# Discussion

We estimated 20-year risk of NHL over the years 1985 through 2004 in the UAW-GM Autoworkers Cohort under different hypothetical interventions targeting exposure to soluble MWFs. The hazard-extended ICE g-formula estimates of NHL cumulative incidence under hypothetical limits on average annual exposure to soluble MWF showed a monotonic dose-dependent relationship with lower cumulative incidence estimates arising from stronger limits. Contrasts in estimated NHL cumulative incidence were statistically significant under the strongest static intervention capping soluble MWF exposure at 0.05 mg/m3 and the two dynamic interventions reducing soluble MWF by an amount dependent on exposure to straight and synthetic MWF. Under the assumptions of conditional exchangeability, positivity, consistency, and correct model specification, our estimates are unbiased for the true counterfactual cumulative incidence under the specified hypothetical limit and no censoring by death, even in the presence of the HWSE.

Conditional exchangeability means that for all time points, there is no confounding of the relationship between exposure/censoring and NHL status at all future time points given the observed past, including past exposure and covariates. A major threat to conditional exchangeability in longitudinal occupational studies is the HWSE. We limit potential bias due to the HWSE by conditioning on all past cumulative exposure, employment status, and cumulative time off at each time point. Cumulative time off and employment status are reasonable mediators of the causal paths linking past health to future exposure and health, but may not be sufficient to fully adjust for the HWSE. Declines in a worker’s health may lead to reductions in work-related exposure without changes in employment status or time off work.1 We expect this absence of good time-varying measures of worker health over the life course to bias results toward the null.

Positivity refers to the need for adequate variation in future exposure among strata formed by past observed covariate and intervention-compliant exposure histories. Even under conditional exchangeability, where exposures within these strata may be considered the result of experimental assignment, expected counterfactual outcomes under different exposures may not be estimable if there is sparsity in the observed distribution of exposures.2 We investigated static and dynamic stochastic interventions on soluble MWF exposure. Since we only intervened when average annual exposure exceeded the hypothetical limit under consideration, our parameters of interest achieve positivity more easily than those for deterministic interventions e.g. setting all to exposed versus unexposed. Nonetheless, violations in positivity were still of concern due to the high dimensionality of covariates, as in common in longitudinal settings. We addressed sparsity by summarizing the 20 years of follow-up over a coarser timescale and by representing covariates using fewer categories. In practice, causal inference using observational data must always balance positivity, covariate adjustment, and outcome modeling.3 Since the g-formula paradigm does not require exposure modeling, concerns over potential bias due to practical violations in positivity may be relaxed if correct model specification is attained.

The consistency assumption, also known as the no-multiple-versions-of-treatment or stable unit treatment value assumption, is that counterfactual outcomes under each possible exposure value take on a unique value.4,5

Our exposure of interest is a complex mixture of diverse components with substantial variation over time due to changes in formulation and due to the natural physical, chemical, and biological changes in a fluid over the course of their use.6 In the last several decades, there have been substantial changes in the formulation of MWF, including the exclusion of nitrogenous compounds and the use of highly-refined mineral oils. The recent establishment of the Ecolabel for lubricants in the European Union reflects the shift toward prioritizing sustainability and health in the design and production of MWF and other lubricants.7,8 However, water-based MWF that satisfy Ecolabel criteria are nonetheless susceptible to the accumulation of dirt, metals, microbes, products of microbial metabolism, and unemulsified oil. These contaminants and their control strategies may also induce risks to human health. Furthermore, adoption of this new class of MWF remains limited, as is evidence of their safety. Considering the evolving complexity in MWF composition, epidemiologic investigations remain an invaluable approach for developing evidence-based occupational health policy and estimating their causal effect.

Much of the existing epidemiologic literature linking occupational and environmental exposures to NHL risk report findings from case-control studies where exposures are measured crudely as binary indicators of exposure or membership in a particular occupational group.9–12 Associations between occupations and NHL risk vary considerably, but one study of working men in Kansas and Nebraska found strong associations between NHL risk and occupations involving metalworking and motor vehicles.13 Both of these occupations may entail exposure to soluble MWFs, which contain a number of additives of concern for human health and NHL risk in particular. Organophosphorous, chlorinated, organosulfur, and amine boron compounds are common additives for the control of microbial growth, improved performance under high heat, and inhibition of corrosion.14

Organophosphate pesticides, have been linked to cancer risk in epidemiologic and animal studies. Some have been classified as possibly carcinogenic by the IARC.15 Studies of occupational exposure to chlorinated solvents and pesticides have also been linked to NHL risk.16–20 In 2014, the International Agency for Research on Cancer classified trichloroethylene, tetrachloroethylene, and other chlorinated agents as Group 1 carcinogens.21 Chlorinated solvents are commonly used as degreasers in industrial settings, but their use in the plants under study here was uncommon and limited to particular operations.22 Although these specific compounds may be absent from MWF, their structural similarity to MWF additives suggest similarities in molecular function and toxicity.

