



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Elagali, A;Drummond, K;Winterton, A;Mulders, Y;Symeonides, C;Seewoo, B;Andersen, G;Ponsonby, AL;Dunlop, S

Title:

Looking inside the black box - Hybrid epidemiology approaches to identify causal inferences

Date:

2025-07-15

Citation:

Elagali, A., Drummond, K., Winterton, A., Mulders, Y., Symeonides, C., Seewoo, B., Andersen, G., Ponsonby, A. L. & Dunlop, S. (2025). Looking inside the black box - Hybrid epidemiology approaches to identify causal inferences. Cambridge Prisms Plastics, 3, <https://doi.org/10.1017/plc.2025.10011>.

Persistent Link:

<https://hdl.handle.net/11343/362939>

License:

[CC-BY](#)

Perspective

Cite this article: Elagali A, Drummond K, Winterton A, Mulders Y, Symeonides C, Seewoo B, Andersen G, Ponsonby A-L and Dunlop S (2025). Looking inside the black box – hybrid epidemiology approaches to identify causal inferences. *Cambridge Prisms: Plastics*, **3**, e19, 1–6
<https://doi.org/10.1017/plc.2025.10011>

Received: 23 December 2024

Revised: 03 June 2025

Accepted: 20 June 2025

Keywords:



plastic chemicals; molecular epidemiology; machine learning; human health and wellbeing

Corresponding author:

Ahmed Elagali;

Email: aelagali@minderero.org

Looking inside the black box – hybrid epidemiology approaches to identify causal inferences

Ahmed Elagali^{1,2,3} , Katherine Drummond³, Adriano Winterton⁴, Yannick Mulders¹, Christos Symeonides¹ , Bhedita Seewoo^{1,2}, Gro Andersen⁴, Anne-Louise Ponsonby³ and Sarah Dunlop^{1,2}

¹Plastics and Human Health, Minderoo Foundation, Perth, WA, Australia; ²School of Biological Sciences, The University of Western Australia, Perth, WA, Australia; ³Florey Institute of Neuroscience and Mental Health, University of Melbourne, Parkville, VIC, Australia and ⁴Division of Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway

Abstract

Plastic chemicals are numerous and ubiquitous in modern life and pose significant risks to human health. Observational epidemiological studies have been instrumental in identifying consistent and statistically significant associations between exposure to certain chemicals and adverse health outcomes. However, these studies often fail to establish causality due to the complexity of real-world chemical mixtures, confounding factors, reverse causation, and study designs that lack measures reflecting underlying genetic and cellular mechanisms indicating causal pathways to harm. Addressing these limitations requires moving beyond traditional ‘black-box’ epidemiology, which mainly focuses on the strength of associations. We propose adopting hybrid epidemiological methodologies that incorporate genetic susceptibility and molecular mechanisms to uncover biological pathways, combined with machine learning and statistical analysis of chemical mixtures, to strengthen the causal evidence linking exposure to harm. By integrating observational multi-omics data with experimental and mechanistic models, hybrid epidemiology offers a transformative path to improve causal evidence and public health interventions. In addition, machine learning and statistical methods provide a more nuanced understanding of the health effects of exposures to plastic chemical mixtures, facilitating the identification of interactions within chemical mixtures and the influence of biological pathways. This paradigm shift is critical addressing the complex challenges of plastic exposure and protecting human health.

Impact statement

This work addresses a critical gap in environmental health research: the inability of traditional observational epidemiology to fully elucidate the causal and interactive effects of plastic chemical exposures. By proposing a hybrid epidemiological framework that incorporates underpinning genetic susceptibility and molecular mechanisms, advanced statistical methods, machine learning and experimental tools, this perspective charts a path towards a more comprehensive understanding of exposure-disease relationships. This approach not only reveals the biological plausibility of the findings but also provides robust evidence for policymakers to regulate hazardous chemical mixtures effectively.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

 Cambridge Prisms

 CAMBRIDGE UNIVERSITY PRESS

Introduction

For decades, observational epidemiological studies have formed the basis for our understanding of environmental health risks, particularly those posed by plastic chemicals (Grandjean and Landrigan, 2006; Woodruff et al., 2011; Landrigan et al., 2023). These studies have provided critical information on the associations between exposure to numerous ubiquitous chemicals, in particular during early life, and various health conditions, including the aetiologies of adverse perinatal and childhood outcomes, endocrine disruption, adverse neurodevelopmental outcomes, obesity, diabetes, and cancer (Symeonides et al., 2024, and references therein). Despite their significant contributions, observational epidemiological studies often face limitations in establishing causal relationships. Such studies typically focus on the health impacts of individual chemicals, an approach that may not fully capture the complexities of endocrine-disrupting chemicals (EDCs) such as phthalates and bisphenols. These chemicals can exert effects through shared and/or overlapping mechanisms on endocrine systems, such as binding to specific hormone receptors in target tissues or interfering with

