Evaluating hypothetical limits on metalworking fluid exposure for reducing non-Hodgkin lymphoma incidence: An application of the hazard-extended iterative conditional expectation parametric g-formula

Kevin T. Chen,a,b, Sally Picciottob, Ellen A. Eisenb

a University of California, Berkeley School of Public Health, Division of Epidemiology & Biostatistics

b University of California, Berkeley School of Public Health, Division of Environmental Health Sciences

Corresponding author:  
Kevin T. Chen  
2121 Berkeley Way  
Room 5302, Desk 5305-1  
Berkeley, CA 94720-7360  
[kevchen@berkeley.edu](mailto:kevchen@berkeley.edu)

# Abstact

Metalworking fluids (MWF) are complex chemical mixtures used to lubricate and cool machinery during manufacturing processes. Straight MWF are the oldest class of MWF, comprising of simple oils or combinations of oils. By the mid 20th century, straight MWF waned in popularity and ceded market share to soluble MWF. Soluble MWF are water-oil mixtures with diverse additives for improving performance and durability. By the 1970s, semisynthetic and synthetic MWF, which contain little to no oil, overtook straight oils in market share. Nonetheless, at the turn of the millennium, soluble MWFs still retained the largest share of the MWF market.1 Carcinogenicity of MWFs has been of concern since as early as the 1970s, when the International Agency for Research on Cancer (IARC) classified mineral oils as carcinogenic, citing studies of occupational exposures among workers in oil, textile, and metal industries.2 Concerns over the health effects of straight MWF hastened the adoption of soluble MWF, which contain less oil per volume, but a richer cocktail of additives including chromates, cycloalkanes, phenols, organochlorines, nitrosamines, and triazines: classes of chemicals with known or plausible carcinogenicity.3–5 Health and environmental considerations led to substantial variation in MWF composition over time, but the elimination of particular etiologic agents often leads to their replacement with other potentially hazardous chemical compounds. Exposure limits for MWF as a mixture encourage limiting occupational exposure to multiple agents simultaneously.

Here, we assess the effect of hypothetical exposure limits on non-Hodgkin lymphoma (NHL) cumulative incidence (risk) from 1985 to 2005 in the United Auto Workers-General Motors (UAW-GM) Occupational Cohort Study. The large size of the study population and rich time-varying, quantitative MWF exposure data provide an opportunity to study this relatively rare cancer while contrasting realistic interventions on MWF exposure in a longitudinal cohort setting. Though NHL accounts for a small proportion of cancer incidence of all sites, it has garnered intense research attention because its incidence nearly quadrupled since 1960.6 Immune modulation due to immunosuppressive therapy and infection by human immunodeficiency virus (HIV) are the most prominent risk factors of NHL, but they cannot fully explain this rise in NHL incidence.7,8 Epidemiologic investigations of environmental and occupational exposures have yielded varied results.6,9 This variation in epidemiologic findings may reflect the substantial etiologic heterogeneity of NHL, but studies of specific NHL subtypes are limited by the lack of quantitative exposure information and the lack of control for the healthy worker survivor effect (HWSE) or other forms of time-varying confounding affected by prior exposure.10,11 The HWSE, a key concern in occupational epidemiology, is the dynamic selection process by which healthier individuals remain at work, where they accumulate more occupational exposure, and less healthy individuals leave work.12 Standard analyses of occupational cohorts affected by the HWSE result in a downward bias on the exposure-outcome associations of interest.

In the present analysis of the UAW-GM Cohort Study, we estimated NHL risk from 1985 to 2005 under hypothetical limits on average annual soluble MWF by applying the novel hazard-extended iterative conditional expectation (ICE) parametric g-formula estimator.13 We leveraged time-varying quantitative MWF exposure data in tandem with employment data to adjust for the HWSE. In particular, we estimated the expected number of NHL cases we would observe if the hypothetical limits of 0.5, 0.25, and 0.05 mg/m3 were enforced for soluble MWF over workers’ entire working lifetimes. Since workers were exposed to MWF of other types as well, we also evaluated the effect of dynamically reducing soluble MWF exposure with the goal of limiting total MWF exposure to 0.5 and 0.25 mg/m3.

