



**Program on Reproductive
Health and the Environment**
1330 Broadway Street, Suite 1135
Oakland, CA 94612
prhe@obgyn.ucsf.edu
(510) 350-1240

January 22, 2014

MEMO TO: OEHHA staff

FROM: Tracey Woodruff, PhD, MPH (UCSF)
Paula Johnson, PhD, MPH (UCSF)
Patrice Sutton, MPH (UCSF)
Erica Koustas, PhD (ORISE Fellow at USEPA)

SUBJECT: DART Data Tabulations

We are part of a group of academic and governmental scientific collaborators who have developed and demonstrated proof of concept of an innovative methodology for systematic and transparent review of environmental health science called the "Navigation Guide" [1-6]. Our recommendations for summarizing data from epidemiological and toxicological studies are based on our experience in applying the methodology in 3 cases studies thus far. We are commenting on the example tables that were presented to the Developmental and Reproductive Toxicant Identification Committee for discussion on December 4, 2013 and posted on the OEHHA website for public comment.

Our overall recommendations are that Cal-EPA should a) abstract more detail on individual studies under evaluation and b) conduct a rigorous and transparent assessment of individual study quality/bias that is uniformly and transparently applied to each study. Specific comments are listed below.

We are including some resources that we think will be helpful, including risk of bias tools that we have developed and applied, which are based on methodology for clinical and healthcare systematic review. These tools are provided as an example of how OEHHA could incorporate aspects of systematic review into hazard assessments. We are currently further developing these tools for our subsequent case studies, and the updated versions will be available on our website once our protocols are finalized. Additionally, we note that the National Toxicology Program (NTP) has collaborative efforts underway to incorporate systematic review methodology in its assessments, and NTP's developer of software tools for systematic review will be holding a demonstration webinar on January 28 (see link in below list).

In addition to the specific comments listed below, we have included:

1. UCSF/PRHE's risk of bias tool for epidemiology studies (separate attachment)
2. UCSF/PRHE's risk of bias tool for toxicology studies (separate attachment)

3. Link to UCSF-PRHE's webpage on the Navigation Guide methodology for systematic review that includes protocols containing the risk of bias tools: <http://prhe.ucsf.edu/prhe/navigationguide.html>
4. Link to January 28th DRAGON webinar by ICF International: <http://www.icfi.com/events/webinars/2014/01/DRAGON-a-suite-of-tools-for-systematic-literature-review-jan-28>
5. Copy of a recent review (separate attachment): "Instruments for Assessing Risk of Bias and Other Methodological Criteria of Published Animal Studies: A Systematic Review" (Krauth, Woodruff and Bero, 2013. Environmental Health Perspectives 121(9):985-992).

Comments on "Agenda Item III: Committee Discussion on How to Tabulate Data from Epidemiological Studies in Hazard Identification Documents"

1. In identifying the study, the table should include if there is a name to the cohort, which can help identify overlapping data. The geographical location should be in a separate column.
2. Study design – This categorization may suffice for general identification, but it is not always easy to fit each study into a certain category. For example, you may have a cross sectional study that estimates retrospective exposure. We recommend specifying WHEN the exposure occurred and WHEN the outcome occurred rather than only the study design name.
3. Separate "exposure measurement method" from "likely routes of exposure." The first is factual and easily reported, but the latter may be uncertain and subjective.
4. Instead of "exposure dosages" – say "exposure concentrations" measured. Be specific about which exposure concentration corresponds to which outcome statistic, e.g. highest quartile of exposure (give actual values) associated with OR=1.5 for low birth weight.
5. Include exposure matrix/metric, e.g. cord serum, maternal urine, drinking water, residency, etc.
6. Include sample size for each outcome.
7. The "Results" column should allow for various measures, i.e. not only those listed, and multiple results from one study.
8. The last column implies that potential biases can be succinctly expressed in one small box, but we recommend removing this column and completing a separate formal in-depth assessment of study bias/quality to aid in the interpretation of the quality/reliability of the overall body of evidence. Application of such a "risk of bias" tool permits a transparent assessment of the same criteria for every included study and transparently demonstrates that each study is held to the same standard. See attachment for example tool for evaluating study quality/bias.

Comments on "Template for tabulating data from developmental toxicity studies" and "Template for tabulating data from male and female reproductive toxicity studies"

1. The specific chemical name should be included in the "chemical" parameter.

2. Include stage for “animal model” parameter to capture pregnancy day, lactation day, etc.
3. Animal source should be included in the “animal model” parameter.
4. For the “study design” parameter, specify that this is to incorporate details of the study design, and not simply a one-word description (e.g., multi-generational). Recording various details of study design may be critical for understanding the context of the study and interpreting the results.
5. Include information about control(s) for “exposure” parameter (important for studies where vehicle is not used as a control, or multiple controls are used).
6. Clarify what information will be included in the results column. Incorporate separate columns for various aspects of the results. For example, the results should include the age/stage of animal at assessment, average litter size (if appropriate), number of animals assessed (this may be different from number allocated), and statistical analyses performed.
7. As with the human epidemiological studies, we recommend a formal in-depth assessment of study bias/quality to aid in the interpretation of the quality/reliability of the overall body of evidence. See attachment for example tool for evaluating study quality/bias.

Thank you for providing the opportunity to give input to your process. We would be pleased to respond to any additional clarifications or questions you may have regarding our recommendations. Please feel free to contact Paula Johnson at 510-350-1245 or johnsonp@obgyn.ucsf.edu.

