**Introduction**

Rapid industrial development has created conditions of cumulative chronic toxicity which pose an acute risk to the wellbeing of humans and our living environment. It has been estimated that human activity releases chemicals at a rate of 220 billion tons per annum (Cribb, 2016). Scholars have recently formally declared that, at this global rate of chemical release, humanity has surpassed the safe operating space of the planetary boundary for novel entities (Persson et al., 2022). As a result, exposure to low levels of pollutants has become an inevitable peril of daily life (Naidu et al., 2021; Vineis, 2018). Hence, it is especially timely that we prompt regulatory control of industrial pollution through studies which investigate the health effects of chemical exposures.

For this, we turn to epidemiological studies. The broad field of preventive epidemiology involves the identification of potentially modifiable risk factors that contribute to the burden of disease within human populations. Environmental epidemiology, in particular, considers the effect of environmental exposures — chemical or otherwise. However, studies concerning chemical pollutants in environmental epidemiology have historically focused on elucidating the effect and mechanisms of exposures to a single pollutant. In reality, humans are invariably exposed to numerous complex chemical mixtures which together contribute to the progression of adverse health outcomes. Therefore, risk assessments of single pollutants likely fail to capture the true consequences of these complex exposures (Heys et al., 2016). Assessing mixtures of chemicals can also have more direct implications for public health interventions. The United States Environmental Protection Agency (U.S. EPA) currently passes regulations for individual pollutants. In practice, though, regulation occurs by controlling the source of pollution, which is responsible for the production of a whole mixture of chemicals with specific joint effects on human health. As a result, the National Academies of Science has advocated for a multipollutant regulatory approach, which is likely to be more protective of human health (Committee on Incorporating 21st Century Science into Risk-Based Evaluations et al., 2017).

There are clear practical motivations for studies that examine the health effects of exposure to co-occurring chemical mixtures, hereafter referred to as exposure mixtures. However, expanding the focus of analysis from one exposure to multiple exposures introduces unique statistical challenges. In addition to a common issue of small effect sizes and small sample sizes present in most exposure analyses, multiple exposure analyses must also contend with high-dimensionality, collinearity, non-linear effects, and non-additive interactions (Yu et al., 2022). In particular, data with numerous pollutants, or predictors, require exponentially greater levels of complexity and time cost in analysis. Collinearity between exposures is common when analyzing pollutants from a single source and can lead to unstable estimates in a generalized linear model if left unaccounted for. Finally, exposures can have both non-linear single effects and non-additive interaction effects, which are difficult to capture unless explicitly specified in the model.

The classic multiple linear regression framework often fails to capture the true effects in this setting. In the past few years, a wide variety of statistical methods have been developed to overcome these challenges (Gibson et al., 2019; Yu et al., 2022), which have been accompanied by a host of comparative simulation studies for general mixture scenarios (e.g., Hoskovec et al., 2021; Lazarevic et al., 2020; Pesenti et al., 2023). However, there is not yet conclusive guidance about the ability of these methods to conduct inference on non-additive interactions between exposures.

The goal of this thesis is to fill this gap in the literature by exploring the theory and performance of Bayesian regression techniques for quantifying complex interactions between environmental exposures. Specifically, I will compare two recently developed models for estimating the health effects of exposure mixtures: Bayesian Kernel Machine Regression (BKMR, Bobb et al., 2015) and Bayesian Semiparametric Regression (BSR, Antonelli et al., 2020).

In an age where anthropogenic actions have radically reshaped the earth, humanistic inquiry can offer critical insights into how we navigate the hazards of our rapidly changing environment. I begin in Chapter 2 by contextualizing this thesis with a brief overview of cultural and social understandings of toxicity. Chapter 3 explains the motivation for studying interactions and provides background on the theory of Bayesian methods for analyzing exposure mixtures. Chapter 4 assesses the performance of these methods using a simulation study, based on a dataset with information on the relationship between prenatal exposure to heavy metals and gestational weight. Chapter 5 explores an application on X data [TBD]. I conclude with a discussion of the implications of this work for the future study of complex interactions in exposure mixture studies.

