR hominidae[tiab] OR ape[tiab] OR apes[tiab] OR pan[tiab] OR paniscus[tiab] OR "pan paniscus"[tiab] OR bonobo[tiab] OR bonobos[tiab] OR troglodytes[tiab] OR "pan troglodytes"[tiab] OR gibbon[tiab] OR gibbons[tiab] OR siamang[tiab] OR siamangs[tiab] OR nomascus[tiab] OR symphalangus[tiab] OR chimpanzee[tiab] OR chimpanzees[tiab] OR prosimians[tiab] OR "bush baby"[tiab] OR prosimian[tiab] OR bush babies[tiab] OR galagos[tiab] OR galago[tiab] OR pongidae[tiab] OR gorilla[tiab] OR gorillas[tiab] OR pongo[tiab] OR pygmaeus[tiab] OR "pongo pygmaeus"[tiab] OR orangutans[tiab] OR lemur[tiab] OR lemurs[tiab] OR lemoridae[tiab] OR horse[tiab] OR horses[tiab] OR equus[tiab] OR cow[tiab] OR calf[tiab] OR bull[tiab] OR chicken[tiab] OR chickens[tiab] OR gallus[tiab] OR quail[tiab] OR bird[tiab] OR birds[tiab] OR quails[tiab] OR poultry[tiab] OR poultries[tiab] OR fowl[tiab] OR fowls[tiab] OR reptile[tiab] OR reptilia[tiab] OR reptiles[tiab] OR snakes[tiab] OR snake[tiab] OR lizard[tiab] OR lizards[tiab] OR alligator[tiab] OR alligators[tiab] OR crocodile[tiab] OR crocodiles[tiab] OR turtle[tiab] OR turtles[tiab] OR amphibian[tiab] OR amphibians[tiab] OR amphibia[tiab] OR frog[tiab] OR frogs[tiab] OR bombina[tiab] OR salientia[tiab] OR toad[tiab] OR toads[tiab] OR "epidalea calamita"[tiab] OR salamander[tiab] OR salamanders[tiab] OR eel[tiab] OR eels[tiab] OR fish[tiab] OR fishes[tiab] OR pisces[tiab] OR catfish[tiab] OR catfishes[tiab] OR siluriformes[tiab] OR arius[tiab] OR heteropneustes[tiab] OR sheatfish[tiab] OR perch[tiab] OR perches[tiab] OR percidae[tiab] OR perca[tiab] OR trout[tiab] OR trouts[tiab] OR char[tiab] OR chars[tiab] OR salvelinus[tiab] OR "fathead minnow"[tiab] OR minnow[tiab] OR cyprinidae[tiab] OR carps[tiab] OR carp[tiab] OR zebrafish[tiab] OR zebrafishes[tiab] OR goldfish[tiab] OR goldfishes[tiab] OR guppy[tiab]

	<p>OR guppies[Tiab] OR chub[Tiab] OR chubs[Tiab] OR tinca[Tiab] OR barbels[Tiab] OR barbuis[Tiab] OR pimephales[Tiab] OR promelas[Tiab] OR "poecilia reticulata"[Tiab] OR mullet[Tiab] OR mullets[Tiab] OR seahorse[Tiab] OR seahorses[Tiab] OR mugil curema[Tiab] OR atlantic cod[Tiab] OR shark[Tiab] OR sharks[Tiab] OR catshark[Tiab] OR anguilla[Tiab] OR salmonid[Tiab] OR salmonids[Tiab] OR whitefish[Tiab] OR whitefishes[Tiab] OR salmon[Tiab] OR salmons[Tiab] OR sole[Tiab] OR solea[Tiab] OR "sea lamprey"[Tiab] OR lamprey[Tiab] OR lampreys[Tiab] OR pumpkinseed[Tiab] OR sunfish[Tiab] OR sunfishes[Tiab] OR tilapia[Tiab] OR tilapias[Tiab] OR turbot[Tiab] OR turbots[Tiab] OR flatfish[Tiab] OR flatfishes[Tiab] OR scuridae[Tiab] OR squirrel[Tiab] OR squirrels[Tiab] OR chipmunk[Tiab] OR chipmunks[Tiab] OR suslik[Tiab] OR susliks[Tiab] OR vole[Tiab] OR voles[Tiab] OR lemming[Tiab] OR lemmings[Tiab] OR muskrat[Tiab] OR muskrats[Tiab] OR lemmus[Tiab] OR otter[Tiab] OR otters[Tiab] OR marten[Tiab] OR martens[Tiab] OR martes[Tiab] OR weasel[Tiab] OR badger[Tiab] OR badgers[Tiab] OR ermine[Tiab] OR mink[Tiab] OR minks[Tiab] OR sable[Tiab] OR sables[Tiab] OR gulo[Tiab] OR gulos[Tiab] OR wolverine[Tiab] OR wolverines[Tiab] OR mustela[Tiab] OR llama[Tiab] OR llamas[Tiab] OR alpaca[Tiab] OR alpacas[Tiab] OR camelid[Tiab] OR camelids[Tiab] OR guanaco[Tiab] OR guanacos[Tiab] OR chiroptera[Tiab] OR chiropteras[Tiab] OR bat[Tiab] OR bats[Tiab] OR fox[Tiab] OR foxes[Tiab] OR iguana[Tiab] OR iguanas[Tiab] OR xenopus laevis[Tiab] OR parakeet[Tiab] OR parakeets[Tiab] OR parrot[Tiab] OR parrots[Tiab] OR donkey[Tiab] OR donkeys[Tiab] OR mule[Tiab] OR mules[Tiab] OR zebra[Tiab] OR zebras[Tiab] OR shrew[Tiab] OR shrews[Tiab] OR bison[Tiab] OR bisons[Tiab] OR buffalo[Tiab] OR buffaloes[Tiab] OR deer[Tiab] OR deers[Tiab] OR bear[Tiab] OR bears[Tiab] OR panda[Tiab] OR pandas[Tiab] OR "wild hog"[Tiab] OR "wild boar"[Tiab] OR fitchew[Tiab] OR fitch[Tiab] OR beaver[Tiab] OR beavers[Tiab] OR jerboa[Tiab] OR jerboas[Tiab] OR capybara[Tiab] OR capybaras[Tiab])</p>
<p>#3 Reproductive/developmental toxicity terms</p>	<p>developmental biology [mh] OR developmental biology [tiab] OR embryonic and fetal development [mh] OR (embryonic [tiab] OR embryonically [tiab]) OR fetal development [tiab] OR growth and development [mh] OR growth and development [subheading] OR (development [tiab] OR developmental [tiab] OR developmentally [tiab]) OR embryology [mh] OR embryology [tiab] OR ecotoxicology [mh] OR ecotoxicology [tiab] OR ecology [mh] OR (ecology [tiab] OR ecological [tiab] OR ecologically [tiab]) OR toxicology [mh] OR (toxicology [tiab] OR toxicological [tiab] OR toxicologically [tiab]) OR toxicogenetics [mh] OR (toxicogenetic [tiab] OR toxicogenetics [tiab]) OR growth [mh] OR growth [tiab] OR environmental pollutants [mh] OR (environmental pollutant [tiab] OR environmental pollutants [tiab]) OR body weight [mh] OR (weight [tiab] OR weights [tiab] OR weighed [tiab]) OR embryo loss [mh] OR (embryo loss [tiab] OR embryo losses [tiab]) OR fetal resorption [mh] OR (fetal resorption [tiab] OR fetal resorptions [tiab]) OR gestational age [mh] OR (gestational age [tiab] OR gestational ages [tiab]) OR litter size [mh] OR (litter size [tiab] OR litter sizes [tiab]) OR endocrine disruptors [mh] OR (endocrine disruptor [tiab] OR endocrine disruptors [tiab] OR endocrine disruption [tiab]) OR reproduction [mh] OR reproduction [tiab] OR toxicity [subheading] OR toxicity [tiab] OR (toxic [tiab] OR toxics</p>

