Alliance for Risk Assessment (*ARA*): Beyond Science and Decision Workshop XV

Case Study 3: Fit-for-purpose systematic review guidelines for toxicity factor derivation, using vanadium as a case study

Presenters: Nnamdi Nnoli & Stony Lo, Toxicology, Risk Assessment, and Research Division, Texas Commission on Environmental Quality, Austin, TX

1. Provide a few sentences summarizing the method illustrated by the case study.

A systematic review (SR) is a methodological approach to answering a research question in a manner that minimizes the risk of bias and error and maximizes transparency. This method involves identifying, appraising, and synthesizing all relevant studies on a particular topic (Uman, 2011; WHO, 2021). This method also provides a scientifically robust approach to the review and interpretation of complex and often contradictory evidence in relation to a research question (WHO, 2021). To achieve this goal, SR methods are defined *a priori* in a comprehensive plan developed during problem formulation exercises. This *a priori* initiative supports answering the relevant research question in a transparent and unbiased fashion (WHO, 2021).

Adapted from the field of evidence-based medicine, SR in toxicology was proposed as a critical practice in determining causation (Guzelian et al., 2005). Since 2005, several organizations have adopted the concept and proposed best practices in the conduct of SR for the purposes of hazard and risk assessment. Notably, the European Food Safety Authority (EFSA), U.S. Environmental Protection Agency (USEPA), and the National Toxicology Program's former Office of Health Assessment and Translation (NTP OHAT) have published methods for conducting SR and evidence integration (EFSA, 2010; USEPA, 2022; NTP OHAT, 2019). The Texas Commission on Environmental Quality (TCEQ) has also previously published guidelines for performing SR for the purpose of developing toxicity factors (Schaefer and Myers, 2017). Among the peer-reviewed literature, Wikoff et al. (2020) proposes a framework that combines and builds on the aforementioned guidance for use by a practitioner in hazard and risk assessment. The World Health Organization (WHO, 2021) offers similar guidance on using systematic review to facilitate the chemical risk assessment process.

The overall objective of the TCEQ SR guidance is to provide a flexible, yet structured, framework for conducting an SR in the context of developing chemical-specific toxicity factors based on evidence from human and/or animal studies, along with supporting available mode-of-action (MOA) studies (when necessary). This case study reflects an update to the previous TCEQ Guidance on Systematic Review and Evidence Integration, originally finalized in 2017 (TCEQ, 2017), which was developed to supplement the TCEQ's 2015 Guidelines to Develop Toxicity Factors (RG-442). This work was done together with the consulting firm ToxStrategies.

Using the updated SR protocol, the TCEQ Toxicology, Risk Assessment, and Research Division (TD) conducted an SR using vanadium as the case study. After problem formulation, an initial literature search was completed and results were uploaded to SWIFT Active Screener, which is a web-based software designed to streamline systematic review, followed by title and abstract screening using set inclusion and exclusion criteria. A collaborative title and abstract (TiAb) screening workflow in SWIFT Active Screener continued until the reviewers reached the

required threshold for screening TiAbs. Thereafter, ToxStrategies and TCEQ scientists worked together to resolve conflicts. Full text pdf files of studies with TiAbs that met the screening criteria were uploaded into Health Assessment Workspace Collaborative (HAWC), an interactive content management system for human health assessments, and reviewed for relevance and data extraction by ToxStrategies and TCEQ scientists. Quality control review of the extracted epidemiology and animal toxicology data was completed. In addition, ToxStrategies updated the vanadium SR case study project in HAWC with refined study quality evaluation criteria and guidance, and piloted study quality evaluation for animal toxicity and human epidemiology studies.

ToxStrategies and TCEQ reviewers completed quality evaluation and quality control of the relevant epidemiology and animal studies. Once quality control review had been completed for all study quality evaluations, the output was integrated into summary evidence tables (see supplemental excel sheet of data extraction and study quality assessment and study quality evaluation heat maps for both epidemiology and animal toxicology literature).

2. Describe the problem formulation process as it applies to this case study

Following the decision to develop a TCEQ development support document (DSD) for a given chemical(s), characterizing the scope of the assessment will begin with a planning phase. In evidence-based methods, this is known as Problem Formulation. Importantly, in the context of risk assessment, "assessment planning" is also incorporated into problem formulation. This includes designing and stating the methods for components of risk assessment that the SR could inform (e.g., hazard identification, toxicity factor derivation, exposure assessment, MOA analysis, toxicokinetics, etc.), recognizing that a systematic review method can be applied to facilitate multiple aspects of risk assessment.

Scoping exercises to aid in this effort are performed during Problem Formulation. Prompting questions may be useful in guiding these exercises. These may include:

- What is the specific context of the assessment?
- What is the timeline, and what resources are available?
- What is the required output to meet the overall goal of the assessment?
- What are the physical and chemical properties of the chemical?
- Are there existing systematic reviews or agency evaluations?
- Are there data available?
- Are the critical effects known?
- Are there known potentially sensitive subpopulations?
- Are the toxicokinetics known, and does route of exposure play a role in toxicity?
- Is the chemical carcinogenic? If so, is the chemical carcinogenic only by a specific route of exposure or when a biologically plausible threshold is exceeded?

Key exercises performed in this phase will include:

- Identification and review of assessments conducted by other organizations or in the peer-reviewed literature
- Scoping the volume and nature of evidence to determine the need for a systematic

evidence map (SEM) or SR

- Definition of risk assessment question and structured Population, Exposure, Comparator, and Outcome (PECO) or Population, Concept, and Context (PCC) elements (defined below)
- Protocol development, including determination of inclusion/exclusion criteria
- Piloting

3. Describe the systematic review protocol and its objective.

The overall objective of the TCEQ SR guidance is to provide a flexible, yet structured, framework for conducting an SR in the context of developing chemical-specific toxicity factors based on evidence from human and/or animal studies, along with supporting available mode-of-action (MOA) studies (when necessary).

