MONDAY, APRIL 27, 2026

Everything You Wanted to Know About Bias in Environmental Health Research (But Were Afraid to Quantify)

SESSION CHAIRS

David A. Savitz
Brown University & HEI Research Committee

Evi Samoli
National and Kapodistrian University of Athens, Greece & HEI Research Committee

SPEAKERS

Matthew Fox
Boston University

In conversation with

David Miller
US EPA (retired)

Risk of bias assessments, evidence synthesis and what has triangulation to do with it?

Evi Samoli, National and Kapodistrian University of Athens

David Savitz, Brown University School of Public Health

Both members of the HEI Research Committee

Health Effects Institute

Key sources of bias in (observational) environmental epidemiology studies

- Residual confounding
- Exposure misclassification/measurement error
- Not selection bias traditionally

Addressing bias in reviews and evidence synthesis

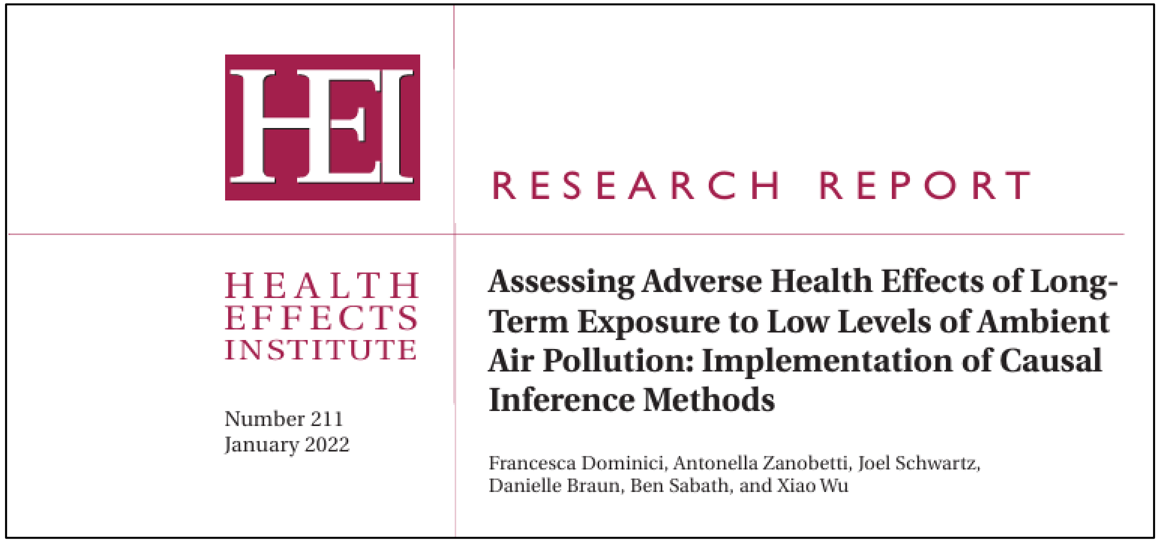
Study specific biases are fed into systematic reviews that use narratives procedures or “algorithmic” tools to assess bias in included studies

Health Effects Institute

Quantitative bias assessment in individual studies

- Intends to provide unbiased adjusted estimates
- Depends on "validation" data and assumptions
- Methods developed on causal inference by enhancing existing ones
- Directed acyclic graphs are supported
- Sensitivity analysis –what happens under different assumptions/scenarios
- Previous literature to describe the amount of error (e.g., direction in confounding, selection proportion for selection bias) is proposed to inform bias parameters and adjust effect estimates
- Previous HEI Requests for Applications have called for attention to causal inference and quantification of measurement error in exposure assessment