# Methods

## Study population

The UAW-GM Cohort includes hourly workers at three automobile manufacturing plants in Michigan who had worked at least 3 years by 1985. Past papers provide detailed descriptions of the cohort.14,15 The present study population (N = 34,738) was restricted to the autoworkers who were at work in 1941 or not yet hired, alive at the start of follow-up, and missing no more than half of their employment history. Autoworkers in the study population were followed for NHL incidence from January 1, 1995 until NHL diagnosis, death, December 31, 2004, or upon reaching the oldest observed age at death, whichever came earlier.

## Outcome and covariates

We identified incident cancers in the UAW-GM Cohort that occurred between 1985 and 2004 by linkage to the Michigan Cancer Registry (MCR). Workers at Plants 1 and 2, located in the greater Detroit metropolitan area, were also linked to the Detroit Regional Registry of the Surveillance, Epidemiology, and End Results (SEER) Program. Cancer types were distinguished using site and histology codes conforming to the International classification of Diseases for Oncology, 3rd edition (ICD-O-3). Non-Hodgkin lymphoma was defined by cancers with any of the following ICD-O-3 Histology codes: 9590-9597, 9670-9671, 9673, 9675, 9678-9680, 9684, 9687-9691, 9695, 9698-9702, 9705, 9708-9709, 9712, 9714-9719, 9724-9729, 9735, 9737-9738, 9811-9818, 9823, 9827, 9837. Details regarding cancer incidence follow-up are described elsewhere.16 Vital status was ascertained by company records and by linkage to Social Security Administration, National Death Index, death certificate, and state mortality files.

Covariates including year of hire, sex, race, and plant location were obtained from company records. Race data were missing for about 16% of the workers. In analyses, missing race was considered a distinct category of race.

## Exposure

For each combination of job type, department, and plant industrial hygienists collected several hundred personal and area samples for particulate matter (mg/m3) over the course of three decades to derive quantitative 8-hour time-weighted average exposure estimates to soluble, straight, and synthetic MWFs. Workers’ time-weighted average annual exposure to each MWF type was determined according to employment records, which recorded time-varying job type, department, and plant. For employment records that were no less than half complete, gaps in the record were interpolated by carrying forward the last known job type. In each year, cumulative exposures to soluble, straight, and synthetic MWFs (mg/m3-years) were calculated by taking cumulative sums of average annual exposures. In analyses, MWF exposure was lagged 10 years. Exposure assessment is described in detail elsewhere.17–19

## Statistical methods

We applied the hazard-extended ICE parametric g-formula with pooling over treatment history13 to estimate 20-year counterfactual risk of non-Hodgkin lymphoma under five hypothetical stochastic interventions on exposure to soluble MWFs with the elimination of censoring/competing risks. The first three interventions were static interventions that limited average annual exposure to soluble MWF to 0.5, 0.25, and 0.05 mg/m3. The remaining two were dynamic interventions that reduced average annual exposure to soluble MWF so that exposure to all MWF types was capped at 0.5 and 0.25 mg/m3, if possible. If those limits on total average annual exposure were not possible by intervention on soluble MWF alone, average annual exposure to soluble MWF was reduced to 0. Interventions were applied at hire, before the start of follow-up, through the end of follow-up. Person-years in which average annual exposure was below the hypothetical limits were not intervened upon. These hypothetical exposure limits reflect the Recommended Exposure Limit (REL) of the National Institute for Occupational Safety and Health (NIOSH) for total particulate mass from MWFs (0.5 mg/m3), half the REL (0.25 mg/m3), and a tenth of the REL (0.05 mg/m3).20

The hazard-extended ICE parametric g-formula estimator is an ICE expression of the parametric g-formula estimator, further extended to incorporate the predicted hazard of past outcomes. This differs from the classical ICE estimator, which incorporates past outcome data as indicators.13 Under the assumptions of conditional exchangeability at all time points, positivity, counterfactual consistency, and correct model specification, the hazard-extended parametric g-formula yields unbiased estimates of counterfactual risk with greater statistical efficiency than both propensity-score based estimators and the classical ICE g-formula.

