TITLE PAGE

PI: Sally Picciotto
UC Berkeley School of Public Health
Environmental Health Sciences
50 University Hall
Berkeley, CA 94720-7360
(510) 643-5716
sallypicciotto@berkeley.edu

"G-estimation methods and applications to quantitative exposure"

R03 OH010202

Dates: 07/01/2012 – 12/31/2014

Date submitted: 03/31/2015

Table of Contents

List of terms and abbreviations

AMI: Acute myocardial infarction

CI: confidence interval

IHD: ischemic heart disease

OEL: Occupational Exposure Limit

UAW-GM: United Autoworkers – General Motors

G-estimation methods and applications to quantitative exposure

PI: Sally Picciotto
UC Berkeley School of Public Health
Environmental Health Sciences
50 University Hall
Berkeley, CA 94720-7360
(510) 643-5716
sallypicciotto@berkeley.edu

In occupational studies, workers more susceptible to the outcome are likely to leave work (or even experience the outcome) earlier, thereby accumulating less exposure—a phenomenon that gives rise to healthy worker survivor bias. This R03 aimed to examine the relationships of exposure to straight (oil-based) metalworking fluids with all-cause and cardiovascular mortality in a cohort of autoworkers, using statistical methods that account for this bias while also incorporating quantitative exposure data.

Traditional Cox regression analysis failed to detect any relationship between straight fluids and cardiovascular mortality, while the exposure appeared slightly protective for all-cause mortality. This is because conventional regression analysis cannot adjust correctly for employment status, a time-varying confounder that affected by prior exposure--even though failure to adjust also leads to bias.

In contrast to traditional regression analysis, g-estimation of a structural accelerated failure time model treats each year of exposure separately and adds together the estimated effects, rather than adding together the exposures over time and attempting to estimate the effect of the cumulative exposure. This allows it to disentangle exposure and confounders over time in a way that traditional regression cannot do. Previous work with g-estimation had suggested that longer durations of straight metalworking fluid exposure shortened the lives of those who died from ischemic heart disease. The present work considered the intensity of exposure, not just its duration, and considered a few related outcomes.

One set of analyses focused on a public health framework by considering a series of hypothetical occupational exposure limits (OELs) and reporting the total number of years of life that would have been saved among those who died of the cause of interest if those limits had been enforced. We estimated that banning straight metalworking fluids could have saved 6047 years of life among those who died during the study period (1941-1994). When we considered more specific outcomes, where the etiologic relationship appeared to be stronger, the estimates were 4003 life-years for IHD, 2932 for AMI, and 916 for cerebrovascular disease, which was the rarest of the outcomes considered. Enforcing exposure limits higher than zero, even if very low, resulted in considerably lower impact (than a ban) on all outcomes except for cerebrovascular disease.

A separate set of analyses made a different assumption about the shape of the exposure-response curve by entering exposure into the outcome model as either a categorical variable treated as continuous, or (when possible) as a continuous variable, rather than a binary variable corresponding to an OEL. These results were consistent with those found above, though some outcomes showed attenuation in the slopes due to the etiologic importance of even low levels of exposure. Cerebrovascular disease did not show this attenuation, which corresponds to the fact that exposure limits above zero still had some impact on this outcome.

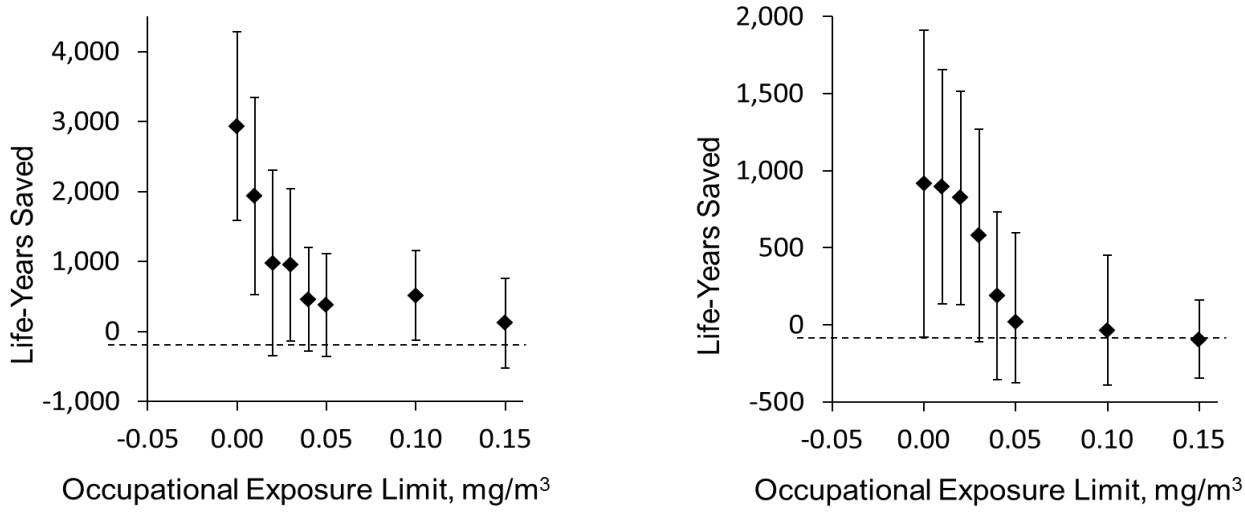
Overall, our results support lower occupational exposure limits for straight metalworking fluids.

