UCSF Program on Reproductive Health and the Environment

EPA must adopt a scientifically valid systematic review method.

Recommendation

To ensure all evidence is properly considered to identify how toxic chemicals impact health, EPA should immediately implement a single, Agency-wide science-based, validated, and transparent systematic review method consistent with the best available science in all scientific assessments.

SUMMARY

Systematic review methods provide a transparent, consistent approach to reduce bias in evidence evaluation and a foundation for environmental health assessments that are more trustworthy and reliable, leading to policy decisions that safeguard public health and advance environmental justice.

Established systematic review methods with clear protocols for literature search, study selection, evidence evaluation, and evidence synthesis can be implemented by EPA without delay to promote health-protective chemical evaluations.

EPA's Integrated Risk Information System (IRIS) program has successfully incorporated many important aspects of consistent and transparent systematic review methods, though there are still some key improvements needed. However, EPA's Toxic Substances Control Act (TSCA) systematic review suffers from critical flaws that have led to significant deficiencies in the Agency's chemical evaluations. EPA must address deficiencies in the IRIS method and more urgently make critical revisions to its flawed approach to TSCA systematic review.

PROPOSED ACTIONS

To ensure EPA is considering all the science and not cherry-picking data when evaluating health harms, EPA should:

1. **Immediately implement a single, Agency-wide transparent, consistent, and science-based systematic review method consistent with NASEM recommendations and the best available science that considers all evidence streams for all scientific assessments that impact policy or regulation.**
2. **Use study assessment tools that are empirically based and should modify the current IRIS approach so that it does not inappropriately discount or exclude studies based on perceived methodological flaws.**
3. **Invest in implementation training and tools for EPA scientists, managers, and consultants across the Agency to ensure best practices and consistency in systematic review.**

SUPPORTING EVIDENCE

EPA should immediately implement a single, Agency-wide transparent, consistent, and science-based systematic review method consistent with NASEM recommendations and the best available science that considers all evidence streams for all scientific assessments that impact policy or regulation.

Empirically proven methods for research synthesis from the medical field were adapted for environmental health through interdisciplinary collaborative efforts, beginning with the development and implementation of the University of California, San Francisco (UCSF) "Navigation Guide Systematic Review Method"[1] and the National Toxicology Program's Office of Health Assessment and Translation (OHAT) "Approach for Systematic Review and Evidence Integration for Health Effects Evaluations."[2] The Navigation Guide method has been demonstrated in seven case studies in the peer-reviewed literature, and OHAT has completed several reports applying its approach to environmental health issues.[3–14] The National Academies of Sciences, Engineering, and Medicine (NASEM) has cited the OHAT and Navigation Guide systematic review methods in several reports as exemplary methods that EPA should use to evaluate environmental chemicals and inform policy and decision-making processes.[15–18]

Currently, EPA has not implemented a consistent or robust framework for systematic review across its offices and programs, and EPA's progress in implementing systematic review methods varies significantly across Agency programs. The programs that have conducted the most systematic reviews are the IRIS and TSCA programs.

The current approach to TSCA systematic review suffers from several critical flaws, including 1) failure to publish a systematic review protocol for each risk evaluation prior to conducting the evaluation, 2) inappropriate methods for study quality evaluation, 3) inappropriate exclusion of studies, 4) failure to adequately define how evidence is integrated from multiple streams to make conclusions about chemical hazards and 5) inadequate documentation of decision-making processes.

EPA's TSCA program first issued a proposed systematic review method in 2018, and after a negative February 2021 review by the NASEM, issued a new draft systematic review method in December 2021.[18] In July 2022, EPA's Science Advisory Committee on Chemicals (SACC) issued over 200 recommendations[19] for improvements to the 2021 draft TSCA method and identified numerous NASEM recommendations from February 2021 that EPA had not addressed. As of January 2025, the TSCA program has not issued a new systematic review methodology document that addresses the July 2022 SACC recommendations. The TSCA program has released several draft risk evaluations since July 2022, but the methods applied are not fully transparent and there is no consistency across these draft evaluations in how various steps in the systematic review process were conducted. For example, some draft risk evaluations issued in 2024 incorporated improved approaches to evaluating the quality of health effects studies, while other 2024 draft evaluations continued to use the flawed metrics for study evaluation that were in the 2021 draft method. Systematic review methods applied by the TSCA program have a significant likelihood of disregarding or downgrading relevant evidence and could result in underestimating health risks of environmental chemicals and pollutants.