	[tiab])
#4	#1 AND #2 AND #3

Search	Web of Science
#1 Substance terms	TS=((perfluorooctanoic acid OR perfluorooctanoic acids) OR (perfluorooctanoic acid OR perfluorooctanoic acids) OR (perfluoro-n-octanoic acid OR perfluoro-n-octanoic acids) OR (pentadecafluorooctanoic acid OR pentadecafluorooctanoic acids) OR APFO OR (perfluorinated AND octanoic acid) OR (perfluorinated AND octanoic acids) OR (perfluorooctanoate OR perfluorooctanoates) OR perfluorooctanoyl chloride OR PFOA OR (fluorinated telomer alcohol OR fluorinated telomer alcohols) OR (fluoro-telomer alcohol OR fluoro-telomer alcohols) OR (fluorocarbon emulsion OR fluorocarbon emulsions) OR (perfluorocarbon OR perfluorocarbons) OR (fluorocarbon polymer OR fluorocarbon polymers) OR (fluorinated polymer OR fluorinated polymers) OR (octanoic acid OR octanoic acids) OR (caprylate OR caprylates) OR (polyfluoroalkyl OR polyfluoroalkyls OR polyfluoroalkylated) OR PFAA OR (perfluoroalkyl chemical OR perfluoroalkyl chemicals) OR (c8 AND perfluorinated) OR (fluoropolymer OR fluoropolymers OR fluoropolymeric) OR (fluorosurfactant OR fluorosurfactants) OR (perfluorochemical OR perfluorochemicals) OR PFCs OR (perfluoroalkyl carboxylate OR perfluoroalkyl carboxylates) OR (perfluorocarboxylate OR perfluorocarboxylates) OR PFCA OR (perfluorinated carboxylic acid OR perfluorinated carboxylic acids) OR FC 143 OR (pentadecafluorooctanoate OR pentadecafluorooctanoates))
#2 Experimental animal terms (modified from Hooijmans, et al., 2010) ⁶¹	TS=(animals OR animal OR mice OR mus OR mouse OR murine OR woodmouse OR rats OR rat OR murinae OR muridae OR cottonrat OR cottonrats OR hamster OR hamsters OR cricetinae OR rodentia OR rodent OR rodents OR pigs OR pig OR swine OR swines OR piglets OR piglet OR boar OR boars OR "sus scrofa" OR ferrets OR ferret OR polecat OR polecats OR "mustela putorius" OR "guinea pigs" OR "guinea pig" OR cavia OR callithrix OR marmoset OR marmosets OR cebuella OR hapale OR octodon OR chinchilla OR chinchillas OR gerbillinae OR gerbil OR gerbils OR jird OR jirds OR merione OR meriones OR rabbits OR rabbit OR hares OR hare OR diptera OR

	<p>flies OR fly OR dipteral OR drosphila OR drosophilidae OR cats OR cat OR carus OR felis OR nematoda OR nematode OR nematodes OR sipunculida OR dogs OR dog OR canine OR canines OR canis OR sheep OR sheeps OR mouflon OR mouflons OR ovis OR goats OR goat OR capra OR capras OR rupicapra OR chamois OR haplorhini OR monkey OR monkeys OR anthropoidea OR anthropoids OR saguinus OR tamarin OR tamarins OR leontopithecus OR hominidae OR ape OR apes OR pan OR paniscus OR "pan paniscus" OR bonobo OR bonobos OR troglodytes OR "pan troglodytes" OR gibbon OR gibbons OR siamang OR siamangs OR nomascus OR symphalangus OR chimpanzee OR chimpanzees OR prosimians OR "bush baby" OR prosimian OR bush babies OR galagos OR galago OR pongidae OR gorilla OR gorillas OR pongo OR pygmaeus OR "pongo pygmaeus" OR orangutans OR lemur OR lemurs OR lemuriidae OR horse OR horses OR equus OR cow OR calf OR bull OR chicken OR chickens OR gallus OR quail OR bird OR birds OR quails OR poultry OR poultries OR fowl OR fowls OR reptile OR reptilia OR reptiles OR snakes OR snake OR lizard OR lizards OR alligator OR alligators OR crocodile OR crocodiles OR turtle OR turtles OR amphibian OR amphibians OR amphibia OR frog OR frogs OR bombina OR salientia OR toad OR toads OR "epidalea calamita" OR salamander OR salamanders OR eel OR eels OR fish OR fishes OR pisces OR catfish OR catfishes OR siluriformes OR arius OR heteropneustes OR sheatfish OR perch OR perches OR percidae OR perca OR trout OR trouts OR char OR chars OR salvelinus OR "fathead minnow" OR minnow OR cyprinidae OR carps OR carp OR zebrafish OR zebrafishes OR goldfish OR goldfishes OR guppy OR guppies OR chub OR chubs OR tinca OR barbels OR barbus OR pimephales OR promelas OR "poecilia reticulata" OR mullet OR mullets OR seahorse OR seahorses OR mugil curema OR atlantic cod OR shark OR sharks OR catshark OR anguilla OR salmonid OR salmonids OR whitefish OR whitefishes OR salmon OR salmons OR sole OR solea OR "sea lamprey" OR lamprey OR lampreys OR pumpkinseed OR sunfish OR sunfishes OR tilapia OR tilapias OR turbot OR turbot OR flatfish OR flatfishes OR sciuridae OR squirrel OR squirrels OR chipmunk OR chipmunks OR suslik OR susliks OR vole OR voles OR lemming OR lemmings OR muskrat OR muskrats OR lemmus OR otter OR otters OR marten OR martens OR martes OR weasel OR badger OR badgers OR ermine OR mink OR minks OR sable OR sables OR gulo OR gulos OR wolverine OR wolverines OR mustela OR llama OR llamas OR alpaca OR alpacas OR camelid OR camelids OR guanaco OR guanacos OR chiroptera OR chiropteras</p>
--	---

	OR bat OR bats OR fox OR foxes OR iguana OR iguanas OR xenopus laevis OR parakeet OR parakeets OR parrot OR parrots OR donkey OR donkeys OR mule OR mules OR zebra OR zebras OR shrew OR shrews OR bison OR bisons OR buffalo OR buffaloes OR deer OR deers OR bear OR bears OR panda OR pandas OR "wild hog" OR "wild boar" OR fitchew OR fitch OR beaver OR beavers OR jerboa OR jerboas OR capybara OR capybaras)
#3 Reproductive/developmental toxicity terms	TS=(developmental biology OR (embryonic OR embryonically) OR (development OR developmental OR developmentally) OR embryology OR ecotoxicology OR (ecology OR ecological OR ecologically) OR (toxicology OR toxicological OR toxicologically) OR (toxicogenetic OR toxicogenetics) OR growth OR (environmental pollutant OR environmental pollutants) OR (weight OR weights OR weighed) OR (embryo loss OR embryo losses) OR (fetal resorption OR fetal resorptions) OR (gestational age OR gestational ages) OR (litter size OR litter sizes) OR (endocrine disruptor OR endocrine disruptors OR endocrine disruption) OR reproduction OR (toxicity OR toxic OR toxics))
#4	#1 AND #2 AND #3

Appendix III. List of Toxicological Websites To Be Searched

- ATSDR Interaction Profiles <http://www.atsdr.cdc.gov/interactionprofiles/>
- ATSDR Toxicological Profiles <http://www.atsdr.cdc.gov/toxproz.html>
- CalEPA Office of Environmental Health Hazard Assessment <http://www.oehha.ca.gov/risk.html>
- Chem ID <http://chem.sis.nlm.nih.gov/chemidplus/>
- Chemfinder www.chemfinder.com/
- Chemspider <http://www.chemspider.com>
- Chemical Carcinogenesis Research Information System
<http://www.nlm.nih.gov/pubs/factsheets/ccrisfs.html>
- DART <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?DARTETIC>
- EPA Acute Exposure Guideline Levels <http://www.epa.gov/oppt/aegl/chemlist.htm>
- EPA IRIS edocket and official records; IRIS Hotline 202-566-1676
- EPA IRIS internet www.epa.gov/iris
- EPA NEPIS <http://www.epa.gov/nepis/>
- EPA NSCEP <http://www.epa.gov/ncepihom/>
- EPA Science Inventory <http://www.epa.gov/gateway/science/>
- EPA Substance Registry System <http://www.epa.gov/srs/>
- Environmental Mutagen Information Center <http://library.wlu.ca/resource/emic>
- European Chemicals Agency http://echa.europa.eu/home_en.asp
- GENETOX <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?GENETOX>
- Health Canada First Priority List Assessments http://www.hcsc.gc.ca/hecs_sesc/exsd/psl1.htm
- Health Canada Second Priority List Assessments http://www.hcsc.gc.ca/hecs_sesc/exsd/psl2.htm
- Hazardous Substances Data Bank <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB>
- IARC <http://monographs.iarc.fr/htdig/search.html>
- ILSI <http://www.ilsi.org/>
- IPCS INCHEM <http://www.inchem.org/>
- ITER http://iter.ctcnet.net/publicurl/pub_search_list.cfm
- NIOSHTIC 2 <http://www2.cdc.gov/nioshtic2/Nioshtic2.htm>
- US National Toxicology Program Management Status Report
http://ntpserver.niehs.nih.gov/main_pages/NTP_ALL_STDY_PG.html
- US National Toxicology Program Results and Status Search
http://ntpserver.niehs.nih.gov/main_pages/NTP_ALL_STDY_PG.html
- US National Toxicology Program Report on Carcinogens
<http://ntpserver.niehs.nih.gov/NewHomeRoc/AboutRoC.html>
- TERA <http://www.tera.org/>
- Toxicology Data Network <http://toxnet.nlm.nih.gov/>
- Toxline <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE>
- RTECS Toxcenter <http://www.cdc.gov/niosh/rtecs/default.html>
- WHO assessments – CICADS, EHC <http://www.who.int/ipcs/assessment/en/>
- USEPA Health and Environmental Studies Online <http://hero.epa.gov/>
- TSCA Test Submissions: <http://www.ntis.gov/products/ots.aspx>