Our implementation of the hazard-extended ICE parametric g-formula involved a series of model-based standardization steps using logistic regression. Post-intervention estimates of the discrete hazard of NHL given all exposures and covariates were combined iteratively from the end of follow-up to the start. In each iteration, predicted discrete hazards were standardized over post-intervention exposure and covariate histories before combining with discrete hazards from the previous iteration. The result of this iterative process was sequentially-standardized estimates of the NHL risk over the entire follow-up period. Averaging over the baseline distribution of covariates yielded the counterfactual risk estimate of NHL when the intervention of interest was enforced for the entire study population.

We split the 20-year follow-up period into eight periods; the first two periods spanned four years each, and the remaining six periods spanned two years each. Post-intervention exposure and exposure history were summarized as cumulative exposure. We modeled discrete hazards by fitting a pooled logistic regression for NHL over at-risk person-periods given cumulative exposure to straight, soluble, and synthetic MWFs, employment status, cumulative time off, year of hire, sex (male/female), race (Black/white/unknown), and plant (Plant 1/Plant 2/Plant 3). Cumulative exposure to MWFs, employment status, and cumulative time off were lagged 10 years. All continuous variables were represented as categorical variables with cut points determined by the tertiles of nonzero values among NHL cases. During the iterative combination of discrete hazards, we performed model-based standardization over baseline covariates and the complete set time-varying covariate histories.

We estimated risk under the observed distribution of soluble MWF exposure (natural course) and under the five interventions. We contrasted the risk under intervention to that under the natural course by computing relative risks. Confidence intervals were computed using the nonparametric bootstrap with 1000 samples and centering on the estimate computed from observed data. In secondary analyses (Appendix), we repeated the analysis to evaluate interventions on exposure to straight and synthetic MWFs.

# Results

Table 1 presents summary statistics of exposure and covariates for the full study population and for those diagnosed with NHL between 1985 and 2004. The cohort is predominantly white (66%) and male (87%). Plant 2 employed the largest number of workers while Plant 1 employed the fewest. The median year of hire among those diagnosed with NHL was 1959 whereas the median year of hire in the full study population was 1967. Age at hire was approximately the same among those with NHL and the full study population. Median lagged cumulative exposure to all three MWF types was higher among NHL cases. Soluble MWFs were the most widely used MWF type, with approximately 90% of workers ever exposed. Median cumulative exposure among exposed was 6.5 times higher for soluble than for straight MWFs. Figure 1 shows median average annual exposure to the three MWF types among exposed workers over calendar time. Exposure to MWF generally followed a downward trend over time. Figure 2 shows the post-intervention distribution of the logarithm of cumulative exposure to soluble MWF at end of follow-up if hypothetical limits on average annual exposure were enforced for all workers starting at hire. The interventions shifted the distribution of cumulative exposure rather than deterministically assigning exposure to a particular value.

Table 1: Summary of population characteristics. Statistics shown above the horizontal line are count (%). Those shown below are median (quartile 1, quartile 3).

|  | Study population | |  | NHL cases | |
| --- | --- | --- | --- | --- | --- |
| N (person-years) | 34,734 | (596,698) |  | 231 | (2,777) |
| Race |  |  |  |  |  |
| White | 22,789 | (66%) |  | 173 | (75%) |
| Black | 6,304 | (18%) |  | 21 | (9%) |
| Unknown | 5,641 | (16%) |  | 37 | (16%) |
| Sex |  |  |  |  |  |
| Male | 30,235 | (87%) |  | 206 | (89%) |
| Female | 4,499 | (13%) |  | 25 | (11%) |
| Planta |  |  |  |  |  |
| Plant 1 | 8,721 | (25%) |  | 68 | (29%) |
| Plant 2 | 14,258 | (41%) |  | 90 | (39%) |
| Plant 3 | 11,755 | (34%) |  | 73 | (32%) |
| Ever exposed to MWFsb |  |  |  |  |  |
| Straight | 19,905 | (57%) |  | 133 | (58%) |
| Soluble | 31,044 | (89%) |  | 210 | (91%) |
| Synthetic | 12,262 | (35%) |  | 72 | (31%) |
| Deceased by end of follow-up | 10,384 | (30%) |  | 33 | (14%) |
| Year of birth | 1940 | (1925, 1950) |  | 1929 | (1919, 1940) |
| Year of hire | 1967 | (1953, 1976) |  | 1959 | (1951, 1969) |
| Age at hire (years) | 23.6 | (20.0, 30.1) |  | 25.4 | (21.1, 33.6) |
| Year of leaving workc | 1979 | (1968, 1989) |  | 1977 | (1964, 1987) |
| Age at leaving work (years)c | 45.0 | (31.9, 57.7) |  | 53.4 | (36.5, 61.2) |
| Years at workc | 15.3 | (7.3, 27.1) |  | 19.2 | (8.0, 29.9) |
| Year of death | 2000 | (1993, 2008) |  | 2001 | (1994, 2005) |
| Age at death (years) | 74.7 | (65.3, 82.5) |  | 73.8 | (66.2, 82.0) |
| Cumulative time off (years)b | 1.05 | (0.30, 1.80) |  | 0.71 | (0.14, 1.40) |
| Cumulative exposure to MWFs (mg/m3-years)d | | | | | |
| Straight | 0.70 | (0.22, 2.56) |  | 0.93 | (0.29, 3.30) |
| Soluble | 4.65 | (1.85, 12.13) |  | 7.16 | (2.86, 20.91) |
| Synthetic | 0.45 | (0.16, 1.64) |  | 0.89 | (0.29, 2.11) |
| NHL: non-Hodgkin lymphoma. | | | | | |
| a Plant of longest employment duration among those who worked at multiple plants; b Lagged 10 years; c Among those who left work by December 31, 1994; d Among ever-exposed individuals, lagged 10 years. | | | | | |

