Web Appendix 1

*A note about the interventions*

The interventions considered here are known as "representative" interventions.[1] They are mathematically equivalent to re-assigning each person-year with exposure above the binary cutoff to an exposure below the binary cutoff at random according to the observed distribution of exposures below the binary cutoff.

*G-estimation details and confidence intervals*

The analysis was based on the following structural accelerated failure time model:

$$T_{\overline{0}} = \int_0^T \exp[\psi A(t)]\, dt \qquad (1)$$

where $T$ is observed survival time, $A(t)$ is observed exposure at time $t$, $T_{\overline{0}}$ is the counterfactual survival time if never exposed, and $\psi$ is the unknown coefficient to be estimated.[2,3] G-estimation of $\psi$ was achieved by solving an estimating equation (see Hernán et al, 2005[2]) using a binary grid search within the search interval -0.9 to 0.9, a range that included all plausible (and many implausible) values for the coefficient.

It would have taken prohibitively long to run 1000 bootstrap samples from which to construct non-parametric, quantile-based confidence limits. We instead assumed the bootstrap estimates to be normally distributed, and used the standard deviation of the 200 bootstrap estimates to compute the confidence interval for each outcome and exposure cutoff. The confidence limits reported are equal to the point estimate plus or minus 1.96 times the standard deviation of the bootstrap estimates. Most of these confidence limits were qualitatively quite similar to the non-parametric confidence intervals (not

shown) obtained by using the 2.5th and 97.5th percentiles of the bootstrap estimates as the lower and upper limits of the 95% confidence interval.

For some bootstrap samples, g-estimation failed to produce an estimate within that interval. For example, this could happen when too few cases who had ever (or never) been exposed above the cutoff were selected into the sample. Since we had fewer than 200 bootstrap estimates, using the standard deviation to calculate confidence limits was inappropriate. Thus, for each cutoff where at least one bootstrap sample (but fewer than 5% of them) failed to produce an estimate, the minimum and maximum estimates were reported instead of adding and subtracting 1.96 times the standard deviation, yielding a possibly conservative confidence interval. This occurred only in the analyses of cerebrovascular disease, for cutoffs 0.05mg/m$^3$, 0.1mg/m$^3$, and 0.15mg/m$^3$.

*Effect measure estimated*

The coefficient $\psi$ represents the log of the ratio of the median survival time that would have been observed if all workers had been exposed for the entire duration of follow-up to the median survival time if all workers had never been exposed. However, this interpretation is not very meaningful in our application: the autoworkers cohort includes follow-up after the end of employment, when exposure can no longer occur, so that "always exposed" is an unrealistic scenario. We reported estimates for a different effect measure derived from the same coefficient: the sum over all deaths from the cause of interest of the difference $T_{\bar{0}} - T$, where $T_{\bar{0}}$ is calculated for each person from the accelerated failure time model (equation (1)) using the g-estimate of $\psi$. Thus, we use the estimate of $\psi$ to calculate the impact of each intervention within the subgroup of the cohort that experienced the outcome of interest during follow-up.

*Administrative censoring*

Since the autoworkers cohort was not followed until all workers had died, administrative censoring would cause bias in the absence of proper adjustment, because survival to the administrative end of follow-up may depend on exposure. We adjusted by artificially censoring those workers whose death from the cause of interest would have been unobserved under at least one possible exposure scenario, as described in the references.[2,3] For most analyses, fewer than 5% of the observed deaths were artificially censored, though more artificial censoring was necessary for cerebrovascular disease than for the other outcomes. For each cutoff, Table A1 displays the number and percent of deaths from each cause that were *not* artificially censored.

Our analysis only includes those workers whose outcome would have been observed under all exposure histories. Thus, under the assumptions described in the Discussion, our g-estimates of $\psi$ are unbiased, and can be thought of as representing the etiologic effect. For each analysis, we then used the g-estimate to calculate the estimated total number of years of life saved among all individuals whose deaths from the cause of interest were observed during the study period, including those artificially censored in the g-estimation step (even though some of those saved life years were after the administrative end of follow-up). Those who were administratively censored (that is, who were still alive at the end of follow-up) were not included in this sum. Different workers may have been artificially censored in the analyses of different cutoffs. Including the artificially censored cases was therefore necessary in order to obtain comparable estimates for different cutoffs, since our sums must refer to the same population for all cutoffs.

*Non-administrative censoring*

We adjusted for loss to follow-up (<4%) and censoring by death from competing risks (13% for cardiovascular disease overall, 17% for IHD, 20% for AMI, and 23% for cerebrovascular disease) using inverse probability of censoring weights.[2] We used stabilized weights equal to the predicted marginal

probability of remaining uncensored divided by the predicted conditional probability of remaining uncensored. The conditional probability of being uncensored at each time was obtained from a pooled logistic model using the following predictors: race, sex, current age and age at baseline (both linear), plant, an indicator for the year being before 1970 (when exposures were dramatically reduced), both current and previous exposures to all three fluid types as described in the text for the model predicting exposure, and intermittent time off work. For all outcomes and cutoffs, the weights were moderate (total range: 0.88 to 1.44) and did not require truncation.