Appendix IV: Exclusion Criteria Screening Forms

Title and Abstract Screening Form

1. Exclude citation from further review because (check only one reason):
 - No original data
Definition: Article does not contain original data (e.g. review article, systematic review, editorial, etc.)
 - Non-animal subjects
Definition: Study subjects were not animals (e.g. humans, in vitro model, etc.)
 - No intentional PFOA exposure
Definition: PFOA was not administered to study subjects
 - No developmental PFOA exposure
Definition: PFOA was not administered during reproductive/developmental time period (before and/or during pregnancy for females, or directly to fetuses)
 - Exclude for other reason
Please explain “other” reason:
2. Please check if abstract is:
 - Non-English
 - Not available
3. Comments:

Full-Text Screening Form

1. Exclude citation from further review because (check only one reason):
 - Study subjects are humans only
 - No animal study subjects (in vitro model, etc.)
 - PFOA was not administered to study subjects
 - Female animals or fetuses were not study subjects
 - No developmental PFOA exposure during reproductive/developmental time period (before and/or during pregnancy for females, or directly to fetuses)
 - Fetal growth was not measured
 - No fetal growth measurements were taken at birth or near term (approximately gestation day 18 for mouse or day 21 for rat)
 - Control (vehicle-treated group) was not included
 - Exclude for other reason
Please explain “other” reason:
2. Could this be a duplicate study?

- Yes
Please explain:

- No

3. Please check if article is:

- Non-English
- Not available
- In English and cannot determine eligibility

4. Check if:

- Include

5. Comments:

DRAFT

Appendix V: Data Collection Form²

Fields are free-form except where choices are shown in italics

Refid:

Reviewer:

- *Erica*
- *Juleen*
- *Other (identify yourself)*

Publication year:

Authors' declared conflicts of interest:

- *None declared*
- *Declared*

If declared, provide details:

Study funding source:

- *Government grant*
- *Industry funded*
- *Nonprofit organization grant*
- *Other (details below)*

Study funding source details:

What are the study objectives?:

Animal transportation mode to facility:

- *Truck*
- *Plane*
- *Other (details below)*

² The source criteria checklists for extraction terms include: gold standard publication checklist (GSPC)[65]; ARRIVE guidelines (ARRIVE)[66]; Cochrane Handbook for Systematic Reviews of Interventions data collection checklist (Cochrane)[64]; GRADE criteria for randomized control trials (GRADE).[63]

- *Not stated*

Animal transportation mode details:

Animal transportation shipping time:

Length of quarantine period after arrival at facility:

Origin and source of animals (commercial breeder, stock #, etc):

Site(s) of experiment (city, state, country):

Are animals immune compromised?:

- *Yes*
- *No*
- *Not stated*
- *Other (details below)*

Immune compromised details:

Microbiological status of facility:

- *SPF*
- *Germ-free*
- *Gnotobiotic*
- *Conventional*
- *Other (details below)*
- *Not stated*

Microbiological status of facility details:

Cage ventilation/setup:

- *Static, filter-top cages*
- *Ventilated caging system*
- *Other (details below)*
- *Not stated*

Cage ventilation/setup details:

Animal room temperature:

Animal room humidity:

Animal room lighting cycle:

- *12-hour light cycle*
- *Other (details below)*
- *Not stated*

Animal room lighting cycle details:

Number of animals per cage during experiment:

Was cage enrichment provided?:

- *Yes*
- *No*
- *Not stated*
- *Other (details below)*

Number of animals and cage enrichment details:

Nutrition (feed brand, ingredients, batch #, etc):

Feeding regimen:

- *Ad libitum*
- *Timed feedings*
- *Other (details below)*
- *Not stated*

Feeding regimen details:

Water provisions:

- *Ad libitum*
- *Timed administration*

- *Limited quantity offered*
- *Other (details below)*
- *Not stated*

Water provision details:

Water system type:

- *Water bottle*
- *Continuous automated watering system*
- *Other (details below)*
- *Not stated*

Water system type details:

Was drinking water purified or pre-treated?:

- *Yes*
- *No*
- *Not stated*
- *Other (details below)*

Water purification/pre-treated details:

Description of breeding methods for dams in experiment:

Have animals undergone previous procedures?:

- *Yes*
- *No*
- *Not stated*
- *Other (details below)*

Previous procedures details:

Euthanasia method for dams, if measuring fetal outcomes:

- *Carbon dioxide*

- *Cervical dislocation or decapitation*
- *Pharmaceutical intervention*
- *Other (details below)*
- *Not stated*

Euthanasia method details:

Description of animal wellbeing throughout study:

Miscellaneous comments by reviewer regarding study info:

Type of study:

- *Preconception*
- *Pregnancy*
- *Both preconception and pregnancy*
- *Multigenerational*
- *Fetal exposure only*
- *Other (details below)*

Type of study details:

Study design:

- *Randomized animal study*
- *Other (details below)*

Study design details:

Animal species:

- *Mouse*
- *Rat*
- *Chicken*
- *Other (details below)*

Animal species details:

Animal strain:

- *CD-1*
- *Sprague-Dawley*
- *Other (details below)*
- *Not stated*

Animal strain details (including exact genetic code, if available):

Animal genetic background:

- *Inbred*
- *Genetically modified (transgenic/knockout)*
- *Other (details below)*
- *Not stated*

Animal genetic background details:

Reason why animal model chosen (characteristics/limitations):

Experimental generation for outcome measurement, if multigenerational study:

Age (at arrival, treatment, mating, pregnancy, etc.):

Developmental stage (at arrival):

- *Virgin*
- *Pregnant*
- *Other (details below)*

Developmental stage (at arrival) details:

Developmental stage (at treatment):

- *Virgin*
- *Pregnant*
- *Other (details below)*

Developmental stage (at treatment) details:

Total number of intervention groups:

Total number of control groups:

Number of animals allocated to each group:

If a power calc was done, was the sample size of the experiment sufficient?:

- *Yes*
- *No*

Doses of PFOA administered:

Frequency of PFOA administration:

- *Daily*
- *Weekly*
- *Other (details below)*

Frequency of PFOA administration details:

How frequently animals weighed to determine dose:

- *Daily*
- *Weekly*
- *More than weekly*
- *Other (details below)*
- *Not stated*

How frequently animals weighed to determine dose details:

Duration of PFOA treatment:

- *Time window prior to pregnancy*
- *Time window prior to and during pregnancy*
- *Time window during pregnancy*
- *Other (details below)*

Duration of PFOA treatment details:

Administration route:

- *Gavage*
- *Other (details below)*

Administration route details:

Chemical name:

- *APFO or PFOA, ammonium salt*
- *Other (details below)*

Chemical name details:

Details on chemical mixture preparation:

Other chemical information:

Outcomes measured:

Method of fetal growth measurement:

- *Weight*
- *Length*
- *Other (details below)*

Method of fetal growth measurement details:

Developmental stage for progeny at outcome measurement:

- *Near term (E18 for mice or E21 for rat)*
- *At birth*
- *Other (details below)*

Developmental stage for progeny outcome measurement details:

Unit of measurement (for weight, etc.):