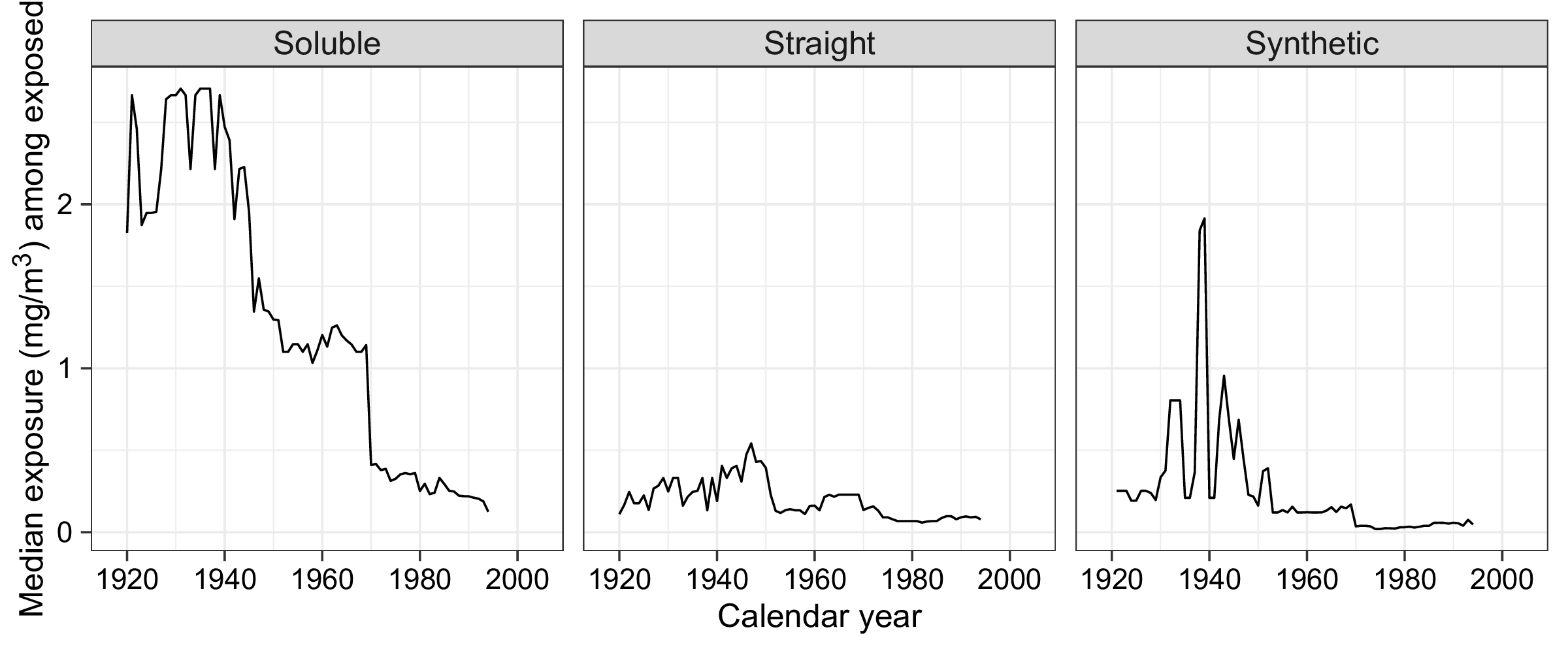


Figure 1. Median average annual exposure to straight, soluble, and synthetic metalworking fluids among exposed workers over time.

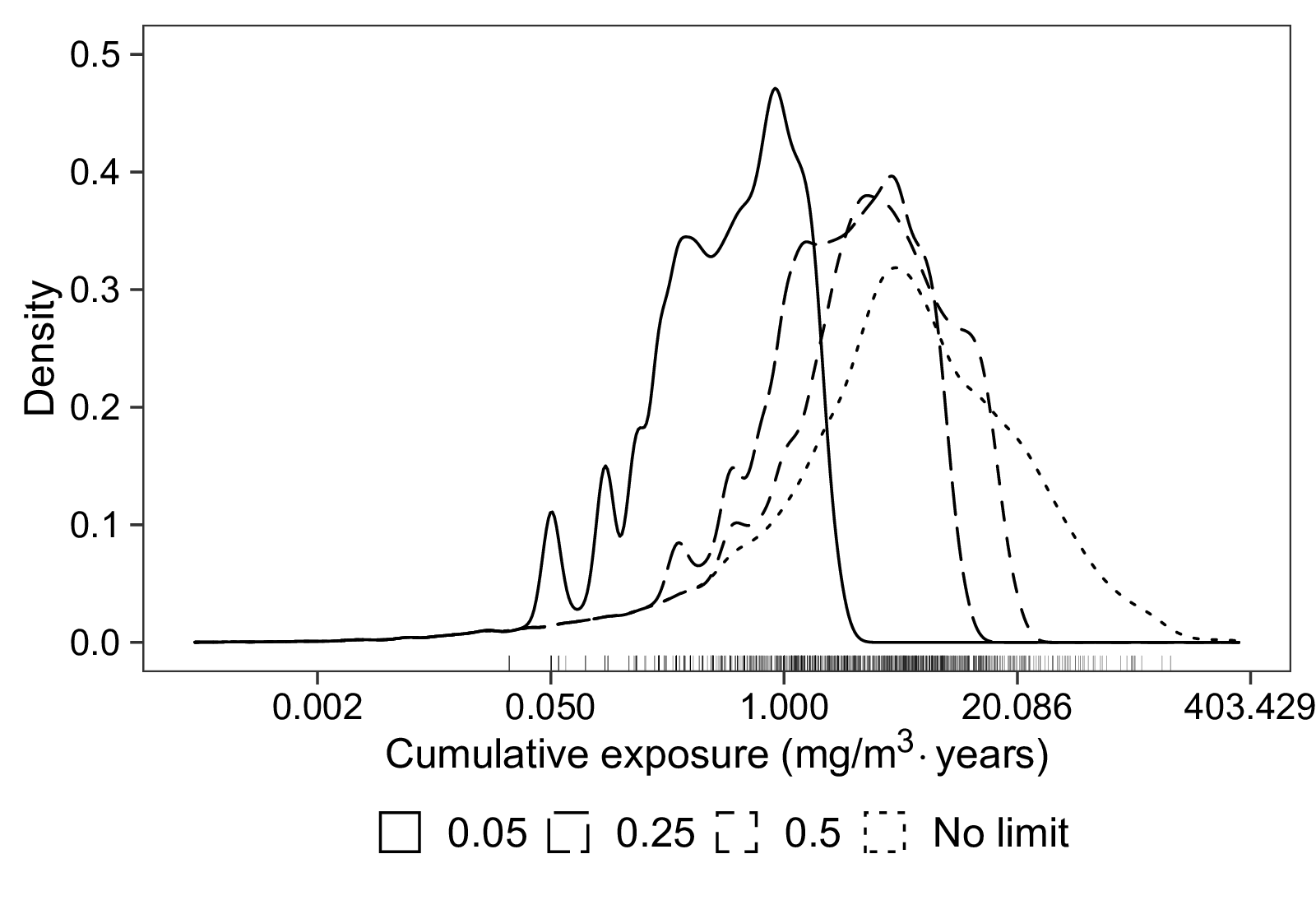


Figure 2. Post-intervention distribution of cumulative exposure to soluble metalworking fluids at end of follow-up under hypothetical limits on average annual exposure (mg/m3). The rugplot shows maximum cumulative exposure accrued by non-Hodgkin lymphoma cases.