In 2022, the IRIS program published a handbook detailing a consistent and transparent approach to systematic review, which incorporates most recommendations from the NASEM and EPA's Science Advisory Board on best practices for each step of systematic review, including evidence identification, evaluation, synthesis, and integration.[20] The IRIS program also publishes a chemical-specific protocol for each chemical it assesses and makes the protocol available for public comment before it conducts the assessment.

There are some shortcomings in the IRIS approach that need to be addressed, in particular the methods for study quality evaluation, which inappropriately exclude or downgrade some epidemiology studies (see below). However, the other components of the IRIS approach currently represent the best practices in a systematic review of chemicals and pollutants. **EPA should rely on methods outlined in the IRIS handbook, with the identified upgrades, as a model for conducting systematic review in all EPA programs.**

The adaptation of the Navigation Guide methods for systematic review of occupational exposure assessment by the World Health Organization and International Labor Organization also provides a model and case studies that can be incorporated in the TSCA program.[21]

The NASEM has further recommended that EPA should adapt these gold-standard methods to develop a systematic review-based framework for evaluating scientific evidence obtained from new approach methodologies (NAMs), which include in vitro, in silico, and certain high-throughput in vivo toxicity testing methods. NAMs data have typically been considered as "mechanistic" evidence and evaluated separately from epidemiological and in vivo animal toxicity testing data in EPA's chemical assessments. A recent NASEM report highlights the need for EPA to integrate NAMs data along with other evidence streams using gold-standard systematic review methods to improve scientific confidence in EPA's chemical hazard or risk assessments.[22]

EPA should use study assessment tools that are empirically based and should modify the current IRIS approach so that it does not inappropriately discount or exclude studies based on perceived methodological flaws.

Assessment of the "risk of bias" is a critical step in systematic review as it evaluates the internal validity of each individual study through a structured assessment of its methodological strengths and limitations (e.g., how well the study measured the exposure and outcome being evaluated). Assessing risk of bias domains provides a transparent approach to assessing methodological features that are important in assessing the overall body of evidence. EPA's IRIS and TSCA programs have adopted methods for evaluating risk of bias that apply inappropriate overall study quality ratings, and then inappropriately use these ratings to determine whether a study should be included in an assessment. This approach can result in artificially undervaluing some scientific studies that could be important in drawing conclusions. The flawed approach is described in the IRIS Handbook, which includes rating several distinct aspects of study quality (e.g., methods for exposure assessment, outcome assessment), similar to the Navigation Guide and OHAT methods.[20] However, the IRIS method then has reviewers assign an overall rating of high, medium, or low confidence, or "uninformative" for each study. Studies with an overall rating of low confidence are given less weight in drawing conclusions, and studies rated as "uninformative" are excluded from further consideration.

This approach is similar to the published "ROBINS-E" approach to study assessment, which was co-authored by several EPA scientists, and also includes an overall study rating.[23] The methods used by EPA and in ROBINS-E to assign an overall rating incorporate several implicit assumptions about the usefulness of a study and the reliability of its findings that are not scientifically supported.[24,25] The use of only "high" and "medium" quality studies can, therefore, lead to a biased evaluation of the overall body of evidence.

The NASEM has recommended that EPA not use methods that assign an overall quality score when evaluating study quality.[16,18,26] Further, studies conducted by UCSF found that the application of the IRIS method would result in the exclusion of multiple studies deemed of sufficient quality by a NASEM review.[27,28]

EPA should remove the unnecessary and misleading steps of applying an overall study quality rating from its systematic review methods. It should instead retain all relevant studies for consideration in the evidence synthesis and evidence integration steps of the systematic review, where conclusions are developed regarding the entire body of evidence.

Additionally, EPA should add financial conflicts of interest to its assessment of study quality. Industry sponsorship can bias research through various mechanisms, including how a study is designed and conducted, selective reporting of the results, skewed or incomplete analyses of study data, misleading or selective presentation of conclusions, and signaling of preferred outcomes in framing the questions to be investigated.[29–32] The NASEM 2014 review of the IRIS program's systematic review method found that "funding sources should be considered in the risk-of-bias assessment conducted for systematic reviews that are part of an IRIS assessment."[16] EPA has not yet acted on this recommendation. To ensure that EPA assessments identify the possible bias in the evidence base, study-funding source and author financial COI should be explicitly identified by the Agency when evaluating the weight of scientific evidence for hazard and risk assessment and industry sponsorship should be flagged as a risk of bias that could affect the validity of a study's findings and conclusions.