- *Grams*
- *Millimeters*
- *Other (details below)*

Unit of measurement (for weight, etc.) details:

If applicable, how progeny weighed (individual or litter, etc.):

Sex (for animals where outcome measured):

- *Males only*
- *Females only*
- *Males and females*
- *Other (details below)*

Sex (for animals where outcome measured) details:

Were dead animals included in outcome assessment?:

- *Yes*
- *No*

How many animals/litters alive vs. dead?:

How was parturition monitored?:

- *Continuously*
- *AM check*
- *Other (details below)*
- *Not stated*

Parturition monitoring details:

Number animals analyzed (for intervention and outcome):

How were data reported (mean, median, raw data, etc.)?:

Growth measurement data for each group (ie, outcome):

How growth measurement data were reported (table, figure, etc):

How was precision reported (standard error, CI, etc.)?:

- *Standard error*
- *Standard deviation*

- *Confidence intervals*
- *Other (details below)*
- *Not stated*

How precision reported details:

Precision estimates:

How precision estimates were reported (table, figure, etc):

Miscellaneous comments by reviewer regarding data analysis:

DRAFT

Appendix VI: Risk of Bias Form

Modified Cochrane Risk of Bias Tool⁶⁷ - Please answer 'yes', 'probably yes', 'probably no', 'no', or 'not applicable' and provide details/justification:

1. Was the allocation sequence adequately generated?

- Criteria for our judgment of 'yes': There is direct evidence that animals were randomly allocated to experimental groups using an appropriate randomization scheme
- Criteria for our judgment of 'probably yes': There is indirect evidence that animals were randomly allocated to experimental groups using an appropriate randomization scheme, as described by the criteria for a judgment of 'yes'
- Criteria for our judgment of 'probably no': There is indirect evidence that animals were allocated to experimental groups using a non-random component, as described by the criteria for a judgment of 'no'
- Criteria for our judgment of 'no': There is direct evidence that animals were allocated to experimental groups using a non-random component
- Criteria for our judgment of 'not applicable':
There is direct evidence that the risk of bias domain is not applicable to the study

2. Was allocation adequately concealed?

- Criteria for our judgment of 'yes': There is direct evidence that investigators could not foresee which animals were allocated to the various experimental groups
- Criteria for our judgment of 'probably yes': There is indirect evidence that investigators could not foresee which animals were allocated to the various experimental groups, as described by the criteria for a judgment of 'yes'
- Criteria for our judgment of 'probably no': There is indirect evidence that investigators performed or were aware of allocation of animals to various experimental group, as described by the criteria for a judgment of 'no'
- Criteria for our judgment of 'no': There is direct evidence that investigators performed or were aware of allocation of animals to various experimental groups

- Criteria for our judgment of ‘not applicable’: There is direct evidence that the risk of bias domain is not applicable to the study
3. Was knowledge of allocated interventions adequately prevented?
- Criteria for our judgment of ‘yes’: There is direct evidence that the study was adequately blinded
 - Criteria for our judgment of ‘probably yes’: There is indirect evidence that the study was adequately blinded, as described by the criteria for a judgment of ‘yes’
 - Criteria for our judgment of ‘probably no’: There is indirect evidence for lack of adequate blinding, as described by the criteria for a judgment of ‘no’
 - Criteria for our judgment of ‘no’: There is direct evidence for lack of adequate blinding
 - Criteria for our judgment of ‘not applicable’: There is direct evidence that the risk of bias domain is not applicable to the study
4. Were incomplete outcome data adequately addressed?
- Criteria for our judgment of ‘yes’: There is direct evidence that incomplete outcome data were adequately addressed
 - Criteria for our judgment of ‘probably yes’: There is indirect evidence that incomplete outcome data were adequately addressed, as described by the criteria for a judgment of ‘yes’
 - Criteria for our judgment of ‘probably no’: There is indirect evidence that incomplete outcome data were not adequately addressed, as described by the criteria for a judgment of ‘no’
 - Criteria for our judgment of ‘no’: There is direct evidence that incomplete outcome data were not adequately addressed
 - Criteria for our judgment of ‘not applicable’: There is direct evidence that the risk of bias domain is not applicable to the study
5. Were study reports free of selective outcome reporting?
- Criteria for our judgment of ‘yes’: There is direct evidence that all of the study’s outcomes of interest outlined in the protocol, methods, abstract, and/or introduction have been reported in a pre-specified way

- Criteria for our judgment of ‘probably yes’: There is indirect evidence that all of the study’s outcomes of interest outlined in the protocol, methods, abstract, and/or introduction have been reported in a pre-specified way, as described by the criteria for a judgment of ‘yes’
 - Criteria for our judgment of ‘probably no’: There is indirect evidence that all of the study’s outcomes of interest outlined in the protocol, methods, abstract, and/or introduction have not been reported in a pre-specified way, as described by the criteria for a judgment of ‘no’
 - Criteria for our judgment of ‘no’: There is direct evidence that all of the study’s outcomes of interest outlined in the protocol, methods, abstract, and/or introduction have not been reported in a pre-specified way
 - Criteria for our judgment of ‘not applicable’: There is direct evidence that the risk of bias domain does not apply to the study
6. Was study free of other problems regarding risk of bias?
- Criteria for our judgment of ‘yes’: There is direct evidence that the study appears to be free of other sources of bias
 - Criteria for our judgment of ‘probably yes’: There is indirect evidence that the study appears to be free of other sources of bias, as described by the criteria for a judgment of ‘yes’
 - Criteria for our judgment of ‘probably no’: There is indirect evidence that there is at least one important risk of bias, as described by the criteria for a judgment of ‘no’
 - Criteria for our judgment of ‘no’: There is direct evidence that there is at least one important risk of bias
 - Criteria for our judgment of ‘not applicable’: There is direct evidence that the risk of bias domain is not applicable to the study
7. Was the study free of support from a company, study author, or other entity having a financial interest in any of the treatments studied?
- Criteria for our judgment of ‘yes’: There is direct evidence that the study was free of support from a company, study author, or other entity having a financial interest in the study
 - Criteria for our judgment of ‘probably yes’: There is indirect evidence that the study was free of support from a company, study author, or other entity

having a financial interest in the study, as described by the criteria for a judgment of 'yes'

- Criteria for our judgment of 'probably no': There is indirect evidence that the study received support from a company, study author, or other entity having a financial interest in the study, as described by the criteria for a judgment of 'no'
- Criteria for our judgment of 'no': There is direct evidence that the study received support from a company, study author, or other entity having a financial interest in the study
- Criteria for our judgment of 'not applicable': There is direct evidence that the risk of bias domain is not applicable to the study

Miscellaneous fields:

Number and reasons animals excluded from experiment:

Number and reasons for premature death:

Number and descriptions of included animals with peculiarities:

Power analysis after adjustment for excluded animals?:

- *Yes*
- *No*

Miscellaneous comments by reviewer regarding risk of bias:

Appendix VII: Instructions for Making Risk of Bias Determinations

1. SEQUENCE GENERATION

Was the allocation sequence adequately generated?

Criteria for a judgment of 'YES' (i.e. low risk of bias):

The investigators describe a random component in the sequence generation process such as:

- Referring to a random number table;
- Using a computer random number generator;
- Coin tossing;
- Shuffling cards or envelopes;
- Throwing dice;
- Drawing of lots.

Criteria for the judgment of 'PROBABLY YES' (i.e. probably low risk of bias):

There is insufficient information about the sequence generation process to permit a judgment of 'YES', but there is indirect evidence that suggests the sequence generation process was random, as described by the criteria for a judgment of 'YES'.

Criteria for the judgment of 'PROBABLY NO' (i.e. probably high risk of bias):

There is insufficient information about the sequence generation process to permit a judgment of 'NO', but there is indirect evidence that suggests a non-random component in the sequence generation process, as described by the criteria for a judgment of 'NO'.

Criteria for the judgment of 'NO' (i.e. high risk of bias):

The investigators describe a non-random component in the sequence generation process or that a random component was not used. Usually, the description would involve some systematic, non-random approach, for example:

- Sequence generated by date of birth;
- Sequence generated by some rule based on date (or day) of arrival at facility;
- Sequence generated by some rule based on record number.

Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgment or some method of non-random categorization of animals, for example:

- Allocation by judgment of the investigator;
- Allocation by availability of the intervention.