Our use of inverse probability weighting to adjust for competing risks gives our results for cause-specific mortality a somewhat odd interpretation: we estimate the impacts of interventions as if it were impossible to die from causes other than cardiovascular disease (or ischemic heart disease, or acute myocardial infarction, or cerebrovascular disease). This scenario is unrealistic, but it is the only way to avoid assuming that death from competing risks is uninformative within the context of g-estimation of an accelerated failure time model. However, a sensitivity analysis without adjustment for competing risks (that is, without using inverse probability weights to adjust for competing risks) yielded very similar estimates, most of which were slightly higher.

*All-cause vs. cause-specific mortality*

The accelerated failure time model in Equation 1 relates counterfactual unexposed survival time to observed exposure and observed survival time via an unmeasured coefficient that can be thought of as the log of the "acceleration factor". Exposure speeds up (or slows down) the time to event, rather than increasing (or decreasing) one's probability of experiencing the outcome. For this reason, the accelerated failure time model is very well-suited to studying all-cause mortality. To illustrate this point, we ran an analysis of all-cause mortality using the same exposure cutoffs. No weights (or assumptions) were needed to adjust for competing risks, since there were none, though we did use weights to adjust

for loss to follow-up. In this analysis, the coefficient $\psi$ from the main analysis represents the log of the factor by which exposure accelerates death in the cohort; this etiologic measure applies to the entire cohort.

For cause-specific mortality, the coefficient $\psi$ from the main analysis represents the log of the factor by which exposure accelerates death from that cause in a pseudo-population in which death from other causes is impossible. The sensitivity analysis without weights for competing risks yields a coefficient representing the log of the factor by which exposure accelerates death from that cause in the cohort, under the assumption that those censored due to death from other causes either (a) share the same distribution of exposure history with or (b) have the same underlying susceptibility to the outcome of interest as those who died of the cause of interest. This complicated interpretation of $\psi$ provides a further motivation for presenting the results as impacts of interventions within subgroups defined by observed outcomes.

*Choice of g-methods*

Another method that correctly handles time-varying confounding affected by prior exposure is the parametric g-formula.[4] This method, rather than adjusting for competing risks, models the outcome, competing outcomes, and all covariates as functions of prior values. Then, the models are used to predict the data distribution under different exposure scenarios; competing outcomes can thus be simulated, possibly giving a more realistic idea of the impact of an intervention. However, this method requires making parametric modeling assumptions not only for the outcome but for every time-varying covariate, which could result in substantial bias if those models are misspecified, especially if errors propagate over a long follow-up period. Unlike g-estimation, the parametric g-formula is also subject to the g-null paradox, meaning that results are expected to be biased in large samples.[5]

Marginal structural models with inverse probability of treatment weighting also adjust correctly for time-varying confounding affected by prior exposure. However, that method gives biased results when positivity is violated. Positivity, or experimental treatment assignment, is the assumption that all non-empty strata of covariates include both exposed and unexposed individuals. Workers who are no longer actively employed cannot be exposed, which represents a structural violation of positivity. G-estimation of our structural model does not require positivity.[5]


1. Taubman SL, Robins JM, Mittleman MA, Hernan MA. Alternative approaches to estimating the effects of hypothetical interventions. Alexandria, VA: American Statistical Association; 2008.

2. Hernan MA, Cole SR, Margolick J, Cohen M, Robins JM. Structural accelerated failure time models for survival analysis in studies with time-varying treatments. Pharmacoepidemiol Drug Saf 2005;14:477-91.

3. Chevrier J, Picciotto S, Eisen EA. A comparison of standard methods with g-estimation of accelerated failure-time models to address the healthy-worker survivor effect: application in a cohort of autoworkers exposed to metalworking fluids. Epidemiology 2012;23:212-9.

4. Cole SR, Richardson DB, Chu H, Naimi AI. Analysis of occupational asbestos exposure and lung cancer mortality using the g formula. Am J Epidemiol 2013;177:989-96.

5. Robins JM, Hernán MA. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, eds. Longitudinal Data Analysis. New York: Chapman & Hall / CRC; 2009:553-99.

Web Table 1. Number and percent of deaths not artificially censored.

| | All-cause (N=9539) | | Cardiovascular disease (N=4153) | | IHD (N=2612) | | AMI (N=1699) | | Cerebrovascular disease (N=501) | |
|---|---|---|---|---|---|---|---|---|---|---|
| Cutoff | N | % | N | % | N | % | N | % | N | % |
| 0 | 9205 | 96 | 3988 | 96 | 2461 | 94 | 1606 | 95 | 460 | 92 |
| 0.01 | 9387 | 98 | 4056 | 98 | 2529 | 97 | 1641 | 97 | 461 | 92 |
| 0.02 | 9430 | 99 | 4092 | 99 | 2548 | 98 | 1664 | 98 | 466 | 93 |
| 0.03 | 9441 | 99 | 4112 | 99 | 2559 | 98 | 1666 | 98 | 473 | 94 |
| 0.04 | 9430 | 99 | 4127 | 99 | 2571 | 98 | 1681 | 99 | 490 | 98 |
| 0.05 | 9458 | 99 | 4143 | 100 | 2579 | 99 | 1684 | 99 | 500 | 100 |
| 0.10 | 9470 | 99 | 4123 | 99 | 2569 | 98 | 1679 | 99 | 459 | 92 |
| 0.15 | 9520 | 100 | 3971 | 96 | 2592 | 99 | 1695 | 100 | 333 | 66 |

IHD: ischemic heart disease; AMI: acute myocardial infarction