This updated SR guidance document builds on previous SR guidance (TCEQ, 2017) with available existing methods in conducting SRs and integrating evidence for the purpose of developing reference values (ReVs), unit risk factors (URFs), oral slope factors (SFos), and reference doses (RfDs). A significant revision in the workflow presented herein compared to previous guidance is the addition of a potential SEM workflow (Figure 1). Similar to SR, an SEM uses robust and transparent methods to systematically explore and describe the literature on a given topic. As stated by WHO (2021), this method can be used to help with prioritization, and the addition of this technique is anticipated to provide guidance when a narrow, more specific SR question cannot be formulated. In contrast to the specific SR question, an open-framed question (or one that lacks specification of some key elements) is posed instead. Practitioners have found that stepwise use of these evidence-based tools is helpful to facilitate the risk assessment process when large evidence bases involving many different outcomes and different evidence streams need to be assessed, as is common in risk assessment and development of toxicity factors.

The following sections provide a summary description of the following aspects of the TCEQ SR protocol:

I. Scoping
II. PECO development
III. Inclusion/exclusion criteria
IV. Tool selection
V. Protocol development
VI. Evidence Identification (literature search, data solicitation, database searching, targeted search)
VII. Literature screening (use of AI/ML-based literature review tools)
VIII. Assessing relevance to risk assessment
IX. Data extraction
X. Study reliability assessment
XI. Evidence Synthesis, Integration, and Derivation of Chemical Toxicity Values

Detailed information can be found in Supplemental material 1.

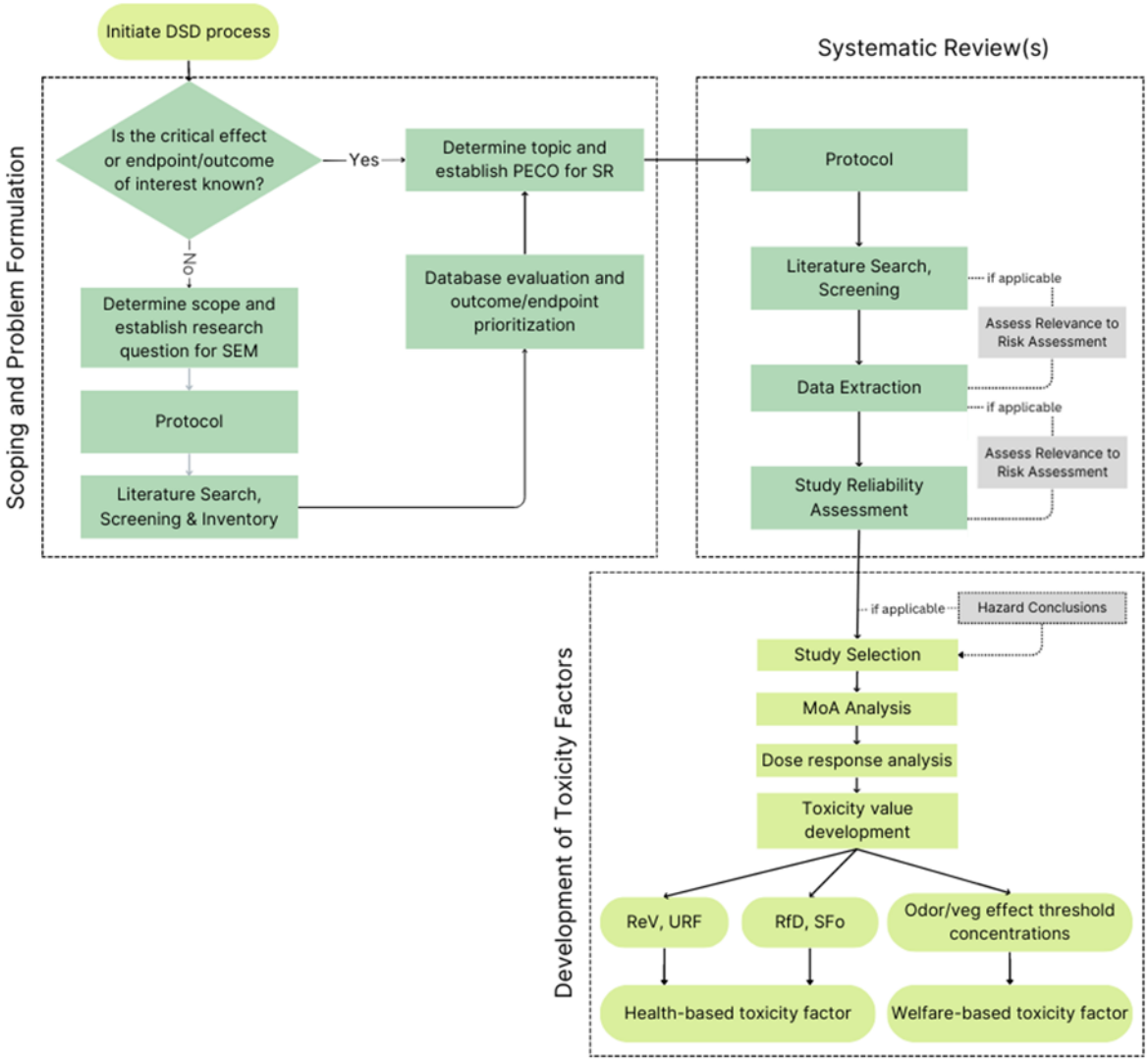

Figure 1. Framework for implementing systematic methods in support of developing toxicity factors

4. Describe the Systematic review protocol as applied to the case study of Vanadium and compounds

The TCEQ Systematic Review protocol was used to conduct a systematic review for vanadium and vanadium compounds as described below:

