TOP TAKEAWAYS

Probability of Causation – A case study comparison of Benzene vs. Formaldehyde and acute myeloid leukemia (AML)

CENTER FOR TRUTH IN SCIENCE

Modern approaches to measuring health risks are limited by reliable methods to quantify the extent to which public, or private, health interventions can change the likelihood of specific adverse health outcomes. A major reason for this is that human health research and the field of epidemiology rely on observational data, as it is often unethical to conduct clinical trials on humans. This limits the ability to identify or characterize associations between exposure and response variables, opening the door to non-causal explanations based on chance, bias, or confounders. Ultimately, decision makers can't effectively determine how risk factors contribute to causation, or how policy changes can improve or worsen health outcomes.

The primary purpose of this study is to explain, advance, and apply interventional probability of causation (IPoC) research methods. The work was conceived and conducted by Tony Cox, PhD, Ken Mundt, PhD, and associates. The paper defines IPoC as the change in probability of a disease (or other harm) for an individual over a specified time interval, caused by a change in exposure to a substance or an intervention, as predicted by a fully specified causal model.

A secondary goal of this case study is to illustrate the practical application of IPoC methods by comparing and contrasting IPoC calculations and uncertainty characterizations for benzene and formaldehyde exposures as potential causes of myeloid leukemia. The paper also discusses modern causal AI and machine learning techniques as enabling technologies for practical IPoC calculation methods.

HIGHLIGHTS FROM THE PUBLISHED WORK:

- Summary of statistical and causal knowledge challenges in estimating and validating IPoCs from available epidemiological data and realistically incomplete observational evidence and knowledge. Summary of challenges in characterizing remaining uncertainties, including uncertainty about correct causal models and partial evidence.
- Constructive solutions to these challenges, building on ideas from modern causal AI and machine learning techniques that are brought together and applied to support practical calculation, validation, and uncertainty characterization for IPoCs.
- An exposition of recent technical innovations for calculating individual-level IPoCs, estimating population distributions of IPoCs, and characterizing uncertainty and variability in both individual- and population-level IPoCs for a broad audience of toxicologists, epidemiologists, risk analysts, regulators, policy analysts, and litigators.
- Illustration of a practical application of IPoC by comparing IPoC calculations and uncertainty calculations for benzene and formaldehyde as potential causes of acute myeloid leukemia.

Summary of the Case Study

- The paper describes recent concerns around whether formaldehyde causes acute myeloid leukemia, identifying key questions about the biological basis for a link between the two.
- The authors outline the need for quantitative risk assessment methods that provide conceptually and statistically sound analyses for both verifying claims of causality as well as quantifying the contributions of exposures to individual-level and population cancer risks.
- Outline of what is known and conjectured about modes of action, pharmacokinetics, and pharmacodynamics for exogenous formaldehyde and benzene and summarizes recent reviews of evidence and risk of bias for causal associations between exposures to benzene and formaldehyde and risk of acute myeloid leukemia (AML).
- The paper also outlines a definition of these chemicals causing AML, both theoretically (unrepaired dysregulation of cell replication) and pragmatically (AML clinical diagnosis).
- The paper discusses the statistical challenges and the causal and empirical knowledge gaps that must be addressed to answer whether realistic levels of exposure to each chemical can increase occupational risk of AML.
- To address the practical questions that risk analysts, regulators, additional public health professionals and attorneys must deal with, the paper focuses on the following: How much, if at all, would an individual's risk (age-specific hazard function) change if exposure levels were changed, or if (counterfactually) they had been changed in the past to be different from the real exposures?

1. **How much, if at all, would a target population's risk (average cases per year and frequency distribution of individual risks) change in the future if exposure levels were changed?**

2. **What direct effects have past changes in exposures had on individual and population risks? (This is sometimes referred to as the accountability problem for regulatory benefits).**

3. **How sure can we be about the answers to these questions, given the realistically limited evidence and knowledge currently available? To what extent do available knowledge and data support conclusions that reducing formaldehyde exposure reduces myeloid leukemia by quantifiable amounts?**

Case Study Findings

- Prolonged, high-intensity exposure to benzene can increase risk of AML.
- No causal pathway leading from formaldehyde exposure to increased risk of AML was found.

"I am thrilled to see our work on IPoC published, offering a new, objective method for evaluating causal relationships in epidemiology. I think IPoC introduces a paradigm shift in public health risk assessment, moving away from untestable assumptions to empirically verifiable causal claims. This can transform how we assess and manage health risks, helping to ensure that policies are grounded in scientifically robust evidence instead of unverified modeling assumptions, benefiting both policymakers and the public."

– Louis Anthony Cox, PhD, Clinical Professor, Department of Biostatistics and Informatics, University of Colorado, Boulder/ Honorary Full Professor, University of Colorado, Denver

"The Center for Truth in Science is excited to support this important work and impressive publication, which is the culmination of more than two years of collective effort by a dedicated team of expert volunteer researchers, Center staff, and graduate students. We believe this powerful method for identifying sources of risk will enable researchers to more accurately determine causation when relying on observational studies. This case study of the chemicals benzene and formaldehyde represents a compelling example of how advanced analytical techniques and artificial intelligence can help decision makers solve challenging problems in the real world."

– Jacob Traverse, President & CEO, Center for Truth in Science

CONCLUSIONS FROM THE PAPER:

- The IPOC approach can differentiate between likely and unlikely causal factors and can provide useful upper bounds for some exposures and diseases of practical importance.
- For causal factors, IPOC can help estimate the quantitative impacts on health risks of reducing exposures, even where there is incomplete mechanistic evidence and uncertain knowledge of individual-level exposure-response.