The observed risk of NHL over the 20-year follow-up period was 6.65 per 1000. Table 2 presents the hazard-extended ICE parametric g-formula estimates of the risk and risk ratios contrasting hypothetical limits on exposure to soluble MWF to no limit, with elimination of competing risks throughout. The estimated risk under no limit on MWF exposure was 9.56 (8.15, 10.89) per 1000. Stronger limits on average annual exposure to soluble MWFs resulted in monotonically stronger reductions in the risk of NHL. Capping average annual exposure to soluble MWFs at the NIOSH REL of 0.5 mg/m3 and a tenth of the REL resulted in a risk of 8.30 (6.52, 10.19) and 7.52 (5.73, 9.51) per 1000, respectively. The risk ratios contrasting these hypothetical limits to no limit were 0.87 (0.72, 1.02) and 0.79 (0.62, 0.97). Dynamic reductions in soluble exposure with the aim of limiting total MWF exposure to the REL and half the REL also yielded protective risk ratios: 0.84 (0.69, 0.99) and 0.80 (0.64, 0.98). These dynamic interventions attain risk reductions of similar magnitude as the static interventions enforcing exposure limits at 0.25 and 0.05 mg/m3, but do so while intervening on a smaller proportion of person-years.

Results for interventions on straight and synthetic MWF are presented in Tables S1 and S2 in the Appendix. The relative risks contrasting interventions on straight and synthetic MWF exposure to the natural course were much smaller in magnitude than those for soluble; none were statistically significant. Since exposure to soluble MWF was greatest in prevalence and magnitude, limits on soluble MWF would result in interventions on a greater percentage of person-years than limits on straight or synthetic MWF. The intervention capping average annual exposure to soluble MWF at the REL affected 23.8% of the person-years in the follow-up period. Limiting straight and synthetic MWF exposure to the REL affected 3.2% and 1.1% of the person-years under follow-up, respectively.

Table 2: Counterfactual risks (per 1000) and risk ratios contrasting interventions on soluble MWF to the observed course.

| Exposure limit on soluble MWF (mg/m3) | Person-years intervened (%) | Risk per 1000 | (95% CI) | RR | (95% CI) |
| --- | --- | --- | --- | --- | --- |
| None | 0.0 | 9.56 | (8.15, 10.89) | 1.00 |  |
| 0.5 | 23.8 | 8.30 | (6.52, 10.19) | 0.87 | (0.72, 1.02) |
| 0.25 | 36.2 | 8.06 | (6.15, 10.15) | 0.84 | (0.68, 1.01) |
| 0.05 | 43.9 | 7.52 | (5.73, 9.51) | 0.79 | (0.62, 0.97) |
| max(0, 0.5 - str - syn) | 28.3 | 8.00 | (6.27, 9.87) | 0.84 | (0.69, 0.99) |
| max(0, 0.25 - str - syn) | 40.0 | 7.69 | (5.88, 9.64) | 0.80 | (0.64, 0.98) |
| MWF: metalworking fluid; str: exposure to straight metalworking fluids; syn: exposure to synthetic metalworking fluids. | | | | | |

# Discussion

We estimated counterfactual 20-year risks of NHL from 1985 to 2005 in the UAW-GM Cohort Study under different hypothetical interventions targeting exposure to soluble MWFs using the hazard-extended ICE parametric g-formula. We found a monotonic dose-dependent relationship with lower risk estimates arising from stronger limits. Contrasts in estimated NHL risk were statistically significant under the strongest static intervention, which limited soluble MWF exposure to 0.05 mg/m3, and the two dynamic interventions, which reduced exposure so that total MWF exposure would be capped at 0.5 or 0.25 mg/m3, if possible. The dynamic interventions, which take into consideration exposure to the other MWF types, may be more efficient at reducing risk of NHL than the static interventions, which ignore exposure to the other MWF types.