Criteria for the judgment of 'NOT APPLICABLE' (risk of bias domain is not applicable to study):

There is evidence that sequence generation is not an element of study design capable of introducing risk of bias in the study.

2. ALLOCATION CONCEALMENT

Was allocation adequately concealed?

Criteria for a judgment of 'YES' (i.e. low risk of bias):

Investigators could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation:

- Sequentially numbered treatment containers of identical appearance to control; or
- Sequentially numbered prepared route of administration (e.g., pre-prepared water dosed with chemical) of identical appearance; or
- Sequentially numbered, opaque, sealed envelopes.

Criteria for the judgment of 'PROBABLY YES' (i.e. probably low risk of bias):

There is insufficient information about allocation concealment to permit a judgment of 'YES', but there is indirect evidence that suggests the allocation was adequately concealed, as described by the criteria for a judgment of 'YES'.

Criteria for the judgment of 'PROBABLY NO' (i.e. probably high risk of bias):

There is insufficient information about allocation concealment to permit a judgment of 'NO', but there is indirect evidence that suggests the allocation was not adequately concealed, as described by the criteria for a judgment of 'NO'.

Criteria for the judgment of 'NO' (i.e. high risk of bias):

Investigators handling experimental animals could possibly foresee assignments and thus introduce selection bias, such as allocation based on:

- Using an open random allocation schedule (e.g. a list of random numbers); or
- Assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or non-opaque or not sequentially numbered); or
- Alternation or rotation; or
- Non-random and known criteria, such as date of birth; or
- Record number; or
- Any other explicitly unconcealed procedure.

Criteria for the judgment of 'NOT APPLICABLE' (risk of bias domain is not applicable to study):

There is evidence that allocation concealment is not an element of study design capable of introducing risk of bias in the study.

3. BLINDING OF PERSONNEL AND OUTCOME ASSESSORS

Was knowledge of the allocated interventions adequately prevented during the study?

Criteria for a judgment of 'YES' (i.e. low risk of bias):

Any one of the following:

- No blinding, but the review authors judge that the outcome and the outcome measurement are not likely to be influenced by lack of blinding; or
- Blinding of key study personnel ensured, and unlikely that the blinding could have been broken; or
- Some key study personnel were not blinded, but outcome assessment was blinded and the non-blinding of others unlikely to introduce bias.

Criteria for the judgment of 'PROBABLY YES' (i.e. probably low risk of bias):

There is insufficient information about blinding to permit a judgment of 'YES', but there is indirect evidence that suggests the study was adequately blinded, as described by the criteria for a judgment of 'YES'.

Criteria for the judgment of 'PROBABLY NO' (i.e. probably high risk of bias):

There is insufficient information about blinding to permit a judgment of 'NO', but there is indirect evidence that suggests the study was not adequately blinded, as described by the criteria for a judgment of 'NO'.

Criteria for the judgment of 'NO' (i.e. high risk of bias):

Any one of the following:

- No blinding or incomplete blinding, and the outcome or outcome measurement is likely to be influenced by lack of blinding; or
- Blinding of key study personnel attempted, but likely that the blinding could have been broken; or
- Some key study personnel were not blinded, and the non-blinding of others likely to introduce bias.

Criteria for the judgment of 'NOT APPLICABLE' (risk of bias domain is not applicable to study):

There is evidence that blinding is not an element of study design capable of introducing risk of bias in the study.

4. INCOMPLETE OUTCOME DATA

Were incomplete outcome data adequately addressed?

Criteria for a judgment of 'YES' (i.e. low risk of bias):

The number of animals assessed for outcome of interest is reported and data is provided indicating adequate follow up of all treated animals. Additional information provided by

authors should be considered when making risk of bias judgments about incomplete outcome data. Additionally, any one of the following:

- No missing outcome data; or
- Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias); or
- Missing outcome data is provided and is balanced in numbers across intervention groups, with similar reasons for missing data across groups; or
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a biologically relevant impact on the intervention effect estimate; or
- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a biologically relevant impact on observed effect size; or
- Missing data have been imputed using appropriate statistical methods.

Criteria for the judgment of 'PROBABLY YES' (i.e. probably low risk of bias):

There is insufficient information about incomplete outcome data to permit a judgment of 'YES', but there is indirect evidence that suggests incomplete outcome data were adequately addressed, as described by the criteria for a judgment of 'YES'.

Criteria for the judgment of 'PROBABLY NO' (i.e. probably high risk of bias):

There is insufficient information about incomplete outcome data to permit a judgment of 'NO', but there is indirect evidence that suggests incomplete outcome data was not adequately addressed, as described by the criteria for a judgment of 'NO'.

Criteria for the judgment of 'NO' (i.e. high risk of bias):

The number of animals allocated not reported and no data is provided to indicate that there was adequate follow up of all treated animals. Additionally, any one of the following:

- Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; or
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce biologically relevant bias in intervention effect estimate; or
- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce biologically relevant bias in observed effect size; or
- 'As-treated' analysis done with substantial departure of the intervention received from that assigned at randomization; or
- Potentially inappropriate application of simple imputation.

Criteria for the judgment of 'NOT APPLICABLE' (risk of bias domain is not applicable to study):

There is evidence that incomplete outcome data is not capable of introducing risk of bias in the study.

5. SELECTIVE OUTCOME REPORTING

Are reports of the study free of suggestion of selective outcome reporting?

Criteria for a judgment of 'YES' (i.e. low risk of bias):

All of the study's pre-specified (primary and secondary) outcomes outlined in the methods, abstract, and/or introduction that are of interest in the review have been reported in the pre-specified way, including the number of animals analyzed for outcomes of interest. Additional information provided by authors should be considered when making risk of bias judgments for selective outcome reporting.

Criteria for the judgment of 'PROBABLY YES' (i.e. probably low risk of bias):

There is insufficient information about selective outcome reporting to permit a judgment of 'YES', but there is indirect evidence that suggests the study was free of selective reporting, as described by the criteria for a judgment of 'YES'.

Criteria for the judgment of 'PROBABLY NO' (i.e. probably high risk of bias):

There is insufficient information about selective outcome reporting to permit a judgment of 'NO', but there is indirect evidence that suggests the study was not free of selective reporting, as described by the criteria for a judgment of 'NO'.

Criteria for the judgment of 'NO' (i.e. high risk of bias):

The study was not free of selective reporting. The following should be considered:

- Authors did not report numbers analyzed for outcomes of interest; or
- Not all of the study's pre-specified primary outcomes (as outlined in the protocol, title, abstract, and/or introduction) that are of interest in the review have been reported; or
- One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); or
- One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); or
- One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; or
- The study report fails to include results for a key outcome that would be expected to have been reported for such a study.

Criteria for the judgment of 'NOT APPLICABLE' (risk of bias domain is not applicable to study):

There is evidence that selective outcome reporting is not capable of introducing risk of bias in the study.

6. OTHER POTENTIAL THREATS TO VALIDITY

Was the study apparently free of other problems that could put it at a risk of bias?

Criteria for a judgment of 'YES' (i.e. low risk of bias):

The study appears to be free of other sources of bias.

Criteria for the judgment of 'PROBABLY YES' (i.e. probably low risk of bias):

There is insufficient information to permit a judgment of 'YES', but there is indirect evidence that suggests the study was free of other threats to validity, as described by the criteria for a judgment of 'YES'.

Criteria for the judgment of 'PROBABLY NO' (i.e. probably high risk of bias):

There is insufficient information to permit a judgment of 'NO', but there is indirect evidence that suggests the study was not free of other threats to validity, as described by the criteria for a judgment of 'NO'.

Criteria for the judgment of 'NO' (i.e. high risk of bias):

There is at least one important risk of bias. For example, the study:

- Had a potential source of bias related to the specific study design used;
- Stopped early due to some data-dependent process (including a formal-stopping rule);
- Had extreme baseline imbalance (improper control group);
- Has been claimed to have been fraudulent;
- The conduct of the study is affected by interim results (e.g. recruiting additional animals from a subgroup showing more benefit);
- There is deviation from the study protocol in a way that does not reflect typical practice (e.g. post hoc stepping-up of doses to exaggerated levels);
- There is pre-randomization administration of an intervention that could enhance or diminish the effect of a subsequent, randomized, intervention; inappropriate administration of an intervention (or co-intervention);
- Occurrence of 'null bias' due to interventions being insufficiently well delivered or overly wide inclusion criteria for animals;
- An insensitive instrument is used to measure outcomes (which can lead to under-estimation of both beneficial and harmful effects);
- Selective reporting of subgroups;
- Had some other problem.