Importantly, including funding as a risk of bias domain does not mean excluding industry-sponsored studies from EPA's hazard and risk assessment; it only means documenting funding as one of many domains of potential bias and evaluating its impact on the overall quality of the body of evidence.

EPA should invest in implementation training and tools for EPA scientists, managers, and consultants across the Agency to ensure best practices and consistency in systematic review.

Implementing a robust systematic review training program at EPA is essential for achieving greater consistency across its programs for how systematic reviews are conducted. Such training would harmonize the methodologies used for conducting scientific assessments, including procedures for the preparation and publication of a systematic review protocol before an assessment is conducted, and methods for evidence identification, evaluation, synthesis, and integration for all disciplines, including hazard assessment, dose-response assessment, and exposure assessment.

A well-structured training program would promote a standardized systematic review approach across the Agency, ensuring that all EPA scientific assessments adhere to the same high standards. This consistency is crucial for maintaining the integrity of the Agency's work, ensuring that its findings and recommendations are reliable and actionable, and facilitating the sharing of knowledge, learning, and resources across Agency programs.

Trainers involved in this initiative must be highly qualified experts in systematic review methods, recognized for their expertise, and free from any financial conflicts of interest. As systematic review methods continue to evolve, ongoing training will enable the EPA to adapt to new methodologies and maintain its mission to safeguard public health.