Introduction

Vanadium (V) is a metal that exists in several oxidation states (Barceloux, 1999). About 80% of the V produced is used as an additive to stabilize steel, and is commonly used in the manufacturing of automobile parts and tools (ATSDR, 2012). Vanadium pentoxide (V_2O_5) is used as a catalyst in the production of sulfuric acid and plastics (Friberg et al., 1986, as cited in OEHHA, 1999). The major anthropogenic point sources of atmospheric emissions are

metallurgical works followed by the burning of crude or residual oil and coal. Incinerators emit V in the form of V_2O_5; the stable form of V, in the presence of atmospheric oxygen. TCEQ has previously developed interim health and welfare-based values from acute and chronic evaluations of V and V_2O_5 to be used for air permitting and air monitoring. The following V compounds are included in this evaluation: V, V_2O_5, vanadium dioxide (VO_2), bismuth orthovanadate (BiO_4V), vanadyl sulfate dihydrate ($VOSO_4$), sodium metavanadate ($NaVO_3$), ammonium metavanadate (NH_4VO_3), and sodium orthovanadate (Na_3VO_4).

Objective

The objective is to perform a SR of literature on toxicity observed in humans and animals exposed to vanadium compounds. Specifically, the findings of this SR will inform the development of toxicity factors in a TCEQ DSD. The protocol contained herein describes the framework of this review and serves as documentation of study design decisions. Any deviations from this protocol will be documented in the change log.

Population, Exposure, Control, and Outcome (PECO) Statement

In humans, what are the apical effects (cancer or non-cancer) of inhaled vanadium at any concentration compared to unexposed individuals.
P: Human population
E: Inhaled vanadium at any concentration
C: No exposure
O: Apical effects (cancer or non-cancer)

Literature Identification

Search strategy

The literature search strategy was developed to identify toxicological data from four primary channels using the identified relevant compounds in Table 1. This includes:

1. PubMed: A detailed search syntax that consists of vanadium compound names, synonyms, CAS numbers, and any other identifying information was used to query the citation database. The search was date-limited (2020 – present) due to the availability of existing resources, specifically the USEPA SEM database for V compounds (2021). Syntax developed during scoping and problem formulation generated 2,938 results; syntax is available in Attachment A.
2. The USEPA SEM database for V compounds (USEPA 2021) was utilized for peer-reviewed literature published prior to 2020.
3. Regulatory and/or authoritative body websites were searched using V compound identifiers for previous risk assessments.
4. Responses to TCEQ's call for data (initial data request notice posted on April 14, 2021, for V and V_2O_5, with a response deadline of July 16, 2021). There were no responses to the public data request, so no additional datasets were added to the results through this channel.

Validation of results

The search results were validated using the following articles to ensure the syntax is complete and thorough: Alvarez-Barrera et al. (2022); Fan et al. (2023); Frawley et al. (2023); He et al.

(2023); Li et al. (2021); Montiel-Flores et al. (2021); Tu et al. (2022); Waidyanatha et al. (2022); Xi et al. (2021); Zhang et al. (2021).

Screening and selection

The TiAb of all references were screened by two reviewers in Swift Active Screener based on inclusion and exclusion criteria. A guide for building screening forms in Swift Active Screener is available in Attachment B. Prior to initiating this effort, a piloting and reviewer calibration phase was performed to calibrate responses across reviewers, make topic-specific adjustments to the inclusion/exclusion criteria, and address any other potential issues related to TiAb screening. Citations that met the inclusion criteria were moved forward to full text review to confirm relevance, using the full manuscript or report, prior to data extraction. The reviewers discussed any inconsistencies with each other's tag selection after both TiAb and full text screening. If the original reviewers could not come to a resolution, a third reviewer was used to provide input.

Inclusion/Exclusion Criteria

Inclusion and exclusion criteria were developed by TCEQ based on the PECO statement and objective of the SR (Table 1). Each identified citation was reviewed in the context of these criteria to determine eligibility, first at TiAb, then full text. For studies excluded at full text screening, one or more exclusion reasons were documented.

Table 1. Inclusion and exclusion criteria used during full text review of the SR

Category	Include	Exclude
Population	Humans Experimental animals (mammalian species)	Non-mammalian species Ecological field studies (e.g., ecotoxicity) Mechanistic evidence [a]
Exposure	Inhalation exposure to any of the Vanadium compounds listed in Table 1 Any exposure duration **Human-specific**: Exposure metrics provided as actual measured air concentrations (e.g., $\mu g/m^3$) **Animal-specific**: Controlled/known exposures	Any route other than inhalation (e.g. injection, oral, dermal) Biological biomarker studies (e.g., vanadium in toenail clippings) Exposure metrics not provided as an actual measured air concentration Exposure concentration unknown or does not have a control group for comparison
Outcomes	Any adverse cancer or non-cancer outcomes such as respiratory, immune, DART, hepatic, renal, cardiometabolic, hematologic, nervous compared to a control population or experimental group.	Any non-apical outcomes such as mode of action [a] or toxicokinetics [a]

Category	Include	Exclude
Reference type	Primary experimental or observational studies that report empirical research	Any case studies, reviews (including SR), meta-analyses, commentaries, editorials, or any other secondary reporting[a]
Additional criteria	See exclusion criteria	Publications not available in English Study reports or publications that are not available for review in full (e.g., only the abstract is available)

[a] *These categories were excluded from systematic review and toxicity factor development; however, they were still categorized for potential contextual information and data used to support interpretation of eligible studies.*

The results of the screening process were documented in the form of a PRISMA chart (Figure 2).