References:

1. Woodruff, T.J., P. Sutton, and The Navigation Guide Work Group, *An Evidence-Based Medicine Methodology To Bridge The Gap Between Clinical And Environmental Health Sciences*. Health affairs, 2011. **30**(5): p. 931-937.
2. Woodruff, T.J. and P. Sutton, *The Navigation Guide: An Improved Method For Translating Environmental Health Science into Better Health Outcomes*. Environmental Health Perspectives, In review.
3. Koustas, E., et al., *The Navigation Guide - Evidence-Based Medicine Meets Environmental Health: Systematic Review of Non-Human Evidence for PFOA Effects on Fetal Growth*. Environmental Health Perspectives, In review.
4. Johnson, P., et al., *The Navigation Guide - Evidence-Based Medicine Meets Environmental Health: Systematic Review of Human Evidence for PFOA Effects on Fetal Growth*. Environmental Health Perspectives, In review.
5. Lam, J., et al., *Integrating human and non-human evidence streams to establish strength of evidence in environmental health*. Environmental Health Perspectives, In review.
6. Vesterinen, H., et al., *The relationship between fetal growth and maternal glomerular filtration rate: a systematic review*. Environ Health Perspect, Submitted.

Appendix VII: Instructions for Making Risk of Bias Determinations

1. Was the strategy for recruiting participants consistent across study groups?

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

Protocols for recruitment and inclusion/exclusion criteria were applied similarly across study groups, and any one of the following:

- Study participants were recruited from the same population at the same time frame; or
- Study participants were not all recruited from the same population, but proportions of participants from each population in each study group are uniform

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

There is insufficient information about participant selection to permit a judgment of 'YES', but there is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria was consistent, as described by the criteria for a judgment of 'YES'.

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

Any one of the following:

- Protocols for recruitment or inclusion/exclusion criteria were applied differently across study groups; or
- Study participants were recruited at different time frames; or
- Study participants were recruited from different populations and proportions of participants from each population in each study group are not uniform

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

There is insufficient information about participant selection to permit a judgment of 'NO', but there is indirect evidence that suggests that participant recruitment or inclusion/exclusion criteria was inconsistent, as described by the criteria for a judgment of 'NO'.

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

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

2. Was knowledge of the exposure groups 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 '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 '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 '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.

3. Were exposure assessment methods robust?

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

The reviewers judge that there is low risk of exposure misclassification and any one of the following:

- There is high confidence in the accuracy of the exposure assessment methods; or
- Less-established or less direct exposure measurements are validated against well-established or direct methods

AND if applicable, appropriate QA/QC for methods are described and are satisfactory, with at least three of the following items reported, or at least two of the following items

reported plus evidence of satisfactory performance in a high quality inter-laboratory comparison:

Limit of detection or quantification; standards recovery; measure of repeatability; investigation and prevention of blanks contamination.

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

There is insufficient information about the exposure assessment methods to permit a judgment of 'YES', but there is indirect evidence that suggests that methods were robust, as described by the criteria for a judgment of 'YES'. Studies only reporting that the QA/QC items above were satisfactory but not reporting all of the actual numbers may receive a judgment of "probably yes."

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

The reviewers judge that there is high risk of exposure misclassification and any one of the following:

- There is low confidence in the accuracy of the exposure assessment methods; or
- Less-established or less direct exposure measurements are not validated and are suspected to introduce bias that impacts the outcome assessment (example: participants are asked to report exposure status retrospectively, subject to recall bias)
- Uncertain how exposure information was obtained

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

There is insufficient information about the exposure assessment methods to permit a judgment of 'NO', but there is indirect evidence that suggests that methods were not robust, as described by the criteria for a judgment of 'NO'.

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. Was confounding adequately addressed?

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

The study accounted for (i.e., matched, stratified, multivariate analysis or otherwise statistically controlled for) important potential confounders, or reported that potential confounders were evaluated and omitted because inclusion did not substantially affect the results. The determination of specific confounders may be informed by the data, including the studies included in the review.

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

The study accounted for most but not all of the important potential confounders
AND this lack of accounting is not expected to introduce substantial bias.

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

The study did not account for or evaluate important potential confounders.

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

The study accounted for some but not all of the important potential confounders
AND this lack of accounting may have introduced substantial bias.

5. Were incomplete outcome data adequately addressed?

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

Participants were followed long enough to obtain outcome measurements and 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 balanced in numbers across exposure 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 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 was adequately addressed, as described by the criteria for a judgment of 'YES'.

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

Participants were not followed long enough to obtain outcome measurements OR
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 exposure 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
- Potentially inappropriate application of imputation.

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 '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.

6. 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 protocol, methods, abstract, and/or introduction that are of interest in the review have been reported in the pre-specified way.

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 'NO' (i.e. high risk of bias):

Any one of the following:

- Not all of the study's pre-specified primary outcomes (as outlined in the protocol, methods, abstract, and/or introduction) 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; or
- One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected effect); or

One or more outcomes of interest are reported incompletely

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 '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.

7. 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.

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; or
- Stopped early due to some data-dependent process (including a formal-stopping rule); or
- Had extreme imbalance of characteristics among exposure groups; or
- Had differential surveillance for outcome between exposure groups or between exposed/unexposed groups

- The conduct of the study is affected by interim results (e.g. recruiting additional participants from a subgroup showing greater or lesser effect); or
- An insensitive instrument is used to measure outcomes (which can lead to under-estimation of both beneficial and harmful effects); or
- Selective reporting of subgroups; or
- Has been claimed to have been fraudulent; or
- Had some other problem

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 '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.

8. Was the study free of support from a company, study author, or other entity having a financial interest in any of the exposures 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. Examples 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', but there is indirect 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 '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;
- Chemicals provided at no cost;
- 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 '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 '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 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. 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. 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;

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):

Any one of the following:

- Authors did not report numbers analyzed for outcomes of interest
- 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.