Causal inference – targeting confounding



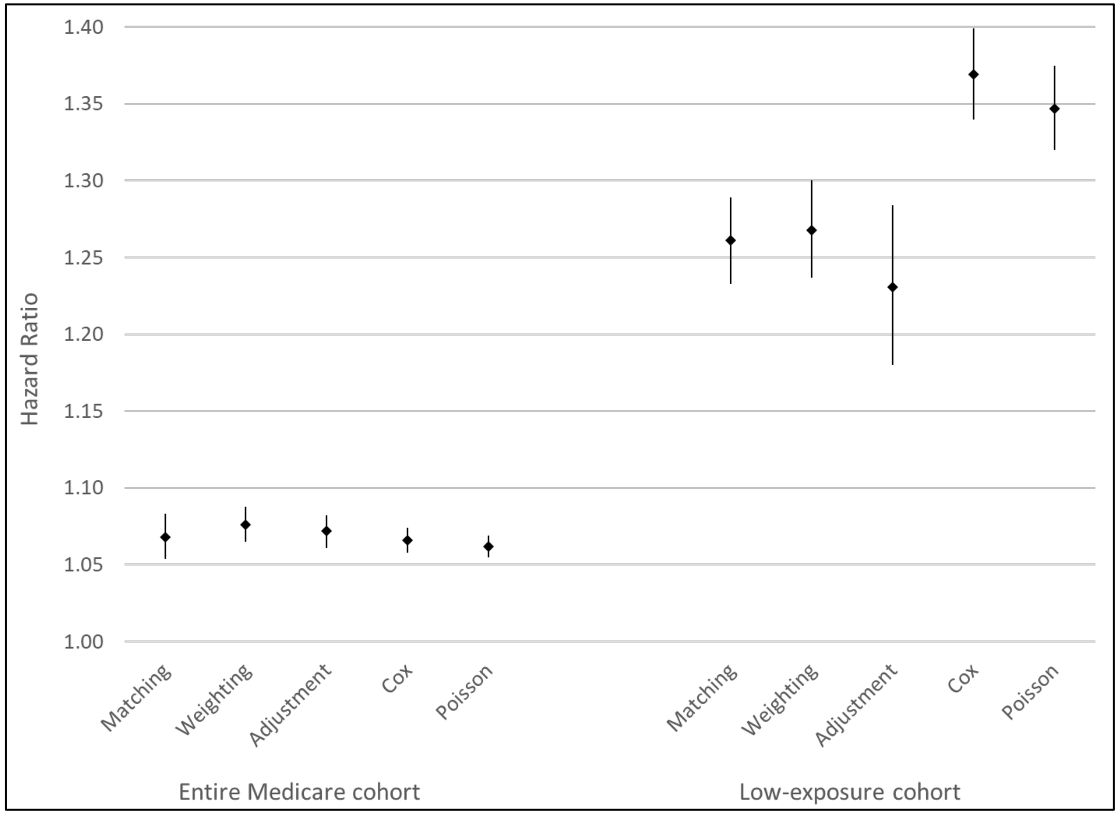
RESEARCH REPORT

HEALTH EFFECTS INSTITUTE

Number 211
January 2022

Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Implementation of Causal Inference Methods

Francesca Dominici, Antonella Zanobetti, Joel Schwartz, Danielle Braun, Ben Sabath, and Xiao Wu