biotransformation and excretion of endogenous hormones, enabling multiple chemicals to act additively or synergistically at lower concentrations to produce outcomes that would require higher concentrations if acting alone (Kortenkamp, 2008; Al-Gubory, 2014; Darbre, 2022). The inability to account for the full exposome (Wild, 2005), encompassing all chemical and non-chemical factors that influence human well-being throughout life, e.g., lifestyle, socio-economic status, diet, adverse life-events and psychosocial factors, can bias findings further and obscure direct causal links between chemical exposure and health outcomes later in life. In addition, the long latency periods associated with conditions such as cancer or reproductive disorders make it difficult to attribute these health impacts to specific chemical exposures many years earlier in life (Amolegbe et al., 2022). Exposures to even very low concentrations during the critical developmental window in the foetal and childhood life stages could be related to a programming effect with an increased risk of disease later in life (Lanphear, 2015). Furthermore, as EDCs often show non-linear exposure–response relationships, assuming only linear relationships will obscure the true impact of these chemicals. These challenges highlight the need for a hybrid epidemiological approach that combines advances in mixture analysis, machine learning and molecular methods to improve causal inference. Integrating experimental models (e.g., *in vitro*, *in vivo*, *in silico* and mechanistic toxicology) with epidemiological data enables testing of biological plausibility, dose–response and temporality. This alignment helps move from statistical associations to identifying mechanistic pathways linking EDC exposure to health outcomes.

The challenge faced by observational studies to establish causality closely resembles those outlined in the Bradford Hill criteria, such as demonstrating strength, consistency, biological plausibility and temporality in observed associations (Woodside III and Davis, 2012; Fedak et al., 2015). Traditional epidemiology often produces small effect sizes (Symeonides et al., 2024), which poses inconclusive and weak confidence in causal links to human harm based on Bradford Hill's criterion of strength of association (Woodside III and Davis, 2012; Fedak et al., 2015). While the consistency across observational epidemiological studies, another Bradford Hill criterion, provides some support for causality, the strength of the association is the primary factor in considering 'cause' (Woodside III and Davis, 2012; Fedak et al., 2015). In contrast, adopting a hybrid epidemiology approach that explores the underlying biological mechanisms is likely to reveal larger effect sizes expressed as several-fold changes (Caporale et al., 2022; Symeonides et al., 2024; Elagali et al., 2025), which increases the evidence for causation. Moving beyond traditional observational studies, this perspective advocates for innovative hybrid epidemiological strategies that will equip researchers with powerful tools to uncover causal relationships and assess the health impacts of complex plastic chemical mixtures. In this perspective, we offer key recommendations to support the shift towards strengthening causal evidence, with substantial potential to drive progress in precision medicine and improve population health outcomes.

Beyond one-chemical-at-a-time: mixture analysis and machine learning methods

Traditional epidemiological approaches are sometimes limited by their focus on assessing health risks associated with single chemical exposure, which fail to capture the complexity of real-world scenarios and treat environmental exposures as isolated events. This reductionist approach has resulted in a limited understanding of

the health impacts associated with chemical mixtures and presents a significant regulatory challenge. In practice, humans are exposed to numerous environmental chemicals throughout their lifespan, often concurrently, as shown by national human biomonitoring programs (Woodruff et al., 2011; Stanfield et al., 2024). These exposures are frequently correlated, may show non-linear dose–response relationships and can lead to cumulative or interactive health effects. Traditional regression models struggle with such complexity due to multicollinearity, limited power to detect interactions and challenges in selecting meaningful variables from high-dimensional data. To address this, advanced statistical techniques have been developed to more accurately model chemical mixtures. These approaches differ in assumptions, interpretability, ability to account for correlation and capacity to detect non-linear or interactive effects. Advancing methods to quantify the disease risks posed by chemical mixtures could help identify modifiable exposures, enabling targeted public health interventions and prevention strategies.

In adapting this approach, epidemiologists must consider that the current research on plastic chemicals is constrained by a focus on substances already known to be hazardous or measurable using existing methods, a limitation often referred to as the 'street-light effect' (J. M. Braun et al., 2016). With the increase in the use of non-targeted approaches to quantify chemical exposures, mixture studies can rely on these methods to prioritize which chemicals are included in mixture-based statistical approaches i.e., '*in silico*' prioritization. To support this advancement, it is crucial to develop efficient high-throughput screening assays that target key endpoints and account for the toxicity of both individual chemicals and their mixtures. G. Braun et al. (2024) demonstrated this approach by analysing the neurotoxic effects of hundreds of chemical mixtures extracted from the blood of pregnant women. Using unbiased extraction methods, automated target screening and high-throughput *in vitro* neurotoxicity assays, Braun et al. uncovered the impact of complex organic mixtures on neurite development and demonstrated how *in silico* prioritization can effectively narrow down chemicals of concern for neurodevelopmental endpoints.