Section 1 (2 page limit)

Significant (key) findings

We investigated the impacts of alternative exposure limits for oil-based metalworking fluids on time to death from cardiovascular disease (Aim 3). We developed a public health framework that built on the structural accelerated failure time models already in use, and applied it to examine a series of hypothetical occupational exposure limits (OELs) in relation to all-cause mortality and to deaths from cardiovascular disease, ischemic heart disease (IHD), acute myocardial infarction (AMI), and cerebrovascular disease. We found that 6047 years of life could have been saved by banning the fluids, but this impact dropped to 2674 years of life if the exposure limit were set to 0.01mg/m^3 and continued to drop lower for increasing limits. For cardiovascular disease, the pattern was similar, though the decrease was slower and statistical significance was lost sooner. Similar patterns were observed for IHD and AMI. The impact on cerebrovascular disease followed a different pattern. Here, the number of years of life saved was about 800-900 for exposure limits from 0 (a ban) to 0.03mg/m^3 before dropping off to null rather abruptly. These results provide some of the first evidence that straight metalworking fluids may be related to cerebrovascular disease. See Figure 1.1.

Figure 1.1. Estimated total number of years of life (with 95% confidence interval bars) that could have been saved by enforcing various occupational exposure limits for straight metalworking fluids, for AMI mortality (left) and cerebrovascular mortality (right). Note that the scales on the y-axis are different in the two graphs.

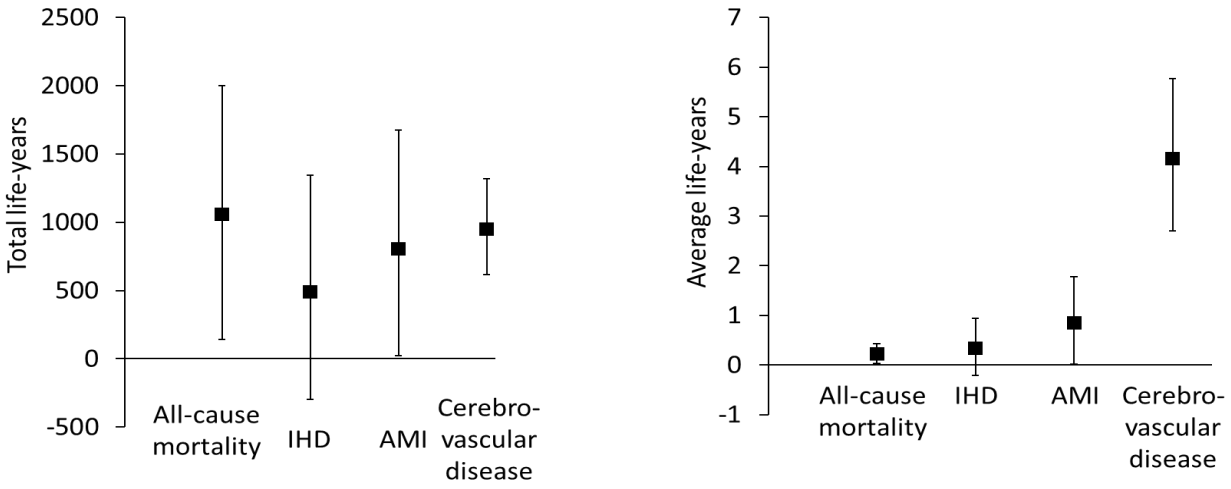


Aim 1 was to develop a way to use continuous exposure in a structural accelerated failure time model. We eventually settled on a method that modeled the effect of exposure on survival time in each year separately and then added them together. This model had not previously been applied in any occupational study using quantitative annual exposure.