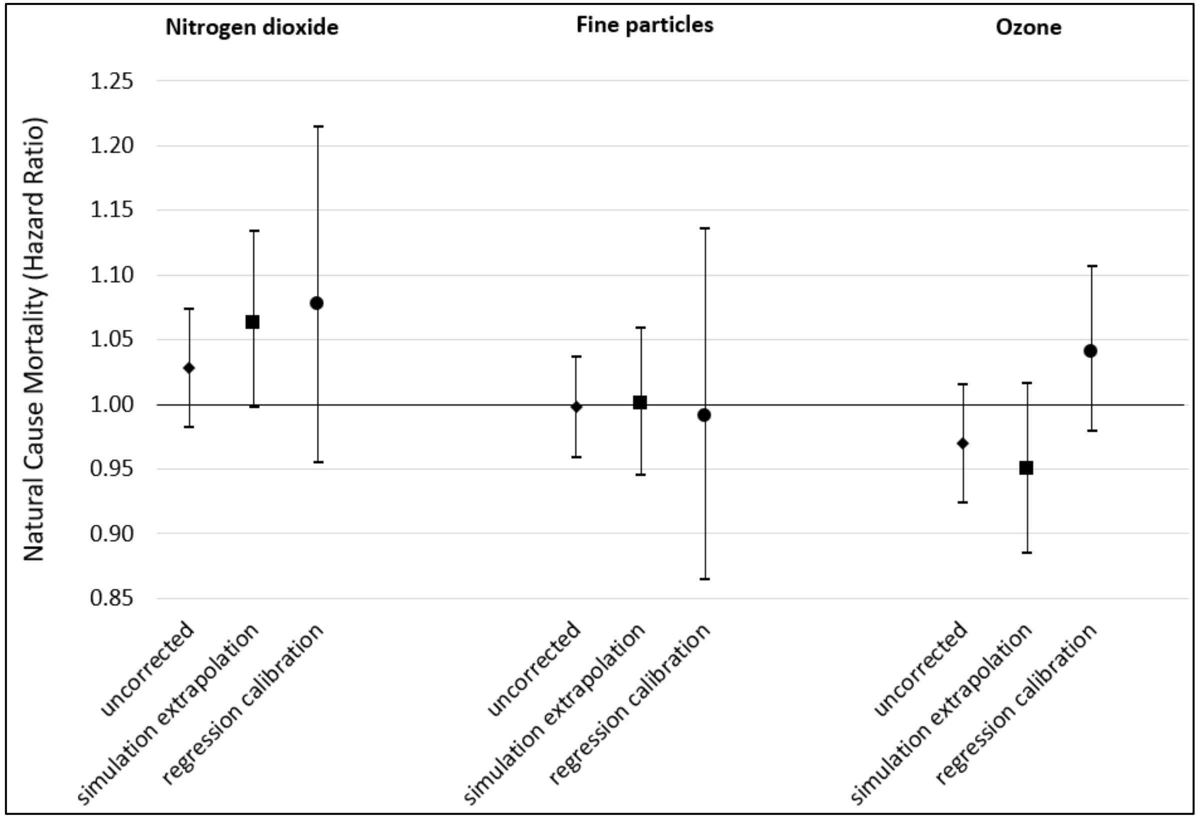


Statement Figure. Associations between longer-term exposures to $PM_{2.5}$ and all-cause mortality among enrollees in the full Medicare cohort and low-exposure cohort. Hazard ratios (calculated per 10-µg/m^3 increase in $PM_{2.5}$ exposure) and 95% confidence intervals were estimated using three causal inference approaches with generalized propensity scores (matching, weighting, and adjustment) and two traditional approaches (Cox and Poisson regression).

Health Effects Institute

Adjusting for exposure measurement error in health estimates in the UK biobank study

HEI
Health Effects Institute
NUMBER 227
MAY 2025
RESEARCH REPORT
Investigating the Consequences of Measurement Error of Gradually More Sophisticated Long-Term Personal Exposure Models in Assessing Health Effects: The London Study (MELONS)
Klea Katsouyanni, Dimitris Evangelopoulos, Dylan Wood, Benjamin Barratt, Hanbin Zhang, Heather Walton, Audrey de Nazelle, Vangelis Evangelou, Sean Beevers, Barbara Butland, Evangelia Samoli, and Joel Schwartz