Colbeth et al. first reported an association between NHL and MWF exposure in this cohort in their standard survival analyses, which investigated associations between MWF exposure and several cancer incidence outcomes.16 We verified and further examined this association in an approach capable of adjusting 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. Under the assumptions of conditional exchangeability, positivity, consistency, and correct model specification, our estimates are unbiased for the true counterfactual risk under the hypothetical interventions.

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 affecting employment status or time off work.21 We expect the absence of good time-varying measures of worker health over the life course to result in bias toward the null.

Positivity refers to the need for adequate variation in future exposure among strata formed by 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.22 We investigated static and dynamic stochastic interventions on soluble MWF exposure which intervened only when average annual exposure exceeded the hypothetical limit under consideration. Hence, 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 with only 8 follow-up periods and by representing covariates using fewer categories. Coarsening limits the comprehensiveness of confounding control, but improves positivity. In practice, causal inference using observational data must always balance positivity, covariate adjustment, and model specification.23 When estimating causal effects using estimators that do not require exposure modeling, as is the case with g-formula methods, 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.24,25 This assumption would be violated if there were multiple versions of treatment causally associated with different outcomes. This basic notion of consistency is violated in our analysis because 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 the MWF over the course of its use.26 However, causal effects estimates under violations in the consistency assumption are still valid and unbiased if there is adequate adjustment for confounders of the exposure-version relationship.25 This may be thought of as conditional consistency within strata, in which there is only one version of treatment. Our analysis indexed time periods over calendar time and adjusted for age, year of hire, and plant. In this way, we limited potential for bias due to variation in MWF composition.

Correct model specification is a standard assumption in all analyses. The estimator we applied offers greater statistical efficiency than the classical ICE parametric g-formula estimator because it leverages greater parametric smoothing. In causal analyses of longitudinal cohort studies, both the hazard-extended and classical ICE parametric g-formula estimators are less common than the non-iterative expression (NICE) of the parametric g-formula.27 However, a major limitation of the NICE g-formula 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.28,29 As with all ICE g-formula estimators, the estimator we applied is not subject to the g-null paradox. Furthermore, simulation studies show that the variance of the hazard-extended ICE parametric g-formula is similar to that of the NICE parametric g-formula, so we expect the former to be no less conservative than the latter.13

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.30–33 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.34 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. Organic compounds containing phosphorous, chlorine, sulfur, nitrogen, and boron are commonly added to soluble MWF to control microbial growth, improve performance under high heat/pressure, and inhibit corrosion.35 Organophosphorous compounds include organophosphate pesticides, which have been linked to cancer risk in epidemiologic and animal studies. Some were classified as possibly carcinogenic by the IARC.5 Studies of occupational exposure to chlorinated solvents and pesticides have also been linked to NHL risk.36–40 In 2014, the International Agency for Research on Cancer classified trichloroethylene, tetrachloroethylene, and other chlorinated agents as Group 1 carcinogens.4 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.41 The structural characteristics shared by MWF additives and known/suspected carcinogens suggest similarities in molecular function and carcinogenesis.

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/m3.20 Without enforcement of the REL, we still observed average annual exposures below the REL for many person-years. If the REL were enforced in the real world, we would not expect reductions in exposure for these low-exposure person-years. Hence, contrasting the counterfactual scenario where all workers experienced annual exposure at the REL to one where all workers experienced annual exposure at some higher level would result in an overestimate of the expected real-world benefit of REL enforcement. By considering stochastic interventions that result in exposure distributions more similar to what we would expect in the real world, our analysis yielded results that more realistically quantify the effects of policy than those investigating deterministic interventions.42,43

Associations between several occupations and risk of NHL have been reported previously, but none evaluated the potential effect of hypothetical limits on occupational exposures.6,8,33,40 We found evidence that exposure to soluble MWF was associated with NHL incidence after adjustment for time-varying confounding affected by prior exposure using the hazard-extended ICE parametric g-formula. Shifting the distribution of cumulative exposure to soluble MWF by limiting average annual exposure resulted in reductions in NHL incidence. Limits on exposure to soluble MWF that take into consideration exposure to straight and synthetic MWFs led to greater reductions NHL incidence while intervening on fewer person-years.