Aim 2 involved applying the model from Aim 1 to the United Autoworkers – General Motors cohort to evaluate the relationship between oil-based metalworking fluids and survival time to death from cardiovascular causes. Results supported those from Aim 3 for all-cause mortality and IHD. Due to convergence issues for AMI and cerebrovascular mortality, exposure was converted to categories that were treated as continuous in a separate set of analyses. Results using categorical exposure supported the findings from Aim 3 for cerebrovascular

disease, while the slopes were attenuated for the other outcomes. This attenuation is consistent with the patterns observed in Aim 3. See Figure 1.2.

Figure 1.2. Estimated total (left) and average (right) number of years of life that would have been saved by banning straight metalworking fluids (among exposed workers who died of various causes) using categorical exposure measure.



Aim 4 was to compare our results to those obtained from conventional regression analysis. Conventional Cox models yielded null or protective associations between straight fluids and all of these outcomes, illustrating the importance of using methods that adjust correctly for employment status and other time-varying confounders affected by prior exposure.

Translation of findings

Our findings that straight metalworking fluids reduce life expectancy overall and specifically for cardiovascular outcomes (most notably cerebrovascular disease and AMI) provide evidence to inform OSHA regulations and industry standards for these fluids. Specifically, our analyses support a lower OEL for oil-based metalworking fluids than those currently in effect.

More broadly, the analytic method and public health framework developed and applied in this project are widely applicable in occupational epidemiology, and this project has helped to determine when they could be most useful. G-estimation of an accelerated failure time model is most appropriate when considering inevitable outcomes, such as all-cause mortality, or common (combined) outcomes such as all cancer incidence/mortality, all cardiovascular mortality, etc. For rarer outcomes, alternative g-methods may be preferable.

Outcomes/Impact

Potential outcomes from this work are new regulations or recommendations that would lower workplace exposure to straight metalworking fluids. Our research indicates that many years of life are lost to cardiovascular mortality as a result of exposure to these fluids. Regulators and industries should be aware of the relationships between straight metalworking fluids and these common outcomes in order to create and enforce effective occupational exposure limits.

Section 2

Scientific report

Background: When workers with poor health decrease their exposure but healthy workers do not, it becomes difficult to detect an association even when exposure causes disease. This bias is called the healthy worker survivor effect, and it is particularly problematic in studies of relationships between long-term occupational exposures and chronic diseases.

There are no satisfactory solutions to the healthy worker survivor effect using conventional methods for analyzing longitudinal occupational data. Fortunately, alternative methods have been developed that can eliminate this type of bias. However, only in the past few years have these methods been applied in an occupational setting, despite several authors' attempts over the previous 15 years to encourage their adoption. Our group previously applied them to a large mortality study, the United Autoworkers – General Motors (UAW-GM) cohort, and found a clear relationship between duration of exposure to oil-based metalworking fluids and heart disease. Our analysis was the first evidence in the literature for this association.

However, these methods for controlling the healthy worker survivor effect and similar biases (known as g-methods) have mostly focused on binary annual exposure measures. If workers are exposed at various levels, higher exposures may be more likely to cause disease, or may cause disease sooner. Unlike binary exposure measures, quantitative exposure measures can distinguish etiologically relevant levels of exposure. Quantitative exposure-response characterization is also necessary for risk assessment providing guidance for policy. Our goal was, therefore, to extend causal methods in order to evaluate the effect of total quantitative exposure on survival time, without bias due to the healthy worker survivor effect, and then apply these methods to the UAW-GM autoworkers cohort. We specifically focused on g-estimation of accelerated failure time models.

Early in the funding period of this grant, while preparing to work on the specific aims, we did some exploratory analyses that did not directly address the specific aims, but that led to a clearer understanding of how g-estimation of an accelerated failure time model handles employment status. This led to a paper published in OEM.[1] It details some subtle points about how g-estimation adjusts for different time-varying confounders affected by prior exposure, such as short-term leaves of absence and employment status. For that paper, we were interested in the difference between (a) continuing follow-up past employment termination and (b) censoring at termination as happens in actively employed cohorts. Using g-estimation of an accelerated failure time model with binary (ever vs. never) exposure in each year, we analyzed the data from the autoworkers cohort, and then reanalyzed them after censoring follow-up at termination of employment. With full follow-up, we observed elevated hazards for all-cause mortality, ischemic heart disease, or all cancers combined, even if exposure duration was only 5 years (compared to the entire population being never exposed). However, these relationships appeared null or protective when we censored workers at employment termination. We concluded that in order to detect harmful effects of workplace exposures, studies of actively employed workers should use inverse probability weighting to adjust for censoring, which in turn requires measuring a well-chosen health-related variable that predicts leaving work.