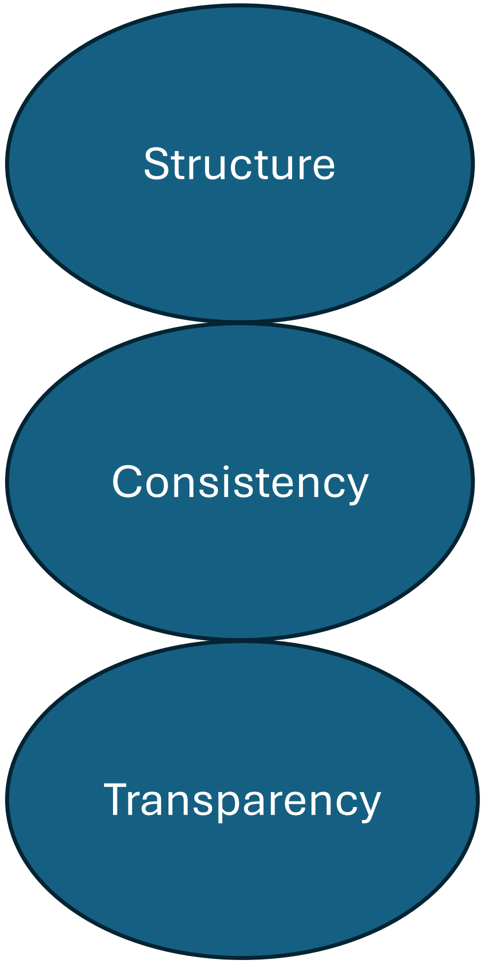


Statement Figure: Associations between estimated annual average pollutant concentrations from a spatiotemporal exposure model and natural cause mortality among UK Biobank cohort participants. Data shown are hazard ratios and 95% confidence intervals estimated per interquartile range increases in exposure. Results are from the main model, which was adjusted for all available individual-level covariates.

HEI
Health Effects Institute

There is a long tradition in environmental health of using frameworks for evidence synthesis

- Four-level hierarchy for classifying the strength of the evidence of the US Surgeon General on smoking and health (1964, 1994, 2004)
- Bradford Hill aspects (Hill et al. 1965)
- Combination of multiple data streams; human, animal, mechanistic from IARC (Samet et al. 2000)
- US EPA weight-of-evidence approach to determine causality in the Integrated Science Assessments (Owens et al. 2017; US EPA Preamble 2015)



HEI
Health Effects Institute

Addressing bias in reviews and evidence synthesis

- Wide array of bias assessment methods available
- This includes developing the process for the systematic review of key biases in individual studies and incorporating the bias assessment into the evidence synthesis
- Two distinct approaches to assess bias are currently in use:, which can be labelled as **“algorithmic”** tools and **triangulation** approaches

HEI Health Effects Institute

The HEI Traffic Review applied a GRADE-type approach

HEI
SPECIAL REPORT
HEALTH EFFECTS INSTITUTE

Systematic Review and Meta-analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution

Number 23
June 2022

HEI Panel on the Health Effects of Long-Term Exposure to Traffic-Related Air Pollution

Commentary

A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP11532.

Evidence Synthesis of Observational Studies in Environmental Health: Lessons Learned from a Systematic Review on Traffic-Related Air Pollution

Hanna Boogaard,[1] Richard W. Atkinson,[2] Jeffrey R. Brook,[3] Howard H. Chang,[4] Gerard Hoek,[5] Barbara Hoffmann,[6] Sharon K. Sagiv,[7] Evangelia Samoli,[8] Audrey Smargiassi,[9] Adam A. Szpiro,[10] Danielle Vienneau,[11,12] Jennifer Weuve,[13] Frederick W. Lurmann,[14] and Francesco Forastiere[15]

Key elements of the OHAT approach	Suggested improvements by the Traffic Review Panel[a]
Evidence synthesis	
Use a GRADE-type approach to assess confidence in the quality of the body of evidence.	Complement the GRADE-type assessment with a broader, "narrative" approach to maximize what can be learned from observational studies in environmental health.
Assign an initial low or moderate level of confidence to all types of observational studies.	Consider that in environmental health, where randomized controlled trials are generally not appropriate, some observational studies can offer high-confidence evidence.
Assess the statistical heterogeneity of results and downgrade the confidence rating if substantial heterogeneity is found.	Sources of heterogeneity can strengthen or weaken the confidence in the evidence and should be carefully explored. Some heterogeneity is expected in studies of the health effects of environmental exposures, due to different populations, locations, and study settings. Consider primarily the direction of the effect estimate rather than its magnitude. Because different methods and study designs that generate similar findings may strengthen the confidence, a separate upgrading factor for consistency should be added to GRADE.[20]
Assess publication bias using Egger's test and funnel plots and downgrade accordingly.	Publication bias is not necessarily expected when large and collaborative (multicenter) studies comprise most of the evidence and/or if evidence has accrued over several decades. Use additional approaches to explore the possibility of publication bias.
Risk of bias in individual studies	
Compare study with randomized controlled trials or hypothetical target experiment as ideal study.	Do not consider randomized controlled trials as ideal study.
Evaluate bias in different domains (e.g., confounding, selection bias, measurement error).	Focus on identifying the most likely influential sources of bias—based on methodologic and subject matter expertise—classifying each study on the basis of how effectively it has addressed each potential bias and determining whether results differ across studies in relation to each hypothesized source of bias.
Rate potential biases (e.g., low, moderate, high) using a risk of bias tool.	Rate biases considering the suggestions in the row above. Those ratings should not be used to dismiss studies based on bias but to conduct sensitivity analyses comparing findings from studies of high bias and low/moderate bias.