REFERENCES

1. Woodruff TJ, Sutton P. The navigation guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environ Health Perspect*. 2014;122(10):1007-1014. doi:10.1289/ehp.1307175
2. National Toxicology Program, Office of Health Assessment and Translation. *Handbook for Conducting a Literature-Based Health Assessment. Using OHAT Approach for Systematic Review and Evidence Integration*. National Institute of Environmental Health Sciences; 2015.
3. Johnson PI, Sutton P, Atchley DS, et al. The Navigation Guide—evidence-based medicine meets environmental health: Systematic review of human evidence for PFOA effects on fetal growth. *Environ Health Perspect*. Published online 2014. doi:10.1289/ehp.1307893
4. Koustas E, Lam J, Sutton P, et al. The Navigation Guide - evidence-based medicine meets environmental health: systematic review of nonhuman evidence for PFOA effects on fetal growth. *Env Health Perspect*. 2014;122(10):1015-1027. doi:10.1289/ehp.1307177
5. Lam J, Koustas E, Sutton P, et al. The Navigation Guide - evidence-based medicine meets environmental health: integration of animal and human evidence for PFOA effects on fetal growth. *Environ Health Perspect*. 2014;122(10):1040-1051. doi:10.1289/ehp.1307923
6. Vesterinen HM, Johnson PI, Atchley DS, et al. Fetal growth and maternal glomerular filtration rate: a systematic review. *J Matern-Fetal Neonatal Med Off J Eur Assoc Perinat Med Fed Asia Ocean Perinat Soc Int Soc Perinat Obstet*. 2015;28(18):2176-2181. doi:10.3109/14767058.2014.980809
7. Johnson PI, Koustas E, Vesterinen HM, et al. Application of the Navigation Guide systematic review methodology to the evidence for developmental and reproductive toxicity of triclosan. *Environ Int*. 2016;92-93:716-728. doi:10.1016/j.envint.2016.03.009
8. Lam J, Lanphear BP, Bellinger D, et al. Developmental PBDE Exposure and IQ/ADHD in Childhood: A Systematic Review and Meta-analysis. *Environ Health Perspect*. 2017;125(8):086001. doi:10.1289/EHP1632
9. Lam J, Koustas E, Sutton P, et al. Exposure to formaldehyde and asthma outcomes: A systematic review, meta-analysis, and economic assessment. *PloS One*. 2021;16(3):e0248258. doi:10.1371/journal.pone.0248258
10. Lam J, Sutton P, Kalkbrenner A, et al. A Systematic Review and Meta-Analysis of Multiple Airborne Pollutants and Autism Spectrum Disorder. *PloS One*. 2016;11(9):e0161851. doi:10.1371/journal.pone.0161851
11. Chartres N, Cooper CB, Bland G, et al. Effects of Microplastic Exposure on Human Digestive, Reproductive, and Respiratory Health: A Rapid Systematic Review. *Environ Sci Technol*. 2024;58(52):22843-22864. doi:10.1021/acs.est.3c09524
12. National Toxicology Program. NTP monograph on the state of the science concerning fluoride exposure and neurodevelopment and cognition: a systematic review. *NTP Monogr*. 2024;(8):NTP-MGRAPH-8. doi:10.22427/NTP-MGRAPH-8
13. National Toxicology Program. *Monograph on Immunotoxicity Associated with Exposure to Perfluorooctanoic Acid (PFOA) and Perfluorooctane Sulfonate (PFOS)*. National Toxicology Program; 2016. https://ntp.niehs.nih.gov/go/mgraph04
14. National Toxicology Program. NTP monograph on the systematic review of traffic-related air pollution and hypertensive disorders of pregnancy. *NTP Monogr*. 2019;(7):NTP-MGRAPH-7. doi:10.22427/NTP-MGRAPH-7
15. National Academies of Sciences, Engineering, and Medicine. *Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals*. The National Academies Press; 2017. doi:10.17226/24758
16. National Academies of Sciences, Engineering, and Medicine. *Review of EPA's Integrated Risk Information System (IRIS) Process*. National Academies Press; 2014. doi:10.17226/18764
17. National Academies of Sciences, Engineering, and Medicine. *Progress Toward Transforming the Integrated Risk Information System (IRIS) Program: A 2018 Evaluation*. The National Academies Press; 2018. doi:10.17226/25086
18. National Academies of Sciences, Engineering, and Medicine. *The Use of Systematic Review in EPA's Toxic Substances Control Act Risk Evaluations*. The National Academies Press; 2021. https://doi.org/10.17226/25952
19. U.S. EPA. *Science Advisory Committee on Chemicals Meeting Minutes and Final Report No. 2022-2*.; 2022:83. https://www.regulations.gov/document/EPA-HQ-OPPT-2021-0414-0044
20. U.S. EPA. *ORD Staff Handbook for Developing IRIS Assessments*.; 2022.
21. Pega F, Momen NC, Gagliardi D, et al. Assessing the quality of evidence in studies estimating prevalence of exposure to occupational risk factors: The QoE-SPEO approach applied in the systematic reviews from the WHO/ILO Joint Estimates of the Work-related burden of disease and Injury. *Environ Int*. 2022;161:107136. doi:10.1016/j.envint.2022.107136
22. National Academies of Sciences, Engineering, and Medicine. *Building Confidence in New Evidence Streams for Human Health Risk Assessment: Lessons Learned from Laboratory Mammalian Toxicity Tests*. National Academies Press; 2023. doi:10.17226/26906
23. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). *Environ Int*. 2024;186:108602. doi:10.1016/j.envint.2024.108602
24. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*. 1999;282(11):1054-1060. doi:10.1001/jama.282.11.1054
25. Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 1st ed. Wiley; 2019. doi:10.1002/9781119536604
26. National Academies of Sciences, Engineering, and Medicine. *Review of U.S. EPA's ORD Staff Handbook for Developing IRIS Assessments*.; 2022. https://doi.org/10.17226/26289
27. Eick SM, Goin DE, Chartres N, Lam J, Woodruff TJ. Assessing risk of bias in human environmental epidemiology studies using three tools: different conclusions from different tools. *Syst Rev*. 2020;9(1):249. doi:10.1186/s13643-020-01490-8
28. Chartres N, Choi G. Assessing risk of bias in human environmental epidemiology studies using risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). In preparation.
29. Odierna DH, Forsyth SR, White J, Bero LA. The Cycle of Bias in Health Research: A Framework and Toolbox for Critical Appraisal Training. *Account Res*. 2013;20(2):127-141. doi:10.1080/08989621.2013.768931
30. Fabbri A, Lai A, Grundy Q, Bero LA. The Influence of Industry Sponsorship on the Research Agenda: A Scoping Review. *Am J Public Health*. 2018;108(11):e9-e16. doi:10.2105/AJPH.2018.304677
31. Psaty BM, Prentice RL. Minimizing bias in randomized trials: the importance of blinding. *JAMA*. 2010;304(7):793-794. doi:10.1001/jama.2010.1161
32. Psaty BM, Kronmal RA. Reporting mortality findings in trials of rofecoxib for Alzheimer disease or cognitive impairment: a case study based on documents from rofecoxib litigation. *JAMA*. 2008;299(15):1813-1817. doi:10.1001/jama.299.15.1813