Criteria for the judgment of 'NOT APPLICABLE' (risk of bias domain is not applicable to study):

There is evidence that other potential threats to validity are not capable of introducing risk of bias in the study.

7. CONFLICT OF INTEREST

Was the study free of support from a company, study author, or other entity having a financial interest in any of the treatments studied?

Criteria for a judgment of 'YES' (i.e. low risk of bias):

The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study. A conflict of interest statement is provided to indicate the authors have no financial interest and there is evidence of the entities not having a financial interest. Examples of this evidence include the following:

- Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations;
- Chemicals or other treatment used in study were purchased from a supplier;
- Company affiliated staff are not mentioned in the acknowledgements section;
- Authors were not employees of a company with a financial interest in the outcome of the study;
- Company with a financial interest in the outcome of the study was not involved in the design, conduct, analysis, or reporting of the study and authors had complete access to the data;
- Study authors make a claim denying conflicts of interest;
- Study authors are unaffiliated with companies with financial interest, and there is no reason to believe a conflict of interest exists;
- All study authors are affiliated with a government agency (are prohibited from involvement in projects for which there is a conflict of interest or an appearance of conflict of interest).

Criteria for the judgment of 'PROBABLY YES' (i.e. probably low risk of bias):

There is insufficient information to permit a judgment of 'YES', for example there is no conflict of interest statement denying financial interests, but there is evidence that suggests the study was free of support from a company, study author, or other entity having a financial interest in the outcome of the study, as described by the criteria for a judgment of 'YES'.

Criteria for the judgment of 'PROBABLY NO' (i.e. probably high risk of bias):

There is insufficient information to permit a judgment of 'NO', but there is indirect evidence that suggests the study was not free of support from a company, study author, or other entity having a financial interest in the outcome of the study, as described by the criteria for a judgment of 'NO'.

Criteria for the judgment of 'NO' (i.e. high risk of bias):

The study received support from a company, study author, or other entity having a financial interest in the outcome of the study. Examples of support include:

- Research funds;
- Writing services;
- Author/staff from study was employee or otherwise affiliated with company with financial interest;
- Company limited author access to the data;
- Company was involved in the design, conduct, analysis, or reporting of the study;
- Study authors claim a conflict of interest.

Criteria for the judgment of 'NOT APPLICABLE' (risk of bias domain is not applicable to study):

There is evidence that conflicts of interest are not capable of introducing risk of bias in the study.

Appendix VIII: Characteristics to Assess Biological Heterogeneity

Characteristic	Indication of major heterogeneity concern	Exclusion from meta-analysis (or requiring separate analyses)
Study source	<ul style="list-style-type: none"> • If study has not gone through formal peer-review process 	<ul style="list-style-type: none"> • Data from any study where the primary methods, data, and analysis has not undergone a formal peer-review process
Experimental design (randomized or not, multigenerational, etc.)	<ul style="list-style-type: none"> • Study does not randomize treatment allocation • Study is multigenerational 	<ul style="list-style-type: none"> • Studies designed to be nonrandomized or studies utilizing questionable randomization methods for allocating animals. • Multigenerational study where F₁ offspring outcome data not measured
Animal model used (species, strain, and genetic background)	<ul style="list-style-type: none"> • Animal model of questionable relevance to humans 	<ul style="list-style-type: none"> • Non-mammalian animal models • Genetic knock-out animals • Mammalian animals will be analyzed separately by species prior to combining
Age of animals (at arrival, start of treatment, mating, and pregnancy)	<ul style="list-style-type: none"> • Extreme difference in length of treatment to PFOA prior to mating and pregnancy • Extreme difference in age (very young or very old) of animal when treated during mating and pregnancy 	<ul style="list-style-type: none"> • Mammalian dams treated during the first 3 weeks of life (i.e., prior to weaning) • Mammalian dams treated for more than 9 weeks prior to pregnancy
Developmental stage of animals at outcome measurement	<ul style="list-style-type: none"> • Fetal growth measured near term versus measured at birth • Time point of measuring near term or definition of time of birth 	<ul style="list-style-type: none"> • Offspring with fetal growth measured near term will be analyzed separately from those measured at birth • Offspring considered near term whose measurement are more than two days from expected birth • Offspring considered at birth whose measurements are more than two days after birth where parturition is monitored <8 hour daily intervals or more than one day after birth where parturition is monitored >8 hour daily intervals or not stated.
PFOA dose levels, frequency, timing, duration, and exposure route	<ul style="list-style-type: none"> • Extreme differences in range of tested PFOA doses • Extreme differences in frequency, timing, and duration of tested PFOA doses • Different routes of exposure to PFOA 	<ul style="list-style-type: none"> • Dams with tested doses considerably higher or lower than the dose ranges of all other studies (defined as when one or more of the study's doses is higher than the highest dose of all other studies by more than the median of the individual dose differences between consecutive doses for all other studies or when one or

		<p>more of the study's doses is lower than the lowest dose of all other studies by more than the median of individual dose differences between consecutive doses for all other studies (or zero, if the difference is negative)), after exclusion by all other characteristics and stratification by outcome and exposure route.</p> <ul style="list-style-type: none"> • Dams where the frequency, timing, and duration of dosing is considerably different (defined as daily doses given for less than 1/6 gestation time consecutively at any time point during pregnancy, or doses given less than daily where the total number of doses given is less than 1/6 the gestation time, at any point during pregnancy). • Dams with inhalation, gavage, dietary, and drinking water exposure will each be analyzed separately. Dams from fetuses treated directly (i.e., direct injection exposure) will not be combined with data from dams exposed during pregnancy.
<p>Outcome reported (fetal weight, birth weight, or other measurement of fetal growth)</p>	<ul style="list-style-type: none"> • Offspring with different fetal growth measurement outcomes reported 	<ul style="list-style-type: none"> • Offspring with weight outcomes will be analyzed separately from those with length (or other) outcomes.

ABOUT THE ARTICLE

Declarations of Interest: No competing interests reported

Sources of Support:

Internal: UCSF - none

External: UCSF through USEPA Contract

Published Notes

ACKNOWLEDGEMENTS

Kristina Thayer, Andy Rooney, and Abee Boyles in the Office of Health Assessment and Translation, National Toxicology Program for assistance with developing risk of bias evaluation criteria and meta-analysis methodologies.

Christopher Lau in the National Health and Environmental Effects Research Laboratory, Office of Research and Development, US Environmental Protection Agency for serving as PFOA subject matter expert.

Lauren Zeise at the California Environmental Protection Agency for assistance with developing risk of bias and evidence evaluation methodologies.

Lisa Bero in the Department of Clinical Pharmacy, University of California, San Francisco for assistance with developing risk of bias and evidence evaluation methodologies.

Kate Guyton in the National Center for Environmental Assessment Washington Division, National Center for Environmental Assessment, Office of Research and Development, US Environmental Protection Agency for assistance with identifying relevant toxicological databases.

Malcolm Macleod in the Experimental Neuroscience Centre for Clinical Brain Sciences, University of Edinburgh for assistance in developing animal-specific systematic review methodology.

Tara Horvath in the Prevention and Public Health Group, and Co-Managing Editor of the Cochrane Review Group on HIV/AIDS at UCSF for training and assistance with developing our search strategy.

Text from *America's Children and the Environment 3rd Edition*⁷⁷ "Perfluorochemicals (PFCs)" topic was used to support the background section of this protocol.