Note: GRADE, Grading of Recommendations Assessment, Development, and Evaluation; OHAT, U.S. Office of Health Assessment and Translation.

[a] HEI appointed an expert Panel to systematically evaluate the epidemiological evidence regarding the associations between long-term exposure (months to years) to TRAP and selected adverse health outcomes. The Panel consisted of 13 experts in epidemiology, exposure assessment, and statistics at institutions in North America and Europe, and are co-authors of this paper.

CONTENTS

Search for algorithms

- Understandable desire to work from a predefined, prescribed script
- Avoid mistakes, controversy, vulnerability to criticism, confusion
- Obligatory for regulatory, legal, and policy applications
  - -Withstand challenge from opponents
  - -Acceptable basis for judgment to those with limited technical expertise

HEI Health Effects Institute

Issues with using algorithms

- Evidence is often inconsistent, complex, incomplete and algorithms may gloss over that
- Ultimately, informed judgment is required
- Attempts to apply generic methods independent of subject matter expertise
- Algorithms create illusion of objectivity and accuracy ("it's right because the algorithm says so")

AJE
American Journal of Epidemiology
© The Author(s) 2019. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.
Vol. 188, No. 9
DOI: 10.1093/aje/kwz131
Advance Access publication: May 30, 2019

Special Article

The Problem With Mechanistic Risk of Bias Assessments in Evidence Synthesis of Observational Studies and a Practical Alternative: Assessing the Impact of Specific Sources of Potential Bias

2019

David A. Savitz*, Gregory A. Wellenius, and Thomas A. Trikalinos

Commentary

2020

ehp Environmental Health Perspectives
A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP6980.

Risk of Bias Assessments and Evidence Syntheses for Observational Epidemiologic Studies of Environmental and Occupational Exposures: Strengths and Limitations

Kyle Steenland,[1] M.K. Schubauer-Berigan,[2] R. Vermeulen,[3] R.M. Lunn,[4] K. Straif,[5,6] S. Zahm,[7] P. Stewart,[8] W.D. Arroyave,[9] S.S. Mehta,[4] and N. Pearce[10]

HEI
Health Effects Institute

Meta-analyses to generate pooled effect estimates

- Studies included in a meta-analysis often represent a **subset** of the available evidence.
- Presumes random error is primary source of uncertainty, pool to increase precision
- Requires sufficient homogeneity of study methods to combine results, rarely true
- Forgoes benefit of informative variation in study methods and results
- Marketing features: Well-defined procedures; Large volume of studies , and Precise pooled estimates

Do pooled estimates from meta- analyses of observational epidemiology studies contribute to causal inference?

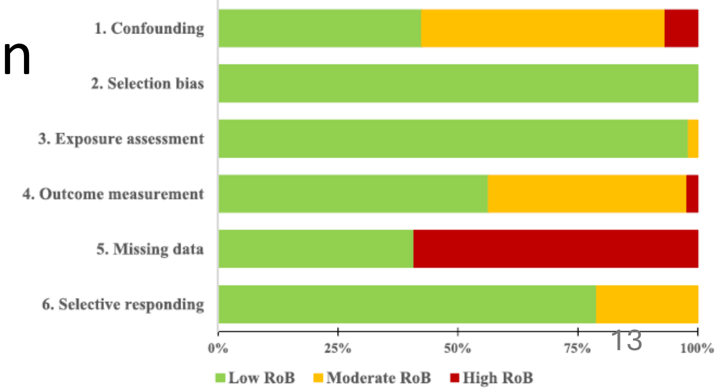
Editorial > Occup Environ Med. 2021 Sep;78(9):621-622. doi: 10.1136/oemed-2021-107702. Epub 2021 Jun 22.