As the study of chemical mixtures remains an evolving field of research, there is currently no established consensus on the most appropriate statistical or machine learning methods to investigate the health effects of chemical mixtures in epidemiological studies (Bobb et al., 2015; Lazarevic et al., 2019; Maitre et al., 2022; Miller and Consortium, 2025). In recent years, numerous methods have been introduced to address the complexities associated with multiple exposures and their interactions (Sun et al., 2013; Forns et al., 2016; Stafoggia et al., 2017). These include techniques for variable selection, shrinkage and grouping of correlated variables, such as the least absolute shrinkage and selection operator (*LASSO*; Sun et al., 2013), elastic net (Lenters et al., 2016) and adaptive elastic net (Zou and Zhang, 2009). Other approaches include dimension reduction methods like principal component analysis (Yang et al., 2013) and partial least squares (Sun et al., 2013), as well as Bayesian frameworks such as Bayesian model averaging (Bobb et al., 2011). Notably, two methods specifically designed for mixture analyses in environmental epidemiology warrant attention: Weighted Quantile Sum Regression (WQSR, Carrico et al., 2015) and Bayesian Kernel Machine Regression (*BKMR*, Bobb et al., 2018). WQSR and *BKMR* provide robust measures of the health effects of mixtures, with *BKMR* offering additional advantages, such as accounting for non-linearity and interaction in multivariate exposure–response relationships. Despite their strengths, all of these methods exhibit some limitations, including instability in model selection

(shrinkage approaches), difficulties in interpreting latent variables (dimension reduction) and the computational intensity of Bayesian models. Furthermore, their application to large and heterogeneous exposome datasets – comprising diverse variable types such as -omics data and mixed categorical and continuous variables – remains limited.

We do not advocate for a rigid, one-size-fits-all approach; however, for complex research questions exploring the health effects of multiple exposures, a multi-staged statistical strategy is recommended. For example, an initial step could involve utilizing methods like *LASSO* as a pre-processing tool (Sun et al., 2013). This preliminary step helps refine the dataset and enhance model performance before implementing Bayesian approaches, such as: Bayesian Hierarchical Models (Gelman et al., 1995), which can handle data with nested or multilevel structures (e.g., repeated measures or multi-cohort studies); *BKMR*, which is particularly effective for modelling non-linear and interactive effects of correlated exposures; or Bayesian Additive Regression Trees (Chipman et al., 2010), a flexible ensemble method that captures complex, non-parametric relationships and has been used to identify high-order interactions in high-dimensional datasets. An alternative to multistage analysis is the use of machine learning (ML) techniques, which are increasingly applied in environmental health to manage high-dimensional exposure data and uncover patterns missed by regression models. ML is well-suited for modelling non-linear relationships and interactions in chemical mixtures. The choice of algorithm depends on the data structure and research aims. For example, neural networks are particularly suited for modelling complex, multilayered interactions in large datasets, while support vector regression (*SVR*) performs well in small- to medium-sized datasets with non-linear boundaries. Ensemble learning methods like gradient boosting (e.g., *XGBoost*, *LightGBM*, *CatBoost*) and adaptive boosting (*AdaBoost*) build highly predictive models and are especially effective for structured data with correlated features (see, for example, Argyri et al., 2024; Guimbaud et al., 2024, and references therein). However, a significant challenge with machine learning models lies in their lack of interpretability. For example, models like gradient boosting, which consist of numerous decision trees, make it difficult to determine how individual features contribute to a specific prediction. To address this limitation, Shapley Additive Explanations (*SHAP*) has emerged as a valuable tool, offering clear and detailed explanations of ML predictions (Lundberg, 2017). Ultimately, the choice of method should reflect the dimensionality of the data, sample size, confounding and research objectives. As environmental health data grow in complexity, so does the relevance of machine learning for analysis and interpretation.

Uncovering mechanisms: -omics and molecular pathways in the shift towards causal inference

As research on plastic chemicals advances, there is an increasing emphasis on the need to identify robust causal evidence linking these exposures to human health outcomes. While correlations between exposure and health effects may represent genuine associations, they do not inherently demonstrate causation. A promising strategy to strengthen causal inference involves the integration of evidence from multiple approaches, each with distinct and unrelated sources of potential bias – a method referred to as triangulation of evidence. For instance, combining observational epidemiological data with experimental or quasi-experimental

approaches, such as toxicological studies, *in vitro* systems and mechanistic models, facilitates the identification of biologically plausible pathways linking plastic exposure to health outcomes. Similarly, the integration of ‘multi-omics’ technologies, encompassing comprehensive biological domains such as genomics, transcriptomics and proteomics, offers a robust framework for uncovering novel causal mediators of disease and revealing biological insights that may remain undetected through single-omics analyses (Hasin et al., 2017; Karczewski and Snyder, 2018; Miller and Consortium, 2025). Since each method operates under unique assumptions and exhibits distinct strengths and limitations, consistent findings across diverse methodologies significantly reduce the likelihood of artefactual results, thereby strengthening the inference of causality (see Figure 1).