Bib so far

Antonelli, J., Mazumdar, M., Bellinger, D., Christiani, D., Wright, R., & Coull, B. (2020). Estimating the health effects of environmental mixtures using Bayesian semiparametric regression and sparsity inducing priors. *The Annals of Applied Statistics*, *14*(1), 257–275. https://doi.org/10.1214/19-AOAS1307

Bobb, J. F., Valeri, L., Claus Henn, B., Christiani, D. C., Wright, R. O., Mazumdar, M., Godleski, J. J., & Coull, B. A. (2015). Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics*, *16*(3), 493–508. https://doi.org/10.1093/biostatistics/kxu058

Committee on Incorporating 21st Century Science into Risk-Based Evaluations, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies, & National Academies of Sciences, Engineering, and Medicine. (2017). *Using 21st Century Science to Improve Risk-Related Evaluations* (p. 24635). National Academies Press. https://doi.org/10.17226/24635

Cribb, J. (2016). *Surviving the 21st century: Humanity’s ten great challenges and how we can overcome them*. Springer.

Gibson, E. A., Nunez, Y., Abuawad, A., Zota, A. R., Renzetti, S., Devick, K. L., Gennings, C., Goldsmith, J., Coull, B. A., & Kioumourtzoglou, M.-A. (2019). An overview of methods to address distinct research questions on environmental mixtures: An application to persistent organic pollutants and leukocyte telomere length. *Environmental Health*, *18*(1), 76. https://doi.org/10.1186/s12940-019-0515-1

Heys, K., Shore, R., Pereira, M., Jones, K., & Martin, F. (2016). Risk assessment of environmental mixture effects. *RSC Advances*, *6*(53), 47844–47857. https://doi.org/10.1039/C6RA05406D

Hoskovec, L., Benka-Coker, W., Severson, R., Magzamen, S., & Wilson, A. (2021). Model choice for estimating the association between exposure to chemical mixtures and health outcomes: A simulation study. *PLOS ONE*, *16*(3), e0249236. https://doi.org/10.1371/journal.pone.0249236

Lazarevic, N., Knibbs, L. D., Sly, P. D., & Barnett, A. G. (2020). Performance of variable and function selection methods for estimating the nonlinear health effects of correlated chemical mixtures: A simulation study. *Statistics in Medicine*, *39*(27), 3947–3967. https://doi.org/10.1002/sim.8701

Naidu, R., Biswas, B., Willett, I. R., Cribb, J., Kumar Singh, B., Paul Nathanail, C., Coulon, F., Semple, K. T., Jones, K. C., Barclay, A., & Aitken, R. J. (2021). Chemical pollution: A growing peril and potential catastrophic risk to humanity. *Environment International*, *156*, 106616. https://doi.org/10.1016/j.envint.2021.106616

Persson, L., Carney Almroth, B. M., Collins, C. D., Cornell, S., De Wit, C. A., Diamond, M. L., Fantke, P., Hassellöv, M., MacLeod, M., Ryberg, M. W., Søgaard Jørgensen, P., Villarrubia-Gómez, P., Wang, Z., & Hauschild, M. Z. (2022). Outside the Safe Operating Space of the Planetary Boundary for Novel Entities. *Environmental Science & Technology*, *56*(3), 1510–1521. https://doi.org/10.1021/acs.est.1c04158

Pesenti, N., Quatto, P., Colicino, E., Cancello, R., Scacchi, M., & Zambon, A. (2023). Comparative efficacy of three Bayesian variable selection methods in the context of weight loss in obese women. *Frontiers in Nutrition*, *10*, 1203925. https://doi.org/10.3389/fnut.2023.1203925

Vineis, P. (2018). From John Snow to omics: The long journey of environmental epidemiology. *European Journal of Epidemiology*, *33*(4), 355–363. https://doi.org/10.1007/s10654-018-0398-4

Yu, L., Liu, W., Wang, X., Ye, Z., Tan, Q., Qiu, W., Nie, X., Li, M., Wang, B., & Chen, W. (2022). A review of practical statistical methods used in epidemiological studies to estimate the health effects of multi-pollutant mixture. *Environmental Pollution*, *306*, 119356. https://doi.org/10.1016/j.envpol.2022.119356