# Citations

1. Childers J. The chemistry of metalworking fluids. In: *Metalworking Fluids*. CRC Press; 2006.

2. IARC. *IARC Monographs on the Evaluation of Carcinogenic Risk of the Chemical to Man: Certain Polycyclic Aromatic Hydrocarbons and Heterocyclic Compounds*. Vol 3. World Health Organization International Agency for Research on Cancer; 1973.

3. IARC. *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans: Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs*. Vol 1-42. World Health Organization International Agency for Research on Cancer; 1987:106-116.

4. 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.

5. 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.

6. Ekström-Smedby K. Epidemiology and etiology of non-hodgkin lymphoma–a review. *Acta oncologica*. 2006;45(3):258-271.

7. Shiels MS, Engels EA, Linet MS, et al. The epidemic of non-hodgkin lymphoma in the united states: Disentangling the effect of HIV, 1992–2009. *Cancer Epidemiology and Prevention Biomarkers*. 2013;22(6):1069-1078.

8. Chiu BCH, Hou N. Epidemiology and etiology of non-hodgkin lymphoma. In: Evens AM, Blum KA, eds. Vol 165. Springer; 2015.

9. Alexander DD, Mink PJ, Adami H-O, et al. The non-hodgkin lymphomas: A review of the epidemiologic literature. *International Journal of Cancer*. 2007;120(S12):1-39.

10. Evens AM, Chiu BC-H. The challenges of epidemiologic research in non-hodgkin lymphoma. *JAMA*. 2008;300(17):2059-2061.

11. Morton LM, Sampson JN, Cerhan JR, et al. Rationale and design of the international lymphoma epidemiology consortium (InterLymph) non-hodgkin lymphoma subtypes project. *J Natl Cancer Inst Monogr*. 2014;2014(48):1-14. doi:[10.1093/jncimonographs/lgu005](https://doi.org/10.1093/jncimonographs/lgu005).

12. 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>.

13. Wen L, Young JG, Robins JM, Hernán MA. Parametric g-formula implementations for causal survival analyses. *Biometrics*. 2020.

14. Eisen EA, Tolbert PE, Monson RR, Smith TJ. Mortality studies of machining fluid exposure in the automobile industry I: A standardized mortality ratio analysis. *American journal of industrial medicine*. 1992;22(6):809-824.

15. Eisen EA, Bardin J, Gore R, Woskie SR, Hallock MF, Monson RR. Exposure-response models based on extended follow-up of a cohort mortality study in the automobile industry. *Scandinavian journal of work, environment & health*. 2001;27(4):240-249.

16. Colbeth HL, Chen KT, Picciotto S, Costello S, Eisen EA. Metalworking fluids and cancer incidence in the united autoworkers-general motors cohort. *American Journal of Epidemiology*. 2022, in press.

17. Hallock MF, Smith TJ, Woskie SR, Hammond SK. Estimation of historical exposures to machining fluids in the automotive industry. *American Journal of Industrial Medicine*. 1994;26(5):621-634. doi:[10.1002/ajim.4700260505](https://doi.org/10.1002/ajim.4700260505).

18. Woskie SR, Smith TJ, Hallock MF, et al. Size-selective pulmonary dose indices for metal-working fluid aerosols in machining and grinding operations in the automobile manufacturing industry. *American Industrial Hygiene Association Journal*. 1994;55(1):20-29.

19. Woskie SR, Virji MA, Hallock M, Smith TJ, Hammond SK. Summary of the findings from the exposure assessments for metalworking fluid mortality and morbidity studies. *Applied occupational and environmental hygiene*. 2003;18(11):855-864.

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

21. 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.

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

23. 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).

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

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

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

27. Keil AP, Edwards JK, Richardson DB, Naimi AI, Cole SR. The parametric g-formula for time-to-event data: Intuition and a worked example. *Epidemiology*. 2014;25(6). <https://journals.lww.com/epidem/Fulltext/2014/11000/The_Parametric_g_Formula_for_Time_to_event_Data_.16.aspx>.

28. 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.

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

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

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

32. 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.

33. ‘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.

34. 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.

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

36. 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.

37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

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