The approach of uncovering the biological pathways linking exposures to outcomes has gained significant attention in Europe and is the focus of ongoing research through initiatives such as the Horizon 2020 ENDpoiNTs project (Lupu et al., 2020) and the Psychiatric Disorders and Comorbidities Caused by Pollution in the Mediterranean Area (PsyCoMed) project. These projects combine scientific expertise in endocrine disruption and developmental neurotoxicity with advanced *in silico* and *in vitro* tools, innovative experimental approaches and sophisticated biostatistical analyses of human epidemiological and biomonitoring data. Their aim is to uncover both correlative and causal links between neurodevelopmental outcomes and endocrine pathways, addressing not only well-studied targets like estrogen, androgen and thyroid systems but also lesser-known pathways such as the retinoic acid system. Symeonides et al. (2024) employed a multimodal approach, combining human observational studies and preclinical mouse models, to investigate the link between prenatal bisphenol A (BPA) exposure and autism spectrum disorder (ASD) in males. In the human cohort, BPA exposure was associated with increased ASD symptoms and diagnosis in males with low aromatase activity, driven by BPA-induced hypermethylation of the *CYP19A1* brain promoter, with over three-fold increased risk of ASD symptoms at age 2 and six-fold increased risk of diagnosis at age 9 (Symeonides et al., 2024). In mouse models, BPA exposure and aromatase knockout led to ASD-like behaviours, amygdala hypoactivation and brain alterations. *In vitro* experiments confirmed BPA’s suppression of aromatase in neuronal cells. These findings provide insights into how prenatal BPA disrupts aromatase signalling, causing changes characteristic of ASD in males.

To inform policy effectively, broader research covering a wider range of health outcomes beyond neurodevelopment is needed (Symeonides et al., 2024). Further, incorporating mixture analyses into triangulation methods is crucial, as interactions between individual EDCs can amplify their effects. For example, BPA has been shown to enhance oestrogen receptor expression, increasing cellular vulnerability to other EDCs (Hayes et al., 2016; Hamid et al., 2021). This complicates the assessment of combined effects based on individual exposures (Kortenkamp, 2008; J. M. Braun et al., 2016). Misattributing health effects to a single exposure rather than to a correlated harmful exposure carries significant regulatory consequences (Caporale et al., 2022; Duh-Leong et al., 2023). Such poorly defined policies can increase global human exposure to plastic chemicals, resulting in considerable public health issues, disability and economic costs (the societal burden of chemical exposure is estimated at US\$340 billion annually, see for example Duh-Leong et al., 2023). We therefore advocate for the integration of triangulation methods with mixture modelling to more accurately reflect real-world exposure scenarios. This combined