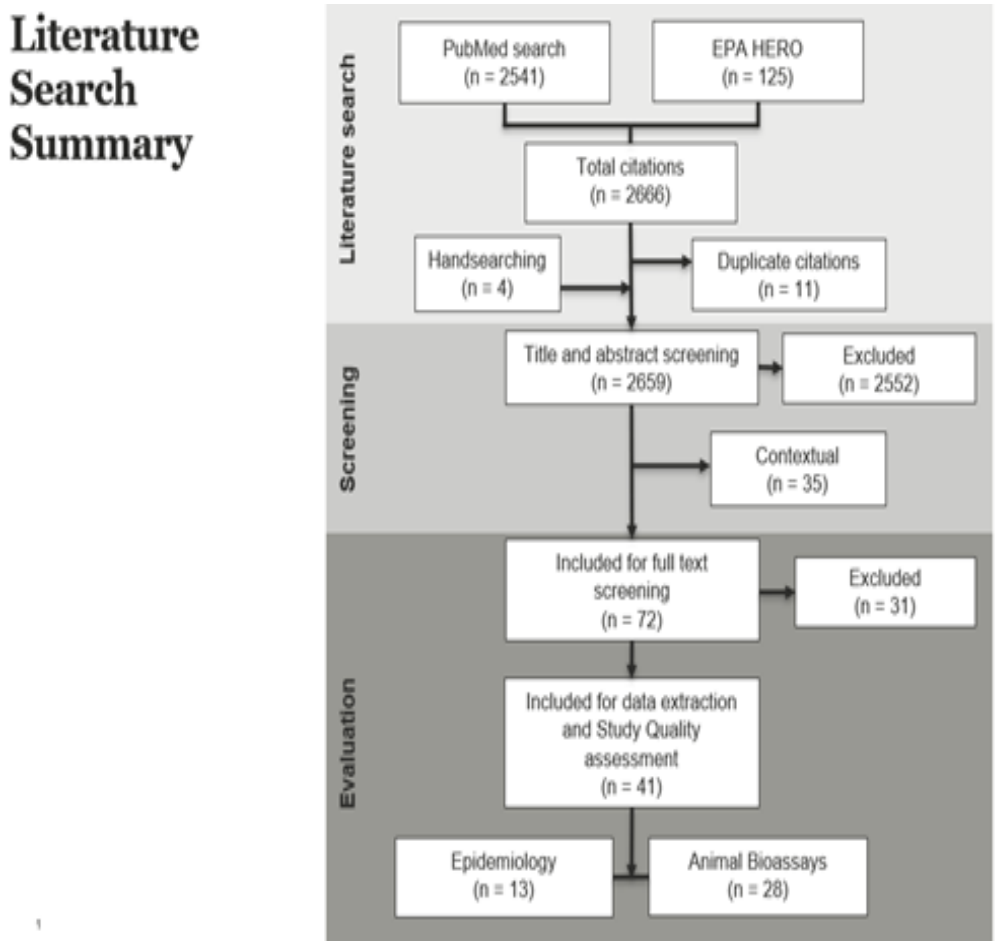

Figure 2: Literature search summary for Vanadium systematic review case study
Note: Handsearching consisted of the manual crosscheck of the V DSD by TCEQ, which resulted in 4 studies being added to HAWC. Contextual studies were not added to HAWC but were noted in the SWIFT Screening Results. A total of 72 studies went through full text evaluation in HAWC. The EPA HERO (Health and Environmental Research Online) database is a searchable repository of scientific studies and references used by the U.S. EPA to develop its risk assessments.

Data Extraction
Data extraction focused on information pertinent to the derivation of toxicity factors as outlined by TCEQ's Guidelines to Develop Toxicity Factors (2015). The review team extracted details of study information, experimental design, and results via pre-populated fields in HAWC. HAWC has detailed extraction pages for each section of human or animal studies. A list of fields relevant to human and animal data extraction are provided in Attachment C. Any changes made to

extraction records during extraction and quality control verification are recorded by HAWC. Quality control of data extraction was performed by the approved designated quality control scientists prior to an assessment of study quality. Details are provided in the data extraction quality control guidance document (Supplemental material 2).

Study Reliability

Study quality evaluation as it relates to the development support document (DSD) for vanadium and compounds will require that each experimental animal and epidemiology study that meets the inclusion criteria will be evaluated for risk of bias (RoB) and study sensitivity. Study quality was not evaluated for mechanistic endpoints.

For experimental animal and epidemiology studies, the quality assessment was conducted at the study design level. The study quality assessment was adapted from study evaluation domains available in HAWC (Shapiro et al., 2018). These domains are arranged into metrics that provide both core and prompting questions to aid the reviewer in assessing the study's reporting, RoB and study sensitivity on an outcome-specific basis. TCEQ has focused the process on the most critical/impactful core questions (Shapiro et al., 2018) to simplify the evaluation without addressing every prompting question in the traditional IRIS-type study quality review. Using the provided guidance, the reviewers scored each domain as Good (++), Adequate (+), Deficient (-), or Critically Deficient. Following domain scoring, each study was given an overall confidence rating of High (++), Medium (+), Low (-), or Uninformative (--) based on these domain scores. Where needed, study quality questions will be refined to target specific aspects of the V literature. A second reviewer performed quality control on the initial reviewer's evaluation. When major conflicts occurred between reviewers during scoring (i.e., inconsistent application of "deficient" or "critically deficient" ratings for domain metrics; inconsistent judgement of overall study quality as "low confidence" or "unreliable"), this was resolved by discussion between the two reviewers. If the two reviewers could not come to a resolution, a third reviewer gave input to make a final determination. Details of the study quality domains are provided in Appendix 1.

This process first went through user reviewer piloting and reviewer calibration, which resulted in iterative refinement of the form. All revisions were documented within the protocol. Details of the Risk of Bias (RoB) assessment tool are provided in Appendix 2. In practice, the study quality evaluation should also include a pilot phase during which assessors apply the refined tool, provide feedback and discussion, and ensure judgements are aligned across the assessment team. The standard output of the evaluation is provided as a heat map of all studies as shown in Figures 3 and 4.