Specific Aim 1: *Extend methods for causal analysis using g-estimation to quantitative exposures. In particular, extend a recently developed multiplicative model so that a final unit-year of exposure decreases survival time by a fixed factor.*

Methodology. The "multiplicative model" we initially proposed to use turned out to have some important limitations that rendered it unsuitable for this application. Our second attempt considered cumulative exposure in an already-established structural model, so that a final unit-year of exposure would decrease remaining survival time by a fixed factor, but closer attention showed that this model required etiologic assumptions that were not appropriate for a study of cardiovascular outcomes. Specifically, exposures that occurred earlier in follow-up had more impact on survival time than exposures occurring late in follow-up; another group did publish a paper using this method in a study of lung cancer.[2] While this emphasis on earlier exposures may be reasonable for cancer outcomes, in which latency is expected, we decided that a model avoiding that assumption would be better for the cardiovascular outcomes we were exploring here.

We therefore used an established accelerated failure time model but substituted a quantitative annual (rather than binary annual or quantitative cumulative) exposure, which had never been applied to an occupational cohort before. In contrast to traditional regressions, g-estimation of our structural accelerated failure time model treats each year of exposure separately and adds together the estimated effects, rather than adding together the exposures and attempting to estimate the effect of the cumulative exposure. This disentanglement of exposure over time is how a structural nested model avoids the bias that occurs in traditional regressions when stratifying on a time-varying confounder that is affected by prior exposure.

For the more specific outcomes in our application, we ran into convergence issues, so we also developed a model that substitutes exposure category (0 or decile) but treats it as a continuous variable in the accelerated failure time model.

The structural model considers only the relationship between the hypothetical survival time if unexposed and the observed exposure and survival time; this relationship is quantified by an unknown coefficient to be estimated. G-estimation relies on a model predicting the exposure in order to adjust for confounding. When using *either* quantitative or categorical exposure in the structural model for the outcome (survival time), we predicted category of exposure level (0 or decile) in each year using a pooled multinomial logistic model based on baseline and time-varying covariates.

The structural accelerated failure time model used in the analysis is:

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

where T is observed survival time, $A(t)$ is observed exposure at time t measured in mg/m^3, $T_{\bar{0}}$ is the counterfactual survival time if never exposed, and ψ is the unknown coefficient to be estimated.[3] [For the categorical exposure analysis, $A(t)$ is observed exposure at time t as a value between 0 (unexposed) and 10 (highest decile of exposure); everything else remains the same except the search interval.]

Follow-up ended before all workers had died. Therefore, administrative censoring would cause bias in the absence of proper adjustment, because exposure may have affected probability of survival to the administrative end of follow-up. We addressed this issue 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.[3,4] 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.

Assuming there are no unmeasured confounders, then counterfactual outcomes should be unrelated to observed exposure (conditional on measured confounders). G-estimation of ψ was therefore achieved by including the binary variable $\Delta(\psi)$ indicating whether the worker's death would have been observed under all possible exposure scenarios in the model for exposure, and searching for the value of ψ that resulted in $\Delta(\psi)$ having a coefficient of zero in the exposure model (or, equivalently, having a *p*-value of 1).

The exposure model was a multinomial logistic regression carried out using PROC GENMOD in SAS. This regression had to be repeated for each candidate value of ψ, making the process somewhat computationally intensive. When exposure was entered into the model for survival time as a continuous variable, the search was conducted in the interval [-1.5, 2.5], first by a coarse check for monotonicity of the coefficient of $\Delta(\psi)$, in which 11 equally-spaced values of ψ were checked, and then by a binary search algorithm in the subinterval where the coefficient changed sign. If necessary to obtain confidence intervals, the search interval was widened. When exposure was categorical, the search interval was [-0.15, 0.15].

We chose multinomial logistic regression for the exposure model because exposure is determined by job, which is inherently more of a qualitative category than a ranking by exposure level. Jobs requiring similar skills may have very different levels of exposure to straight metalworking fluids, because different facilities, machines, and operations used different fluid types or different mixtures of the fluid types. Multinomial logistic regression is therefore more suitable than linear regression or ordinal logistic regression for predicting exposure in this cohort.