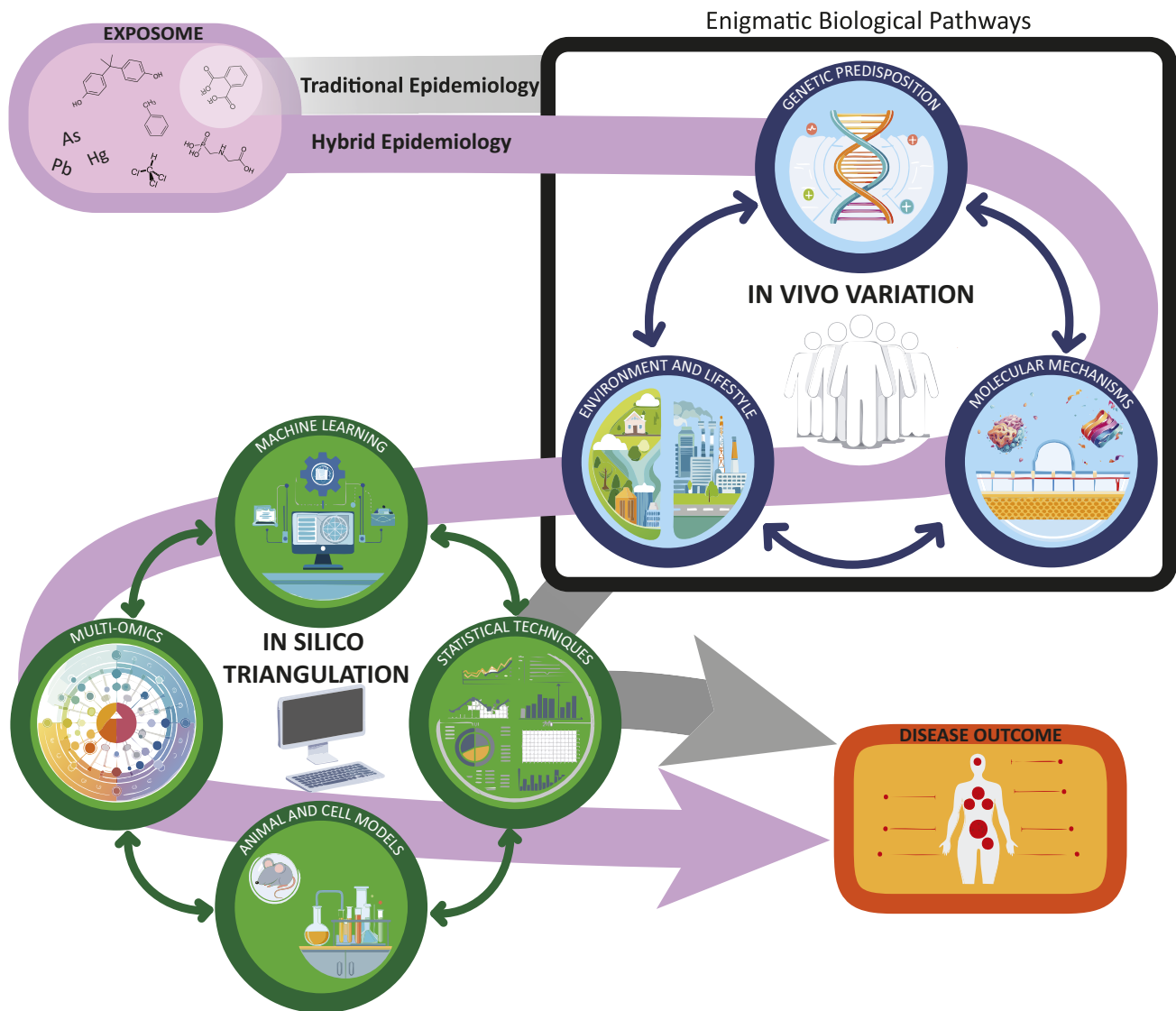


Figure 1. Unlike the traditional epidemiological approach (represented by the grey arrow), which examines the relationship between a single exposure and a disease outcome while many critical aspects of the biological pathway remain uncovered, hybrid epidemiology (depicted by the purple arrow) adopts a more comprehensive view. The hybrid epidemiology approach considers the broader exposome and examines how factors such as genetic predisposition, molecular mechanisms, environmental factors and lifestyle variations within individuals can modulate this effect. Hybrid epidemiology goes a step further by integrating multiple analytical approaches and techniques – including machine learning, multi-omics, advanced statistical methods and animal or cell models – to triangulate and identify the most parsimonious effect of exposure on disease aetiology.

approach enhances the identification of the most harmful chemicals within mixtures and provides insights into whether their effects are additive, antagonistic or synergistic – thus supporting more precise and actionable interventions. The applicability of this hybrid epidemiological framework is particularly relevant in areas where environmental exposures act through diverse, interacting biological pathways — including endocrine disruption, neurodevelopment and immune modulation. In this context, the eight hallmarks of environmental insults proposed by Peters et al. (2021) — namely oxidative stress and inflammation, genomic alterations and mutations, epigenetic alterations, mitochondrial dysfunction, endocrine disruption, altered intercellular communication, altered microbiome communities and impaired nervous system function — offer a valuable conceptual scaffold. Integrating multi-omics and mechanistic evidence with these hallmarks enables researchers to better identify plausible biological pathways and prioritize environmental exposures for regulatory action.

Additionally, the integration of advanced biomolecular techniques, such as genomics, proteomics and transcriptomics, plays a key role in enhancing the biological plausibility of exposure–disease associations and strengthening causal evidence. For instance, utilizing genetic data can effectively reduce confounding in epidemiological research through methods like pedigree-based analyses (e.g., sibling comparisons) and genetically informed approaches (e.g., polygenic scoring, genetic pathway score functions and Mendelian randomization). This capability to identify genotype-specific or individual responses to environmental exposures is crucial to unravelling gene–environment interactions (see, for example, Tanner et al., 2022; Elagali et al., 2025), as the aetiology of many diseases involves intricate interplays between genetic susceptibilities and environmental factors (Manrai et al., 2017; Vermeulen et al., 2020; Chang et al., 2024). Advancing this understanding is important for the development of predictive gene–environment interaction (*GxE*) models, with transformative potential in robust

chemical regulation and large-scale public health monitoring (Motsinger-Reif et al., 2024). Further, recent advances in single-cell omics and human organoid modelling provide powerful opportunities to validate and contextualize associations identified through epidemiological studies (Cuomo et al., 2023; Farbehi et al., 2024; Caporale et al., 2025). For example, single-cell expression quantitative trait loci (eQTL) mapping reveals how genetic variation influences gene expression in specific cell types or states, helping identify where environmental exposures exert their effects. Likewise, organoid systems derived from human stem cells offer controlled platforms to model tissue-specific responses to plastic chemicals and test hypotheses about developmental disruption, endocrine signalling or neurotoxicity (Caporale et al., 2025). These tools strengthen causal inference by enabling experimental testing of epidemiologically derived hypotheses in human-relevant models and are a key complement to the hybrid epidemiology framework. By incorporating these technologies, researchers can bridge the gap between population-level associations and cell-level mechanisms, ultimately producing more translatable, policy-relevant insights (Caporale et al., 2022).