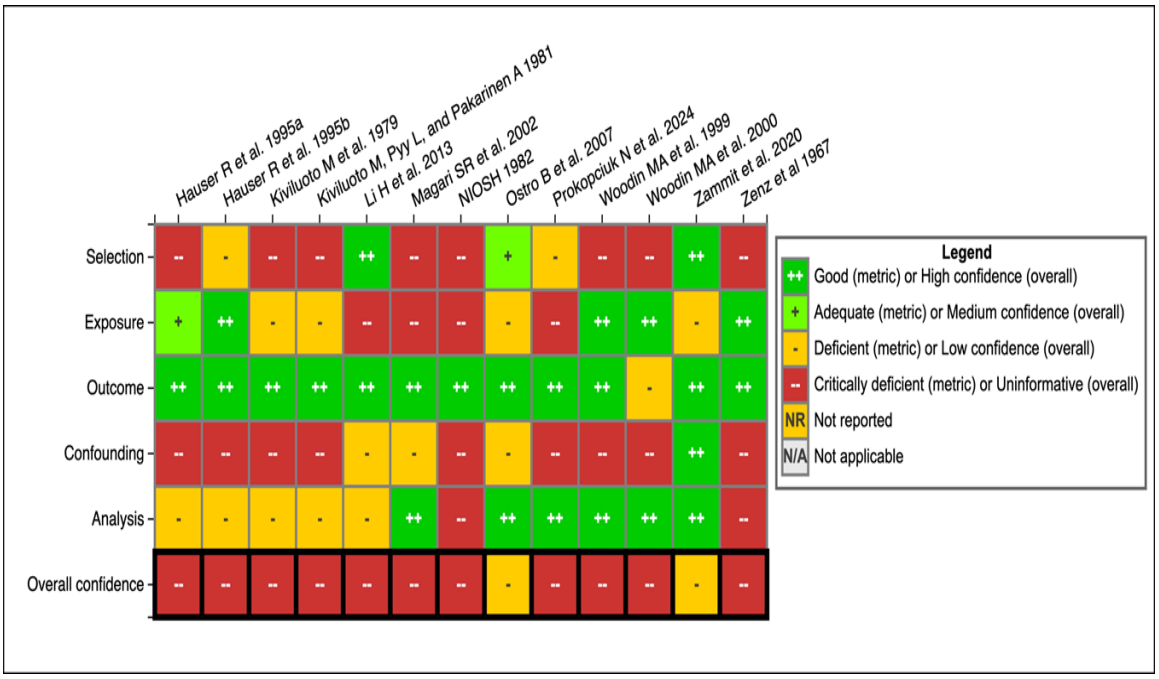

Figure 3: Vanadium Epidemiology Study Quality Heatmap

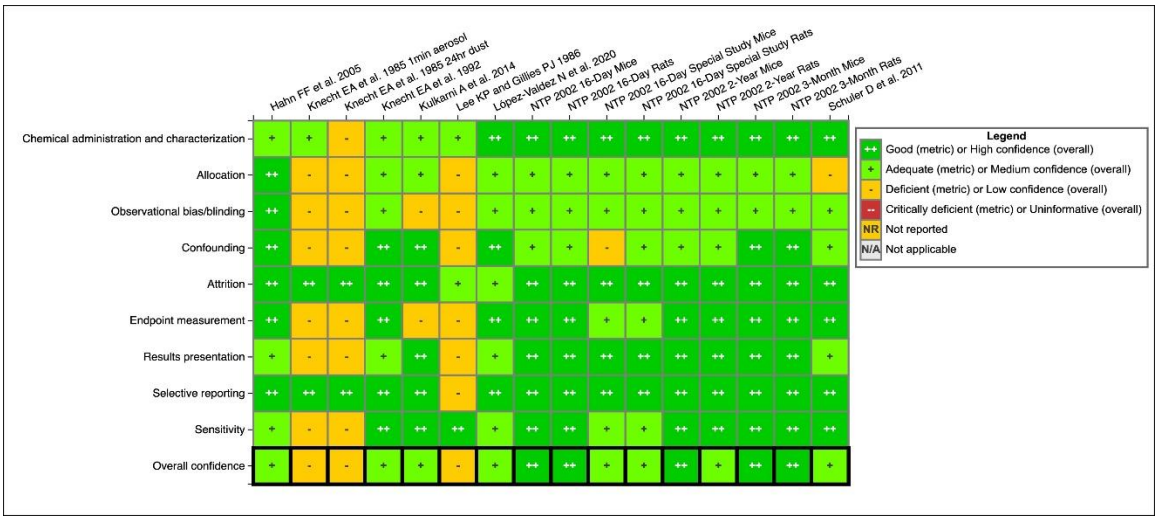

Figure 4: Vanadium Animal Study Quality Heatmap

Evidence Integration and Synthesis

After addressing the study quality and RoB for each of the selected studies, the information from each of the data streams (human, animal, mechanistic) was compiled together and assessed for use as key, supporting, and informative studies. This information is put into an evidence integration table (Supplemental material 3).

Ultimately, the synthesis and integration process are guided by TCEQ's Guidelines to Develop Toxicity Factors (2015) and typically results in the development of toxicity factors such as inhalation reference value (ReV), inhalation unit risk factor (URF), oral reference dose (RfD), and/or oral slope factor (SFo) values. Using the available data identified and assembled in the SR, key and supporting studies will be identified based on attributes of study design, study reliability, and needs of dose-response modeling. The TCEQ DSD process provides guidance on evaluating the weight of evidence including factors such as toxicokinetics and mode of action that can affect human relevance and quantitative dose response, as well as study quality and database confidence (TCEQ, 2015). Consequently, this approach will support the development of chemical risk assessments that are transparent, consistent, and reliable, conferring confidence in the DSD process.