References

1. Woodruff TJ, Sutton P. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health Aff (Millwood)*. May 2011;30(5):931-937.
2. Prevedouros K, Cousins IT, Buck RC, Korzeniowski SH. Sources, fate and transport of perfluorocarboxylates. *Environ Sci Technol*. Jan 1 2006;40(1):32-44.
3. CDC. Fourth National Report on Human Exposure to Environmental Chemicals: Atlanta, GA: Centers for Disease Control and Prevention; 2009.
4. U.S. Environmental Protection Agency. Long-Chain Perfluorinated Chemicals (PFCs) Action Plan. http://www.epa.gov/oppt/existingchemicals/pubs/pfcs_action_plan1230_09.pdf. Accessed 3 February, 2013.
5. ATSDR. Toxicological Profile for Perfluoroalkyls. (Draft for Public Comment). In: Services USDoHaH, ed; 2009.
6. Kato K, Wong LY, Jia LT, Kuklennyik Z, Calafat AM. Trends in exposure to polyfluoroalkyl chemicals in the U.S. Population: 1999-2008. *Environ Sci Technol*. Oct 1 2011;45(19):8037-8045.
7. Fromme H, Tittlemier SA, Volkel W, Wilhelm M, Twardella D. Perfluorinated compounds--exposure assessment for the general population in Western countries. *Int J Hyg Environ Health*. May 2009;212(3):239-270.
8. Kelly BC, Ikonomidou MG, Blair JD, et al. Perfluoroalkyl contaminants in an Arctic marine food web: trophic magnification and wildlife exposure. *Environ Sci Technol*. Jun 1 2009;43(11):4037-4043.
9. Lau C, Anitole K, Hodes C, Lai D, Pfahles-Hutchens A, Seed J. Perfluoroalkyl acids: a review of monitoring and toxicological findings. *Toxicol Sci*. Oct 2007;99(2):366-394.
10. Martin JW, Mabury SA, Solomon KR, Muir DC. Bioconcentration and tissue distribution of perfluorinated acids in rainbow trout (*Oncorhynchus mykiss*). *Environ Toxicol Chem*. Jan 2003;22(1):196-204.
11. Conder JM, Hoke RA, De Wolf W, Russell MH, Buck RC. Are PFCAs bioaccumulative? A critical review and comparison with regulatory criteria and persistent lipophilic compounds. *Environ Sci Technol*. Feb 15 2008;42(4):995-1003.
12. Kelly BC, Ikonomidou MG, Blair JD, Morin AE, Gobas FA. Food web-specific biomagnification of persistent organic pollutants. *Science*. Jul 13 2007;317(5835):236-239.
13. Bartell SM, Calafat AM, Lyu C, Kato K, Ryan PB, Steenland K. Rate of decline in serum PFOA concentrations after granular activated carbon filtration at two public water systems in Ohio and West Virginia. *Environ Health Perspect*. Feb 2010;118(2):222-228.
14. Brede E, Wilhelm M, Goen T, et al. Two-year follow-up biomonitoring pilot study of residents' and controls' PFC plasma levels after PFOA reduction in public water system in Arnsberg, Germany. *Int J Hyg Environ Health*. Jun 2010;213(3):217-223.
15. Harada K, Inoue K, Morikawa A, Yoshinaga T, Saito N, Koizumi A. Renal clearance of perfluorooctane sulfonate and perfluorooctanoate in humans and their species-specific excretion. *Environ Res*. Oct 2005;99(2):253-261.
16. Olsen GW, Burriss JM, Ehresman DJ, et al. Half-life of serum elimination of perfluorooctanesulfonate, perfluorohexanesulfonate, and perfluorooctanoate in retired fluorochemical production workers. *Environ Health Perspect*. Sep 2007;115(9):1298-1305.

17. Seals R, Bartell SM, Steenland K. Accumulation and clearance of perfluorooctanoic acid (PFOA) in current and former residents of an exposed community. *Environ Health Perspect.* Jan 2011;119(1):119-124.
18. Betts KS. Perfluoroalkyl acids: what is the evidence telling us? *Environ Health Perspect.* May 2007;115(5):A250-256.
19. U.S. Environmental Protection Agency. News Release: EPA Announces Substantial Decrease of PFOA.
<http://yosemite.epa.gov/opa/admpress.nsf/68b5f2d54f3eefd28525701500517bf/8f9dbdd044050f71852573e50064439f!OpenDocument>. Accessed 3 February, 2013.
20. U.S. Environmental Protection Agency. News Release: Industry Progressing in Voluntary Effort to Reduce Toxic Chemicals / Companies have successfully developed over 150 alternatives to long-chain perfluorinated chemicals.
<http://yosemite.epa.gov/opa/admpress.nsf/79c090e81f0578738525781f0043619b/335d29e8d5e40864852579a0005f3760!OpenDocument>. Accessed 11 February, 2013.
21. Trudel D, Horowitz L, Wormuth M, Scheringer M, Cousins IT, Hungerbuhler K. Estimating consumer exposure to PFOS and PFOA. *Risk Anal.* Apr 2008;28(2):251-269.
22. Karrman A, Ericson I, van Bavel B, et al. Exposure of perfluorinated chemicals through lactation: levels of matched human milk and serum and a temporal trend, 1996-2004, in Sweden. *Environ Health Perspect.* Feb 2007;115(2):226-230.
23. Llorca M, Farre M, Pico Y, Teijon ML, Alvarez JG, Barcelo D. Infant exposure of perfluorinated compounds: levels in breast milk and commercial baby food. *Environ Int.* Aug 2010;36(6):584-592.
24. Tao L, Kannan K, Wong CM, Arcaro KF, Butenhoff JL. Perfluorinated compounds in human milk from Massachusetts, U.S.A. *Environ Sci Technol.* Apr 15 2008;42(8):3096-3101.
25. Thomsen C, Haug LS, Stigum H, Froshaug M, Broadwell SL, Becher G. Changes in concentrations of perfluorinated compounds, polybrominated diphenyl ethers, and polychlorinated biphenyls in Norwegian breast-milk during twelve months of lactation. *Environ Sci Technol.* Dec 15 2010;44(24):9550-9556.
26. Volkel W, Genzel-Boroviczeny O, Demmelmair H, et al. Perfluorooctane sulphonate (PFOS) and perfluorooctanoic acid (PFOA) in human breast milk: results of a pilot study. *Int J Hyg Environ Health.* Jul 2008;211(3-4):440-446.
27. Bjorklund JA, Thuresson K, De Wit CA. Perfluoroalkyl compounds (PFCs) in indoor dust: concentrations, human exposure estimates, and sources. *Environ Sci Technol.* Apr 1 2009;43(7):2276-2281.
28. Kato K, Calafat AM, Needham LL. Polyfluoroalkyl chemicals in house dust. *Environ Res.* Jul 2009;109(5):518-523.
29. Kubwabo C, Stewart B, Zhu J, Marro L. Occurrence of perfluorosulfonates and other perfluorochemicals in dust from selected homes in the city of Ottawa, Canada. *J Environ Monit.* Nov 2005;7(11):1074-1078.
30. Strynar MJ, Lindstrom AB. Perfluorinated compounds in house dust from Ohio and North Carolina, USA. *Environ Sci Technol.* May 15 2008;42(10):3751-3756.
31. Konwick BJ, Tomy GT, Ismail N, et al. Concentrations and patterns of perfluoroalkyl acids in Georgia, USA surface waters near and distant to a major use source. *Environ Toxicol Chem.* Oct 2008;27(10):2011-2018.
32. Moody CA, Hebert GN, Strauss SH, Field JA. Occurrence and persistence of perfluorooctanesulfonate and other perfluorinated surfactants in groundwater at a fire-training area at Wurtsmith Air Force Base, Michigan, USA. *J Environ Monit.* Apr 2003;5(2):341-345.