The first time an association between NHL and MWF exposure in this cohort was reported was in Colbeth et alia (in press), which presented standard survival analyses investigating the associations between MWF and several cancer incidence outcomes. We verified and further examined this association in an approach that adjusted for time-varying confounding possibly affected by past exposure. The present approach also yielded marginal rather than conditional estimates of effect while evaluating realistic interventions, and is thus more compatible with a population health framework. We contrasted the counterfactual cumulative incidence under interventions on soluble MWF to that under the natural course by computing risk ratios, but our approach allows for the estimation of any risk contrast, including risk difference and attributable risk.

Previous studies in occupational epidemiology have successfully estimated causal effects of interventions using propensity score weighting and non-iterative expressions of the parametric g-formula.23,24 The use of propensity score-based weighting is not always appropriate or possible because of violations or practical violations of the positivity assumption.3 Causal inference methods relying on the specification of outcome models e.g. parametric g-formula and g-estimation avoid modeling the exposure mechanism altogether and thus, are more robust to near violations of the positivity assumption. The non-iterative expression of the parametric g-formula is particularly attractive because its intuitive approach for handling competing events.25 However, a major limitation of that approach is the g-null paradox: the guaranteed misspecification of parametric models resulting in the false rejection of the null hypothesis when the null is true and there is time-varying confounding affected by past exposure.26

The g-null paradox is not relevant to the estimator we applied or any other iterative conditional expectation expression of the g-formula. However, by forgoing the use of the non-iterative parametric g-formula, we also sacrificed the prospect of estimating counterfactual risks over larger windows and adjusting for confounding over a finer grid of time points. Limiting our analysis to eight time points within the 20-year window between 1985 and 2004 was necessary to prevent excessive sparsity in the high-dimensional estimation problem. There were two or more years of follow-up time between time points, so our analysis may not have adequately controlled for time-varying confounding by measured covariates possibly affected by prior exposure. However, since healthier workers tend to accrue more exposure over the course of their working lifetime, we expect coarser control of time-varying confounding affected by prior exposure to result in bias toward the null.27

This study investigated the effect of hypothetical limits on average annual exposure to MWF by comparing the standardized distribution of NHL under various distributions of cumulative exposure induced by applying upper bounds to annual exposure. When average annual exposure did not exceed the hypothetical limit, it was left at the observed level of exposure. We selected these hypothetical limits based on the NIOSH REL of 0.5 mg/m3.28 If the REL were enforced in the workplace, we would not expect changes in exposure associated with low-exposure operations. Likewise, when the REL was not enforced, we still observed average annual exposure below the REL for many workers. Hence, considering the counterfactual scenario where all workers experienced annual exposure at the REL to that where all workers experienced annual exposure at some level higher than the REL would result in an over-estimate of the expected benefit of REL enforcement in the real world. Unrealistic contrasts are difficult to avoid when considering static deterministic interventions. Instead, the interventions applied here are stochastic interventions resulting in a post-intervention distribution of exposure that better reflects what we would expect if average annual exposure were capped at some hypothetical limit. The interventions which ignore covariate values such as exposure to straight and synthetic MWF are static. The interventions which reduce exposure to soluble MWF by a quantity determined by exposure to other MWF are dynamic. The dynamic stochastic interventions respect the fact that the REL is for total MWF, not just for soluble MWF. If an individual experienced a large amount of exposure to straight and synthetic MWF in one year, then their exposure to soluble MWF should be bounded at a lower level so that the sum of their exposure to soluble, straight, and synthetic MWF is no greater than the hypothetical limit. By specifying stochastic interventions, our analysis yielded results that better evaluate policy than those investigating deterministic interventions.29,30

# Conclusion

1. Garcia E, Picciotto S, Costello S, Bradshaw PT, Eisen EA. Assessment of the healthy worker survivor effect in cancer studies of the united autoworkers-general motors cohort. *Occupational and environmental medicine*. 2017;74(4):294-300.

2. Maldonado G, Greenland S. Estimating causal effects. *International journal of epidemiology*. 2002;31(2):422-429.

3. Petersen ML, Porter P, Gruber S, Wang Y, van der Laan MJ. Diagnosing and responding to violations in the positivity assumption. *Statistical Methods in Medical Research*. 2012;21(1):31-54. doi:[10.1177/0962280210386207](https://doi.org/10.1177/0962280210386207).

4. Cole SR, Frangakis CE. The consistency statement in causal inference: A definition or an assumption? *Epidemiology*. 2009;20(1):3-5.