Evidence integration and synthesis procedures as they relate to the development of the DSD for vanadium and compounds focused on studies pertaining to the inhalation route of exposure.

Selected key studies from vanadium systematic review are thus:
For acute toxicity factors: Schuler et al. First steps towards an understanding of a mode of carcinogenic action for vanadium pentoxide. Journal of Toxicologic Pathology, 2011; 24: 149 - 162
For chronic toxicity factors: National Toxicology Program. Technical report on the toxicology and carcinogenesis studies of vanadium pentoxide (CAS No. 1314-62-1) in F344/N rats and B6C3F1 mice (Inhalation studies) NTP, 2002. (2-Year rat and mice studies).
Evidence integration table for the selected studies is provided in Appendix 3.

5. Discuss the overall strengths and limitations of the systematic review protocol.

Strengths:

- **Transparency and Reproducibility**: The clear, detailed and defined methodology employed in SR allows for transparency and ensures that the review process can be replicated by other researchers.
- **Minimized Risk of Bias**: The thorough and organized process using preset inclusion and exclusion criteria as well as the use of independent study quality reviewers reduced the risk of bias in the review process (i.e., concerns about "cherry-picking" studies).

- **Increased confidence in findings**: The data extraction as well as evidence synthesis and integration process from multiple studies provide more robust and generalizable information on which to base conclusions.
- **Study quality assessment**: The thorough appraisal of included studies helps to identify potential biases and study limitations to ensure that the overall conclusions are based on reliable evidence.
- **Broad literature search**: The SR process helps to identify and document a broader body of literature on a topic than an ordinary literature search.
- **Identification of future research topics**: The detailed nature of the SR process helps to identify inconsistencies and limitations present in the conducted studies, thereby highlighting areas for future research/investigation.
- **Informed Decision-Making**: The SR process helps to identify, document, and synthesize a large body of knowledge necessary for informed decision-making on the researched subject.
- **Comprehensive review of study results**: Systematic reviews assemble and synthesize a large body of studies across different study settings providing the ability and justification for the generalization of study results.

Weaknesses:

Systematic reviews are a rigorous method for synthesizing research, but they are not without weaknesses.

- **Resource and time intensive:** The SR process consumes a lot of time and resources. A detailed SR process can last for several months and require both human and financial resources per reviewers and monetary subscription for applicable software tool use in the process. The process presented here is intended to provide some streamlining to decrease resource/time needs where possible.
- **Publication bias**: Studies having statistically significant results are more likely to be published compared to studies without any statistically significant findings. The list of literature retrieved through the SR process may not necessarily represent the complete picture of the available evidence. Similarly, some authors may preferentially report only results that support their hypothesis and skip others. The SR process cannot fix the problem around these unreported outcomes.

- **Language bias**: Due to language barriers, many SRs focus on studies published in specific language(s). This language restriction can result in missing pertinent evidence regarding the subject matter being studied.
- **Inappropriate use:** SRs may be inappropriately designed, especially when researching a topic that is novel. Thus, systematic evidence mapping is often conducted initially to gain some insight that will inform the problem formulation and development of protocol for the SR.
- **Conclusions may be the same without or without SR**: SRs are sometimes treated like an essential tool for conducting evaluations, without which an analysis is fundamentally flawed. However, particularly for toxicity factor derivations that come down to a few key and supporting studies, these studies may have been identified, the quality evaluated, etc. without a SR. So, the resources used for the SR would have come to the same end.

6. Outline the requirements needed for a systematic review.

A systematic review requires a transparent, reproducible, and rigorous methodology to synthesize available evidence and minimize the potential for bias. Unlike ordinary literature review, the systematic review process adheres to a defined protocol at every stage of the process.

A. **Requirements for the reviewers:**

I. Number of reviewers: The systematic review process requires a minimum of two independent reviewers who will select and assess the studies for inclusion or exclusion based on set criteria. A third reviewer may be required to help resolve conflicting opinions between the two reviewers.

II. Expertise: The reviewers should be individuals who have a reasonable knowledge base on the subject matter under review to ensure that concepts are well-understood.

III. Librarian: A librarian or information management expert may help with establishing a detailed literature search strategy across several databases and to help with retrieving the full text copies of the studies needed for review.

B. **Requirements for the review process:**

I. Protocol development: Prior to conducting a systematic review, the review team must design a comprehensive protocol that outlines the methodology for the review process. Among other things, the protocol should include:

- **Research question:** The protocol should define the goal of the review. A clear and specific research question at the problem formulation stage and a well-defined PECO (Population,

Exposure, Comparison/Control, Outcome) statement will help streamline the review process from the outset.

- **Inclusion and exclusion criteria:** The review team must set a clear and detailed inclusion and exclusion criteria to guide which studies to include in the review and those to exclude from the review. These criteria may include language restrictions, study design, type of study, study population, and nature of exposure measurement.
- **Search strategy:** The systematic review protocol must state the databases and search-terms employed in retrieving literature from the databases to make the search reproducible.
- **Data collection and analysis plan:** The review team must have a data collection tool/platform so that extracted from selected studies will be organized in a format that makes it easy to synthesize.
- **Risk-of-bias assessment:** There must be a set guideline for assessing the risk of bias and quality of studies selected for data extraction.

II. Comprehensive literature searching: The search must be systematic and thorough to find all relevant evidence.

- **Literature sources:** The search for relevant literature must be comprehensive and should include various databases, grey literature, and other non-journal publications. (e.g., conference proceedings, government reports and archives).
- **Document search strategy:** The full electronic search strategy for all the searched databases must be accurately documented to ensure reproducibility.
- **Manage citations:** The review team should manage study references using appropriate citation management tools such as Endnote, to enable study collection, removal of duplicate studies and management of all retrieved records.