Limitations

While hybrid epidemiological approaches offer promising pathways for strengthening causal inference, several methodological and practical challenges remain. Data availability and accessibility are key limitations, particularly for high-resolution multi-omics and single-cell datasets, which are often derived from small cohorts and may be restricted due to privacy, consent or regulatory concerns. Even when accessible, harmonizing data across studies is difficult due to differences in sample processing, measurement platforms and annotation standards, hindering integrative and meta-analytic analyses.

Modelling the health effects of chemical mixtures also presents significant challenges. Although methods like WQSR, BKMR and elastic net are useful, they often assume additive effects, despite the likelihood of synergistic or antagonistic interactions. Accurately identifying and quantifying such interactions typically requires large sample sizes, which in turn lead to substantial financial costs – a major limiting factor. These financial constraints are compounded by the need for substantial biospecimen volumes (e.g., blood, urine or plasma), which can be especially difficult to obtain from vulnerable populations such as neonates or young children. As the field progresses, overcoming data and modelling limitations will be essential to bridge existing gaps in our understanding of how environmental exposures affect human health.

Conclusion

Although observational epidemiological studies have provided important information on the health impacts of exposure to plastic chemicals, it is time to move beyond simple ‘black-box’ associations. Hybrid epidemiology, powered by examining mixture analyses, including underpinning -omics and molecular mechanisms, as well as machine learning, offers the opportunity to examine causal evidence, the underlying biological pathways and understand the complex interactions between multiple chemical exposures. This shift is crucial for making informed decisions that protect human health in the face of increasing environmental challenges. By accounting for the totality of exposures and integrating mechanistic evidence, we can more effectively understand and prevent

disease, ultimately improving population health, quality of life and economic outcomes such as Gross Domestic Product (GDP, Landrigan, 2017).

Open peer review. To view the open peer review materials for this article, please visit <http://doi.org/10.1017/plc.2025.10011>.

Data availability statement. No data were used for the research described in this article.

Acknowledgements. All authors made a substantial contribution to the conception, preparation, writing and editing of the manuscript. All authors reviewed the manuscript before submission.

Competing interests. The authors declare no competing interests.

Financial statement. This research is funded by Minderoo Foundation (Australia), an independent not-for-profit philanthropic organization. Neither the foundation, nor its benefactors, had any influence in the planning and conduction of this work.

References

- Amolegbe SM, Carlin DJ, Henry HF, Heacock ML, Trottier BA and Suk WA (2022) Understanding exposures and latent disease risk within the national institute of environmental health sciences superfund research program. *Experimental Biology and Medicine* **247**(7), 529–537.
- Argyri KD, Gallos IK, Amditis A and Dionysiou DD (2024) Exposomics and cardiovascular diseases: a scoping review of machine learning approaches. *medRxiv*, 2024–07.
- Bobb JF, Henn BC, Valeri L and Coull BA (2018) Statistical software for analyzing the health effects of multiple concurrent exposures via bayesian kernel machine regression. *Environmental Health* **17**, 1–10.
- Bobb JF, Dominici F and Peng RD (2011) A bayesian model averaging approach for estimating the relative risk of mortality associated with heat waves in 105 us cities. *Biometrics* **67**(4), 1605–1616.
- Bobb JF, Valeri L, Henn BC, Christiani DC, Wright RO, Mazumdar M, Godleski JJ and Coull BA (2015) Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics* **16**(3), 493–508.
- Braun G, Herberth G, Krauss M, König M, Wojtysiak N, Zenclussen AC and Escher BI (2024) Neurotoxic mixture effects of chemicals extracted from blood of pregnant women. *Science* **386**(6719), 301–309.
- Braun JM, Gennings C, Hauser R and Webster TF (2016) What can epidemiological studies tell us about the impact of chemical mixtures on human health? *Environmental Health Perspectives* **124**(1), A6–A9.
- Caporale N, Castaldi D, Rigoli MT, Cheroni C, Valenti A, Stucchi S, Lessi M, Bulgheresi D, Trattaro S, Pezzali M, et al. (2025) Multiplexing cortical brain organoids for the longitudinal dissection of developmental traits at single-cell resolution. *Nature Methods* **22**(2), 358–370.
- Caporale N, Leemans M, Birgersson L, Germain P-L, Cheroni C, Borbély G, Engdahl E, Lindh C, Bressan RB, Cavallo F, et al. (2022) From cohorts to molecules: Adverse impacts of endocrine disrupting mixtures. *Science* **375**(6582), eabe8244.
- Carrico C, Gennings C, Wheeler DC and Factor-Litvak P (2015) Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. *Journal of Agricultural, Biological, and Environmental Statistics* **20**, 100–120.
- Chang L, Ewald J, Hui F, Bayen S and Xia J (2024) A data-centric perspective on exposomics data analysis. *Exposome* **4**(1), osae005.
- Chipman HA, George EI and McCulloch RE (2010) Bart: bayesian additive regression trees. *Annals of Applied Statistics* **4** (1), 266–298. <https://projecteuclid.org/journals/annals-of-applied-statistics/volume-4/issue-1/BART-Bayesian-additive-regression-trees/10.1214/09-AOAS285>.
- Cuomo ASE, Nathan A, Raychaudhuri S, MacArthur DG and Powell JE (2023) Single-cell genomics meets human genetics. *Nature Reviews Genetics* **24**(8), 535–549.