33. Post GB, Louis JB, Cooper KR, Boros-Russo BJ, Lippincott RL. Occurrence and potential significance of perfluorooctanoic acid (PFOA) detected in New Jersey public drinking water systems. *Environ Sci Technol*. Jun 15 2009;43(12):4547-4554.
34. Sinclair E, Mayack DT, Roblee K, Yamashita N, Kannan K. Occurrence of perfluoroalkyl surfactants in water, fish, and birds from New York State. *Arch Environ Contam Toxicol*. Apr 2006;50(3):398-410.
35. Skutlarek D, Exner M, Farber H. Perfluorinated surfactants in surface and drinking waters. *Environ Sci Pollut Res Int*. Sep 2006;13(5):299-307.
36. Steenland K, Jin C, MacNeil J, et al. Predictors of PFOA levels in a community surrounding a chemical plant. *Environ Health Perspect*. Jul 2009;117(7):1083-1088.
37. Shin HM, Vieira VM, Ryan PB, et al. Environmental Fate and Transport Modeling for Perfluorooctanoic Acid Emitted from the Washington Works Facility in West Virginia. *Environ Sci Technol*. Jan 12 2011.
38. Apelberg BJ, Goldman LR, Calafat AM, et al. Determinants of fetal exposure to polyfluoroalkyl compounds in Baltimore, Maryland. *Environ Sci Technol*. Jun 1 2007;41(11):3891-3897.
39. Fei C, McLaughlin JK, Tarone RE, Olsen J. Perfluorinated chemicals and fetal growth: a study within the Danish National Birth Cohort. *Environ Health Perspect*. Nov 2007;115(11):1677-1682.
40. Loccisano A, Campbell JL Jr, Butenhoff J, Andersen M, Clewell HJ 3rd. Evaluation of placental and lactational pharmacokinetics of PFOA and PFOS in the pregnant, lactating, fetal and neonatal rat using a physiologically based pharmacokinetic model. *Reprod Toxicol*. Jul 2012;33(4):468-490.
41. Midasch O, Drexler H, Hart N, Beckmann MW, Angerer J. Transplacental exposure of neonates to perfluorooctanesulfonate and perfluorooctanoate: a pilot study. *Int Arch Occup Environ Health*. Jul 2007;80(7):643-648.
42. Monroy R, Morrison K, Teo K, et al. Serum levels of perfluoroalkyl compounds in human maternal and umbilical cord blood samples. *Environ Res*. Sep 2008;108(1):56-62.
43. Woodruff TJ, Zota AR, Schwartz JM. Environmental chemicals in pregnant women in the United States: NHANES 2003-2004. *Environ Health Perspect*. Jun 2011;119(6):878-885.
44. Mondal D, Lopez-Espinosa MJ, Armstrong B, Stein CR, Fletcher T. Relationships of perfluorooctanoate and perfluorooctane sulfonate serum concentrations between mother-child pairs in a population with perfluorooctanoate exposure from drinking water. *Environ Health Perspect*. May 2012;120(5):752-757.
45. Inoue K, Okada F, Ito R, et al. Perfluorooctane sulfonate (PFOS) and related perfluorinated compounds in human maternal and cord blood samples: assessment of PFOS exposure in a susceptible population during pregnancy. *Environ Health Perspect*. Aug 2004;112(11):1204-1207.
46. Calafat AM, Kuklenyik Z, Reidy JA, Caudill SP, Tully JS, Needham LL. Serum concentrations of 11 polyfluoroalkyl compounds in the u.s. population: data from the national health and nutrition examination survey (NHANES). *Environ Sci Technol*. Apr 1 2007;41(7):2237-2242.
47. Apelberg BJ, Witter FR, Herbstman JB, et al. Cord serum concentrations of perfluorooctane sulfonate (PFOS) and perfluorooctanoate (PFOA) in relation to weight and size at birth. *Environ Health Perspect*. Nov 2007;115(11):1670-1676.
48. Fei C, McLaughlin JK, Tarone RE, Olsen J. Fetal growth indicators and perfluorinated chemicals: a study in the Danish National Birth Cohort. *Am J Epidemiol*. Jul 1 2008;168(1):66-72.

49. Hamm MP, Cherry NM, Chan E, Martin JW, Burstyn I. Maternal exposure to perfluorinated acids and fetal growth. *J Expo Sci Environ Epidemiol*. Nov 2010;20(7):589-597.
50. Washino N, Saijo Y, Sasaki S, et al. Correlations between prenatal exposure to perfluorinated chemicals and reduced fetal growth. *Environ Health Perspect*. Apr 2009;117(4):660-667.
51. Hines EP, White SS, Stanko JP, Gibbs-Flournoy EA, Lau C, Fenton SE. Phenotypic dichotomy following developmental exposure to perfluorooctanoic acid (PFOA) in female CD-1 mice: Low doses induce elevated serum leptin and insulin, and overweight in mid-life. *Mol Cell Endocrinol*. May 25 2009;304(1-2):97-105.
52. Lau C, Butenhoff JL, Rogers JM. The developmental toxicity of perfluoroalkyl acids and their derivatives. *Toxicol Appl Pharmacol*. Jul 15 2004;198(2):231-241.
53. Lau C, Thibodeaux JR, Hanson RG, et al. Effects of perfluorooctanoic acid exposure during pregnancy in the mouse. *Toxicol Sci*. Apr 2006;90(2):510-518.
54. Institute of Medicine. *Preterm Birth: Causes, Consequences, and Prevention*. Washington, DC: The National Academies Press; 2007.
55. Goldenberg RL, Culhane JF. Low birth weight in the United States. *Am J Clin Nutr*. Feb 2007;85(2):584S-590S.
56. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2005 period linked birth/infant death data set. *National Vital Statistics Reports*. 2008;57(2).
57. JAMA. JAMA patient page. Low birth weight. *JAMA*. Jan 9 2002;287(2):270.
58. Cosmi E, Fanelli T, Visentin S, Trevisanuto D, Zanardo V. Consequences in infants that were intrauterine growth restricted. *J Pregnancy*. 2011;2011:364381.
59. Newbold R, Heindel J. Developmental Exposures and Implications for Early and Latent Disease. In: Woodruff T, ed. *Environmental impacts on reproductive health and fertility*. New York: Cambridge University Press, Cambridge; 2010:92-102.
60. Rinaudo PF, Lamb J. Fetal origins of perinatal morbidity and/or adult disease. *Semin Reprod Med*. Sep 2008;26(5):436-445.
61. Hooijmans CR, Tillema A, Leenaars M, Ritskes-Hoitinga M. Enhancing search efficiency by means of a search filter for finding all studies on animal experimentation in PubMed. *Lab Anim*. Jul 2010;44(3):170-175.
62. U.S. Environmental Protection Agency. Toxic Substance Control Act (TSCA) Section 8(e) Notices. <http://www.epa.gov/opptintr/tsca8e/>. Accessed 6 February, 2013.
63. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol*. Apr 2011;64(4):407-415.
64. Higgins JPT, Deeks JJ (editors). Chapter 7: Selecting studies and collecting data. In: *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*, (Higgins JPT, Green S, eds). www.cochrane-handbook.org. Accessed 7 February, 2013.
65. Hooijmans CR, Leenaars M, Ritskes-Hoitinga M. A gold standard publication checklist to improve the quality of animal studies, to fully integrate the Three Rs, and to make systematic reviews more feasible. *Altern Lab Anim*. May 2010;38(2):167-182.
66. Killkenny C, Browne W, Cuthill IC, Emerson M, Altman DG. Animal research: reporting in vivo experiments: the ARRIVE guidelines. *Br J Pharmacol*. Aug 2010;160(7):1577-1579.
67. Higgins JPT, Altman DJ (editors). Chapter 8: Assessing risk of bias in included studies. In: *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*, (Higgins JPT, Green S, eds). www.cochrane-handbook.org. Accessed 7 February, 2013.
68. Roseman M, Milette K, Bero LA, et al. Reporting of conflicts of interest in meta-analyses of trials of pharmacological treatments. Mar 9 2011; *JAMA* 305(10):1008-1017.

69. R Development Core Team. R: A language and environment for statistical computing. <http://www.R-project.org/>. Accessed 13 February, 2013.
70. Kimmel CA, Holson JF, Hogue CJ, Carlo GL. *Reliability of experimental studies for predicting hazards to human development*. NCTR technical report for experiment 6015. Jefferson, AR: National Center for Toxicological Research; 1984.
71. DuMouchel W, Harris J. Bayes method for combining the results of cancer studies in humans and other species. *J Am Stat Assoc*. 1983;78(382):293-308.
72. Jones DR, Peters JL, Rushton L, Sutton AJ, Abrams KR. Interspecies extrapolation in environmental exposure standard setting: A Bayesian synthesis approach. *Regul Toxicol Pharmacol*. Apr 2009;53(3):217-225.
73. Peters JL, Rushton L, Sutton AJ, Jones DR, Abrams KR, Mugglestone MA. Bayesian methods for the cross-design synthesis of epidemiological and toxicological evidence. *J Royal Statistical Society*. 2005;54(1):159-172.
74. IARC. Preamble to the IARC Monographs.: Lyon, France: International Agency for Research on Cancer; 2006.
75. US EPA. Guidelines for Reproductive Toxicity Risk Assessment. Fed Reg 61:56274-56322. In: U.S. Environmental Protection Agency, ed; 1996.
76. US EPA. Guidelines for Developmental Toxicity Risk Assessment. Fed Reg 56:63798-63826. In: U.S. Environmental Protection Agency, ed; 1991.
77. U.S. Environmental Protection Agency. America's Children and the Environment, 3rd Edition. http://www.epa.gov/ace/publications/ACE3_2013.pdf. Accessed 19 February 2013.