III. Study selection based on set criteria: Two independent reviewers must screen all identified studies against predetermined criteria in a two-stage process.

- **Title and abstract screening:** Every reviewer independently screens the titles and abstracts of all retrieved literature to identify studies for inclusion or exclusion.
- **Full-text screening:** Full-text articles of the selected studies are obtained and screened against the inclusion/exclusion criteria. Should there be disagreements between the two reviewers, this should be resolved through discussion between the two reviewers. If the two reviewers cannot reach a conclusion, a third reviewer is required to break the tie to reach a final resolution.

IV. Data extraction and quality assessment

- **Standardized data extraction:** Relevant data are extracted from the studies selected after the full text screening using data tables/log already set during the design of the systematic review protocol.
- **Risk of bias:** Risk of bias evaluation should be conducted on all selected studies to identify potential biases so that results from selected studies would be reliable and valid.

V. Synthesize and document results

- **Data integration and synthesis:** The extracted data is integrated and evaluated to arrive at a specific conclusion based on the research question.
- **Limitations:** It is important to discuss the limitations of both the individual studies and the overall review process to ensure that findings are interpreted and presented within the appropriate context.
- **Transparent documentation:** The systematic review methodology, findings, and conclusions must be reported in a transparent manner to ensure that credibility of the process and findings are justified.

7. Questions for the panel.

Following review of the updated SR method and its application in the development of toxicity factors, the TCEQ poses the following questions for the panel.

a. Does this systematic review protocol entail reasonable considerations and bases for study selection in development of toxicity factors?

b. Can you suggest any alternate considerations/bases for study selection in development of toxicity factors?

c. Does the panel have specific recommendations concerning TCEQ's updated SR?

a. Does the application of SR in study identification and evaluation in development of toxicity factors imply that toxicity factors developed without a SR process would be considered unacceptable?

b. Can you suggest possible ways to address some of the weaknesses of the SR process?

c. With the recent trend in artificial intelligence (AI), how can AI be used to advance SR?

References:

EFSA (European Food Safety Authority). (2010). Application of systematic review methodology to food and feed safety assessments to support decision making. *EFSA 8*(6), 1637. doi:10.2903/j.efsa.2010.1637.

Guzelian PS, Victoroff MS, Halmes NC, James RC, & Guzelian CP. (2005). Evidence-based toxicology: A comprehensive framework for causation. *Human & Experimental Toxicology*, *24*(4), 161–201. https://doi.org/10.1191/0960327105ht517oa.

Klimisch HJ, Andreae M, & Tillmann U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regul Toxicol Pharmacol, 25*(1), 1-5. https://doi.org/10.1006/rtph.1996.1076.

LaKind JS, Burns CJ, Johnson GT, & Lange SS. (2023). Epidemiology for risk assessment: The US Environmental Protection Agency quality considerations and the Matrix. *Hygiene and Environmental Health Advances, 6*, 100059. https://doi.org/10.1016/j.heha.2023.100059.

Meek MEB & Wikoff D. (2023, Jun 28). The need for good practice in the application of mechanistic constructs in hazard and risk assessment. *Toxicol Sci, 194*(1), 13-22. https://doi.org/10.1093/toxsci/kfad039.

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

Morgan RL, Whaley P, Thayer KA, & Schunemann HJ. (2019). Identifying the PECO: A framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. *Environment International, 121*(Pt 1), 1027-1031. https://doi.org/10.1016/j.envint.2018.07.015.

NRC (National Research Council). (2001). Standing Operating Procedures (SOP) of the National Advisory Committee on Acute Exposure Guideline Levels for Hazardous Substances (PDF). www.epa.gov/opptintr/aegl/pubs/sop.pdf

NRC (National Research Council). (2014). Review of the Environmental Protection Agency's Integrated Risk Information System (IRIS) process. Washington (DC): National Academies Press; 2014.

NTP OHAT (National Toxicology Program, Office of Health Assessment and Translation). (2019). Handbook for conducting a literature-based health assessment using OHAT approach from systematic review and evidence integration. OHAT, Division of the National Toxicology Program, National Institute of Environmental Health Sciences. Available at: https://ntp.niehs.nih.gov/sites/default/files/ntp/ohat/pubs/handbookmarch2019_508.pdf

Schaefer HR & Myers JL. (2017). Guidelines for performing systematic reviews in the development of toxicity factors. *Regulatory Toxicology and Pharmacology*, 91, 124–141. https://doi.org/10.1016/j.yrtph.2017.10.008.

TCEQ (Texas Commission on Environmental Quality). (2015). Guidelines to develop toxicity factors (RG-442). Chief Engineer's Office. Available at: https://www.tceq.texas.gov/toxicology/esl/guidelines/about

TCEQ (Texas Commission on Environmental Quality). (2017). TCEQ Guidelines for systematic review and evidence integration. White Paper. Toxicology Division.

Uman L. (2011). Systematic reviews and meta-analyses. Journal of the Canadian Academy of Child & Adolescent Psychiatry, 20(1): 57-59. PMID: 21286370.

USEPA (U.S. Environmental Protection Agency). (1994). *Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry.* EPA/600/8-90/066F. Washington, DC. Available at: https://www.epa.gov/risk/methods-derivation-inhalation-reference-concentrations-and-application-inhalation-dosimetry.