- Darbre PD** (2022) How could endocrine disrupters affect human health? In *Endocrine Disruption and Human Health*. Cambridge, Massachusetts: Academic Press, pp. 31–56. <https://doi.org/10.1016/B978-0-12-821985-0.00002-5>
- Duh-Leong C, Maffini MV, Kassotis CD, Vandenberg LN and Trasande L** (2023) The regulation of endocrine-disrupting chemicals to minimize their impact on health. *Nature Reviews Endocrinology* **19**(10), 600–614.
- Elagali A, Eisner A, Tanner S, Drummond K, Symeonides C, Love C, Tang MLK, Mansell T, Burgner D, Collier F, et al.** (2025) A pathway-based genetic score for inflammation: An indicator of vulnerability to phthalate-induced adverse neurodevelopment outcomes. *International Journal of Hygiene and Environmental Health* **264**, 114514.
- Farbehi N, Neavin DR, Cuomo ASE, Studer L, MacArthur DG and Powell JE** (2024) Integrating population genetics, stem cell biology and cellular genomics to study complex human diseases. *Nature Genetics* **56**(5), 758–766.
- Fedak KM, Bernal A, Capshaw ZA and Gross S** (2015) Applying the Bradford hill criteria in the 21st century: How data integration has changed causal inference in molecular epidemiology. *Emerging Themes in Epidemiology* **12**, 1–9.
- Forns J, Mandal S, Iszatt N, Polder A, Thomsen C, Lyche JL, Stigum H, Vermeulen R and Eggesbø M** (2016) Novel application of statistical methods for analysis of multiple toxicants identifies ddt as a risk factor for early child behavioral problems. *Environmental Research* **151**, 91–100.
- Gelman A, Carlin JB, Stern HS and Rubin DB** (1995) *Bayesian Data Analysis*. Boca Raton, Florida: Chapman/Hall/CRC.
- Grandjean P and Landrigan PJ** (2006) Developmental neurotoxicity of industrial chemicals. *The Lancet* **368**(9553), 2167–2178.
- Al-Gubory KH** (2014) Environmental pollutants and lifestyle factors induce oxidative stress and poor prenatal development. *Reproductive Biomedicine Online* **29**(1), 17–31.
- Guimbaud J-B, Siskos AP, Sakhi AK, Heude B, Sabidó E, Borràs E, Keun H, Wright J, Julvez J, Urquiza J, et al.** (2024) Machine learning-based health environmental-clinical risk scores in european children. *Communications Medicine* **4**(1), 98.
- Hamid N, Junaid M and Pei D-S** (2021) Combined toxicity of endocrine-disrupting chemicals: A review. *Ecotoxicology and Environmental Safety* **215**, 112136.
- Hasin Y, Seldin M and Lusic A** (2017) Multi-omics approaches to disease. *Genome Biology* **18**, 1–15.
- Hayes L, Weening A and Morey LM** (2016) Differential effects of estradiol and bisphenol a on set8 and sirt1 expression in ovarian cancer cells. *Dose-Response* **14**(2), 1559325816640682.
- Karczewski KJ and Snyder MP** (2018) Integrative omics for health and disease. *Nature Reviews Genetics* **19**(5), 299–310.
- Kortenkamp A** (2008) Low dose mixture effects of endocrine disrupters: Implications for risk assessment and epidemiology. *International Journal of Andrology* **31**(2), 233–240.
- Landrigan PJ** (2017) Air pollution and health. *The Lancet Public Health* **2**(1), e4–e5.
- Landrigan PJ, Raps H, Cropper M, Bald C, Brunner M, Canonizado EM, Charles D, Chiles TC, Donohue MJ, Enck J, et al.** (2023) The minderoo-Monaco commission on plastics and human health. *Annals of Global Health* **89**(1), 1–215.
- Lanphear BP** (2015) The impact of toxins on the developing brain. *Annual Review of Public Health* **36**(1), 211–230.
- Lazarevic N, Barnett AG, Sly PD and Knibbs LD** (2019) Statistical methodology in studies of prenatal exposure to mixtures of endocrine-disrupting chemicals: A review of existing approaches and new alternatives. *Environmental Health Perspectives* **127**(2), 026001.
- Lenters V, Portengen L, Rignell-Hydbom A, Jönsson BAG, Lindh CH, Piersma AH, Toft G, Bonde JP, Heederik D, Rylander L, et al.** (2016) Prenatal phthalate, perfluoroalkyl acid, and organochlorine exposures and term birth weight in three birth cohorts: Multi-pollutant models based on elastic net regression. *Environmental Health Perspectives* **124**(3), 365–372.
- Lundberg S** (2017) A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* **30**. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