5. VanderWeele TJ, Shpitser I. On the definition of a confounder. *Annals of statistics*. 2013;41(1):196.

6. Howell JK, Lucke WE, White EM. Metalworking fluids. In: Byers JP, ed. CRC Press; 2006.

7. Commission TE. Establishing the EU ecolabel criteria for lubricants. *Official Journal of the European Union*. 2018.

8. Skerlos SJ, Hayes KF, Clarens AF, Zhao F. Current advances in sustainable metalworking fluids research. *International Journal of Sustainable Manufacturing*. 2008;1(1-2):180-202. doi:[10.1504/IJSM.2008.019233](https://doi.org/10.1504/IJSM.2008.019233).

9. Cano MI, Pollán M. Non-hodgkin’s lymphomas and occupation in sweden. *International archives of occupational and environmental health*. 2001;74(6):443-449.

10. Costantini AS, Miligi L, Kriebel D, et al. A multicenter case-control study in italy on hematolymphopoietic neoplasms and occupation. *Epidemiology*. 2001:78-87.

11. Karunanayake CP, McDuffie HH, Dosman JA, Spinelli JJ, Pahwa P. Occupational exposures and non-hodgkin’s lymphoma: Canadian case-control study. *Environmental Health*. 2008;7(1):1-9.

12. ‘t Mannetje A, De Roos AJ, Boffetta P, et al. Occupation and risk of non-hodgkin lymphoma and its subtypes: A pooled analysis from the InterLymph consortium. *Environmental health perspectives*. 2016;124(4):396-405.

13. Zheng T, Blair A, Zhang Y, Weisenburger DD, Zahm SH. Occupation and risk of non-hodgkin’s lymphoma and chronic lymphocytic leukemia. *Journal of occupational and environmental medicine*. 2002;44(5):469-474.

14. Evans R, Hooijman J, van der Veer J. High-speed machining. In: Gupta K, Davim P, eds. Academic Press; 2020.

15. IARC. *IARC Monographs on the Evaluation of Carcinogenic Risk of the Chemical to Humans: Some Organophosphate Insecticides and Herbicides*. Vol 112. World Health Organization International Agency for Research on Cancer; 2017.

16. Cocco P, Brennan P, Ibba A, et al. Plasma polychlorobiphenyl and organochlorine pesticide level and risk of major lymphoma subtypes. *Occupational and Environmental Medicine*. 2008;65(2):132-140.

17. Purdue MP, Bakke B, Stewart P, et al. A case-control study of occupational exposure to trichloroethylene and non-hodgkin lymphoma. *Environmental health perspectives*. 2011;119(2):232-238.

18. Cocco P, Vermeulen R, Flore V, et al. Occupational exposure to trichloroethylene and risk of non-hodgkin lymphoma and its major subtypes: A pooled IinterLlymph analysis. *Occupational and environmental medicine*. 2013;70(11):795-802.

19. Vlaanderen J, Straif K, Pukkala E, et al. Occupational exposure to trichloroethylene and perchloroethylene and the risk of lymphoma, liver, and kidney cancer in four nordic countries. *Occupational and environmental medicine*. 2013;70(6):393-401.

20. Callahan CL, Stewart PA, Friesen MC, et al. Case-control investigation of occupational exposure to chlorinated solvents and non-hodgkin’s lymphoma. *Occupational and environmental medicine*. 2018;75(6):415-420.

21. IARC. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Trichloroethylene, Tetrachloroethylene, and Some Other Chlorinated Agents*. Vol 106. World Health Organization International Agency for Research on Cancer; 2014:106-116.

22. Shrestha D, Liu S, Hammond SK, et al. Risk of renal cell carcinoma following exposure to metalworking fluids among autoworkers. *Occupational and environmental medicine*. 2016;73(10):656-662.

23. Neophytou AM, Costello S, Brown DM, et al. Marginal structural models in occupational epidemiology: Application in a study of ischemic heart disease incidence and PM2. 5 in the US aluminum industry. *American journal of epidemiology*. 2014;180(6):608-615.

24. Neophytou AM, Picciotto S, Costello S, Eisen EA. Occupational diesel exposure, duration of employment, and lung cancer: An application of the parametric g-formula. *Epidemiology*. 2016;27(1):21-28. doi:[10.1097/EDE.0000000000000389](https://doi.org/10.1097/EDE.0000000000000389).

25. Naimi AI, Tchetgen Tchetgen EJ. Invited commentary: Estimating population impact in the presence of competing events. *American journal of epidemiology*. 2015;181(8):571-574.

26. McGrath S, Young JG, Hernán MA. Revisiting the g-null paradox. *Epidemiology*. 2022;33(1):114-120.

27. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology*. 1994;5(2):189-196. <http://www.jstor.org/stable/3702361>.

28. Rosenstock L, ed. *What You Need to Know about Occupational Exposure to Metalworking Fluids*. Department of Health; Human Services (NIOSH); 1998.

29. Bembom O, van der Laan MJ. A practical illustration of the importance of realistic individualized treatment rules in causal inference. *Electronic journal of statistics*. 2007;1:574.

30. Muñoz ID, van der Laan M. Population intervention causal effects based on stochastic interventions. *Biometrics*. 2012;68(2):541-549.