USEPA (U.S. Environmental Protection Agency). (2005). United States Environmental Protection Agency. Science Issue Paper: Mode of Carcinogenic Action for Cacodylic Acid (Dimethylarsinic Acid, DMA) and Recommendations for Dose Response Extrapolation. Washington, DC. Available at: https://archive.epa.gov/pesticides/reregistration/web/pdf/dma_moa-2.pdf

USEPA (U.S. Environmental Protection Agency). (2022). *Office of Research and Development (ORD) Staff Handbook for Developing IRIS Assessments.* EPA/600/R-22/268. Washington, DC. Available at: https://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=356370

WHO (World Health Organization). (2021). *Framework for the use of systematic review in chemical risk assessment*. ISBN 978-92-4-003448-8. Geneva: World Health Organization. Available at: https://www.who.int/publications/i/item/9789240034488

Attachment A - Draft Literature Search Syntax

Draft literature search syntax to be used in PubMed:

((Vanadium[Mesh] OR "Vanadates"[Mesh] OR "Vanadium Compounds"[Mesh] OR "vanadium iodoperoxidase"[Supplementary Concept] OR "7440-62-2"[rn] OR "Vanadium pentoxide" OR "vanadium pentoxide" [Supplementary Concept] OR "N,N-bis(salicylidene)-o-phenylenediamine vanadium (IV) oxide"[Supplementary Concept] OR "1314-62-1"[rn] OR "Vanadium oxide" OR "vanadic anhydride" OR "divanadium pentoxide" OR "vanadyl sulfate"[Supplementary Concept] OR "27774-13-6"[rn] OR "Vanadyl sulfate dihydrate"[tiab:~1] OR "vanadic sulfate"[tiab:~1] OR "vanadium oxide sulfate"[tiab:~1] OR "Sodium metavanadate" OR "vanadic acid monosodium"[tiab:~1] OR "ammonium metavanadate"[Supplementary Concept] OR "7803-55-6"[rn] OR "Ammonium metavanadate" OR "Ammonium vanadate" OR "Ammonium monovanadate" OR (("ammonium vanadium") AND (oxide OR trioxide)) OR "vanadic acid" OR "ammonium salt" OR "Sodium orthovanadate" OR "Sodium o-vanadate" OR "Sodium pervanadate" OR "sodium vanadium oxide" OR (("vanadic acid") AND ("trisodium salt"))) OR (Bismuth orthovanadate OR "bismuth orthovanadate" [Supplementary Concept] OR "14059-33-7"[rn] OR "Bismuth vanadate" OR ("bismuth tetraoxidovanadate"[tiab:~2]) OR "bismuth vanadium oxide" OR "vanadic acid")) OR (vanadium dioxide OR "vanadium dioxide"[Supplementary Concept] OR "12036-21-4"[rn] OR (Bis AND oxidanylidene AND vanadium)) AND (2020:2024[pdat])

Attachment B - Title and Abstract Screening Guide

This is a starting place that should be revised on a case-by-case basis. Depending on the chemical, PECO, and needs of the risk assessment, inclusion and exclusion reasoning will need to be adjusted based on the protocol and piloting.

For example, the protocol may state mechanistic/mode of action studies are important to the review process. The Reason for inclusion option would no longer be "Contextual - Mechanistic/Mode of Action" and could be changed to "Mechanistic/Mode of Action". Similarly, the Reason for exclusion should be adjusted based on the PECO and what is decided as not relevant during piloting and throughout screening.

- Does this reference meet the inclusion/exclusion criteria defined by the PECO statement?
    - Yes, this study is a [insert type of studies want to include]
        - (example: Yes, this study is a human or animal study measuring apical outcomes from vanadium inhalation exposure)
        - Included
            - Reason for inclusion:
                - Epidemiological
                - Toxicological
                - Contextual - Mechanistic/Mode of Action
                - Contextual - Kinetics
                - Contextual - Relevant Review
    - No, study does not meet the PECO requirements
        - Excluded
            - Reason for exclusion:
                - No apical outcomes or mechanistic evidence
                - Not in English
                - Exposure route not relevant
                - Biological biomarker
                - Non-mammalian species
                - Ecological study
                - No measured exposure
                - Only abstract available

Attachment C - Data Extraction

Data extraction fields available in HAWC:
Note: Fields an asterisk (*) must be filled out before HAWC will allow the page to be saved.

- Human Study Extraction in HAWC -
- Study design fields:
    - Summary*
    - Name
    - Design*
    - Source*
    - Age category*
    - Age details
    - Sex*
    - Race/ethnicity
    - Population N*
    - Enrollment period
    - Follow up
    - Countries
    - Other geographical information
    - Inclusion/Exclusion criteria
    - Susceptibility
    - Additional comments

- Exposure details fields: multiple rows can be added with additional fields for data
    - Chemicals – name, CASRN, DTXSID
    - Exposure measurements – name, measurement type, timing
    - Exposure levels – name, chemical, exposure, central tendency
    - Outcomes – system, effect, effect detail, outcome
    - Adjustment factors – name, description
    - Data extractions – group, outcome exposure level, timing, Estimate type, N, value, confidence

- Animal Study Extraction in HAWC -
- Study fields:
    - Name*
    - Type*
    - Multiple Generations (yes/no)
    - Chemical Name
    - Chemical Identifier
    - DSSTox substance identifier
    - Source of chemical
    - Purity (check box)
        - Purity qualifier (<, >, etc.)
        - Chemical purity (percentage)
        - Chemical vehicle
    - Guideline compliance (OCED, GLP, etc.)

- Methods fields:
    - Species*
    - Strain*
    - Sex*
    - Animal source
    - Exposure life stage
    - Life stage at assessment
    - Siblings
    - Husbandry
    - Diet
    - Route*
    - Exposure duration (days and additional notes)
    - Exposure-outcome duration
    - Number of dose groups*
    - Positive and/or negative control*
    - Dose groups*
- Endpoint fields:

- System
- Organ/tissue
- Effect
- Effect subtype
- Observation time and units*
- Dataset type*
- Variance type*
- Response units
- Data location
- Expected response adversity direction*
- NOEL*
- LOEL*
- FEL (Frank Effect Level)*
- Monotonicity*
- Statistical tests
- Trend results*
- Trend value
- Notes
- Additional methods
- Dose response data