- Lupu D, Andersson P, Bornehag C-G, Demeneix B, Fritsche E, Gennings C, Lichtensteiger W, Leist M, Leonards PEG, Ponsoy A-L, et al.** (2020) The endpoints project: Novel testing strategies for endocrine disruptors linked to developmental neurotoxicity. *International Journal of Molecular Sciences* **21**(11), 3978.
- Maitre L, Guimbaud J-B, Warembourg C, Güil-Oumrait N, Petrone PM, Chadeau-Hyam M, Vrijheid M, Basagaña X, Gonzalez JR, Exposome Data Challenge Participant Consortium, Consortium EDCP, et al.** (2022) State-of-the-art methods for exposure-health studies: Results from the exposome data challenge event. *Environmental International* **168**, 107422.
- Manrai AK, Cui Y, Bushel PR, Hall M, Karakitsios S, Mattingly CJ, Ritchie M, Schmitt C, Sarigiannis DA, Thomas DC, et al.** (2017) Informatics and data analytics to support exposome-based discovery for public health. *Annual Review of Public Health* **38**(1), 279–294.
- Miller GW and Consortium BE** (2025) Integrating exposomics into biomedicine. *Science* **388**(6745), 356–358.
- Motsinger-Reif AA, Reif DM, Akhtari FS, House JS, Campbell CR, Messier KP, Fargo DC, Bowen TA, Nadadur SS, Schmitt CP, et al.** (2024) Gene-environment interactions within a precision environmental health framework. *Cell Genomics* **4**(7), 100591. <https://doi.org/10.1016/j.xgen.2024.100591>
- Peters A, Nawrot TS and Baccarelli AA** (2021) Hallmarks of environmental insults. *Cell* **184**(6), 1455–1468.
- Stafoggia M, Breitner S, Hampel R and Basagaña X** (2017) Statistical approaches to address multi-pollutant mixtures and multiple exposures: The state of the science. *Current Environmental Health Reports* **4**, 481–490.
- Stanfield Z, Setzer RW, Hull V, Sayre RR, Isaacs KK and Wambaugh JF** (2024) Characterizing chemical exposure trends from nhanes urinary biomonitoring data. *Environmental Health Perspectives* **132**(1), 017009.
- Sun Z, Tao Y, Li S, Ferguson KK, Meeker JD, Park SK, Batterman SA and Mukherjee B** (2013) Statistical strategies for constructing health risk models with multiple pollutants and their interactions: Possible choices and comparisons. *Environmental Health* **12**, 1–19.
- Symeonides C, Aromataris E, Mulders Y, Dizon J, Stern C, Barker TH, Whitehorn A, Pollock D, Marin T and Dunlop S** (2024) An umbrella review of meta-analyses evaluating associations between human health and exposure to major classes of plastic-associated chemicals. *Annals of Global Health* **90**(1), 1–54.
- Symeonides C, Vacy K, Thomson S, Tanner S, Chua HK, Dixit S, Mansell T, O’Hely M, Novakovic B, Herbstman JB, et al.** (2024) Male autism spectrum disorder is linked to brain aromatase disruption by prenatal bpa in multimodal investigations and 10hda ameliorates the related mouse phenotype. *Nature Communications* **15**(1), 6367.
- Tanner S, Thomson S, Drummond K, O’hely M, Symeonides C, Mansell T, Saffery R, Sly PD, Collier F, Burgner D, et al.** (2022) A pathway-based genetic score for oxidative stress: An indicator of host vulnerability to phthalate-associated adverse neurodevelopment. *Antioxidants* **11**(4), 659.
- Vermeulen R, Schymanski EL, Barabási A-L and Miller GW** (2020) The exposome and health: Where chemistry meets biology. *Science* **367**(6476), 392–396.
- Wild CP** (2005) Complementing the genome with an “exposome”: The outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiology Biomarkers & Prevention* **14**(8), 1847–1850.
- Woodruff TJ, Zota AR and Schwartz JM** (2011) Environmental chemicals in pregnant women in the United States: Nhanes 2003–2004. *Environmental Health Perspectives* **119**(6), 878–885.
- Woodside III FC and Davis AG** (2012) The Bradford hill criteria: The forgotten predicate. *Thomas Jefferson Law Review* **35**, 103.
- Yang Y, Li R, Li W, Wang M, Cao Y, Wu Z and Xu Q** (2013) The association between ambient air pollution and daily mortality in Beijing after the 2008 olympics: A time series study. *PLoS One* **8**(10), e76759.
- Zou H and Zhang HH** (2009) On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics* **37**(4), 1733.