David A Savitz ,[1] Francesco Forastiere[2]

HEI Health Effects Institute

Risk of bias scales

- Scores to characterize the study's quality
  - -Comprehensive checklist of methods
  - -Declaration of vulnerability to bias
- Common examples: GRADE, Newcastle-Ottawa, ROBINS-I
- Assigns equal value to each source of bias regardless of subject matter
- Risk of bias indicates the **potential** for bias rather than the direction and magnitude of actual biases.

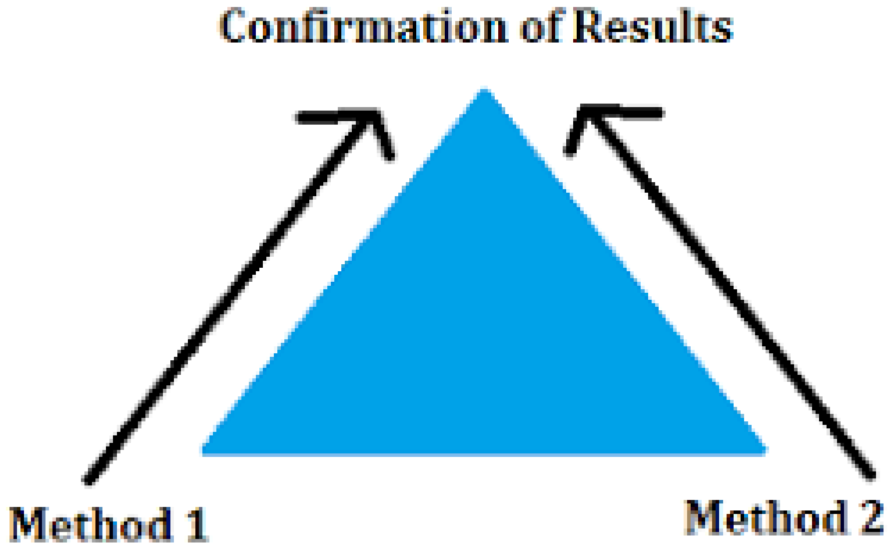


Health Effects Institute

Triangulation for bias assessment and evidence synthesis

BOX 1-1
Definition of Triangulation

"The practice of strengthening causal inferences by integrating results from several different approaches, where each approach has different (and assumed to be largely unrelated) key sources of potential bias."
—Lawlor et al. (2016)



Approaches to causal inference

Triangulation in aetiological epidemiology

Debbie A Lawlor,[1,2,*] Kate Tilling[1,2] and George Davey Smith[1,2]

[1]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK and [2]School of Social and Community Medicine, University of Bristol, Bristol, UK

"Triangulation integrates data from different methods, designs, and theoretical approaches, as well as data with different and unrelated sources of potential bias, to determine if findings converge on one conclusion"

“Algorithmic” tools

- Systematic review
  - -“a methodology that increases the scientific rigor, transparency, and objectivity in a risk evaluation by using a predefined, multi-step processes to identify, critically assess, and synthesize evidence” (from EPA)
- World Health Organization – GRADE
- National Toxicology Program – OHAT
- Environmental Protection Agency – Systematic review protocol for TSCA, IRIS
- Navigation Guide

HEI Health Effects Institute

Appealing features of quantitative bias assessment

- Systematic, quantitative application of triangulation
- Deep dive into critical issues for a specific topic, driven by subject matter expertise
- Explicit "what if?" exercise, links assumptions and implications
- Identifies drivers of uncertainty to prioritize future research
- Tool for informing judgment rather than a substitute for judgment

HEI
Health Effects Institute

Tension between prescribed methods and strategic improvisation

- Reviews may be constrained by predefined protocols for regulatory purposes
- Subject matter experts needed to identify critical questions and methods to address those challenges
- Perhaps quantitative bias assessments could be complementary approach for algorithm-based assessments

MONDAY, APRIL 27, 2026

Everything You Wanted to Know About Bias in Environmental Health Research (But Were Afraid to Quantify)

SESSION CHAIRS

David A. Savitz

Brown University &
HEI Research Committee

Evi Samoli

National and Kapodistrian
University of Athens, Greece &
HEI Research Committee

SPEAKERS

Matthew Fox
Boston University

In conversation with

David Miller
US EPA (retired)