Interpretation of ψ

If exposure is constant over all of follow-up, then Equation (1) reduces to

$$T_{\bar{0}} = T \exp[\psi A(t)] \qquad (2)$$

where $A(t)$ is the constant exposure and does not depend on t. If we set $A(t) = 1$, then we have

$$\exp[\psi] = \frac{T_{\bar{0}}}{T}. \qquad (3)$$

Thus, the coefficient ψ represents the log of the ratio of survival times comparing no exposure to constant exposure at a concentration of 1mg/m^3.

While higher values of ψ represent stronger etiologic relationships, interpreting the coefficient in the literal way described above does not describe a meaningful comparison in our cohort, since only those who are actively employed can be exposed in each year. We therefore used Equation 1 to calculate $T_{\bar{0}}$ for each individual, using the estimate of ψ. The difference between this and the observed survival time is the estimated number of years of life that would have been saved for that individual, if they had never been exposed and if they did not die of a different cause. These numbers are summed over the workers who died of the cause of interest to obtain an estimate that addresses the impact of banning the fluids in a public health framework. The average among all workers who died of that cause and had nonzero exposure can also be calculated.

Aim 1 was purely methodological and therefore has no results. The methods, along with the application (Aim 2), are described in detail in a paper currently under review at the *American Journal of Epidemiology*.

Specific Aim 2: *Apply method from Aim 1 to evaluate the causal effect of quantitative exposure to oil-based metalworking fluids on cardiovascular disease mortality (ICD-9 codes 390-459) in the UAW-GM autoworkers cohort. Also evaluate the effects on mortality from ischemic heart disease (codes 410-414), acute myocardial infarction (code 410), and cerebrovascular disease/stroke (codes 430-438). For each outcome, calculate the median ratio of survival time if never exposed to observed survival time.*

Methodology. Aim 2 was an application of the method developed for Aim 1. We studied several mortality outcomes: all-cause mortality, IHD, AMI, and cerebrovascular disease. We decided not to use this method to analyze cardiovascular disease overall because previous work for Aim 3 indicated that the results for cardiovascular disease were mainly driven by the relationship between straight metalworking fluids and IHD. The effect metrics reported were the total and average number of years of life that could have been saved if the fluid had been banned, under the assumptions of model (1) (described above). The confounders adjusted for in this analysis were the covariates included in the model predicting exposure: sex, race, facility, calendar year, age, past exposures and time taken off work. For cerebrovascular disease and AMI, g-estimation using a continuous exposure variable in the accelerated failure time model did not converge, so we also ran all analyses using a categorical variable treated as continuous in the accelerated failure time model.

Results. The analysis using a continuous measure of exposure indicated that workers who died during follow-up and were ever exposed to straight metalworking fluids would have lived, on average, 2.08 years longer (95% CI: 0.43, 3.51) if they had never been exposed. Of course, some of these workers may have died from diseases unrelated to metalworking fluid exposure. Those who died of IHD would have lived an estimated average of 4.22 (-0.07, 28.14) years longer if they had never been exposed. Using the categorical measure of exposure, results were attenuated for all-cause mortality and for IHD and AMI, but we estimated that those who died of AMI or cerebrovascular disease would have lived 0.854 (0.025, 1.779), or 4.16 (2.70, 5.77) years longer, respectively, if they had never been exposed. Using categorical analysis for all outcomes yielded the following public health impacts of a ban (with 95% CIs): 1058 (144, 2002), 493 (-298, 1346), 805 (24, 1677), and 951 (619, 1321) years of life saved for all causes, IHD, AMI, and cerebrovascular disease respectively. The categorical analysis also suggested that the etiologic effect may be strongest for cerebrovascular disease other than hemorrhagic stroke, though hemorrhagic stroke also appeared to be strongly related to metalworking fluid exposure.

Specific Aim 3: *Investigate the impact of alternative exposure limits on time to death from cardiovascular disease, considering a series of plausible specific OSHA standards for oil-based metalworking fluids.*

Methodology.

Aim 3 used a related method within a public health framework that we developed. While still using g-estimation of an accelerated failure time model, for this set of analyses we dichotomized exposure into a series of binary variables, each indicating whether exposure exceeded a different “cutoff” level representing a hypothetical occupational exposure limit. This approach allowed us to evaluate the impacts of hypothetical interventions to enforce those limits in terms of the number of years of life saved for each of the following causes: all causes combined, all cardiovascular causes, IHD, AMI, and cerebrovascular disease. By comparing

the survival times under each limit to the observed survival times in each separate analysis, we were then able to compare *across* interventions using the estimates of the total numbers of years of life saved. Using multiple cutoff levels was also a way to take advantage of quantitative exposure information without using a continuous exposure variable or making strong assumptions about linearity.

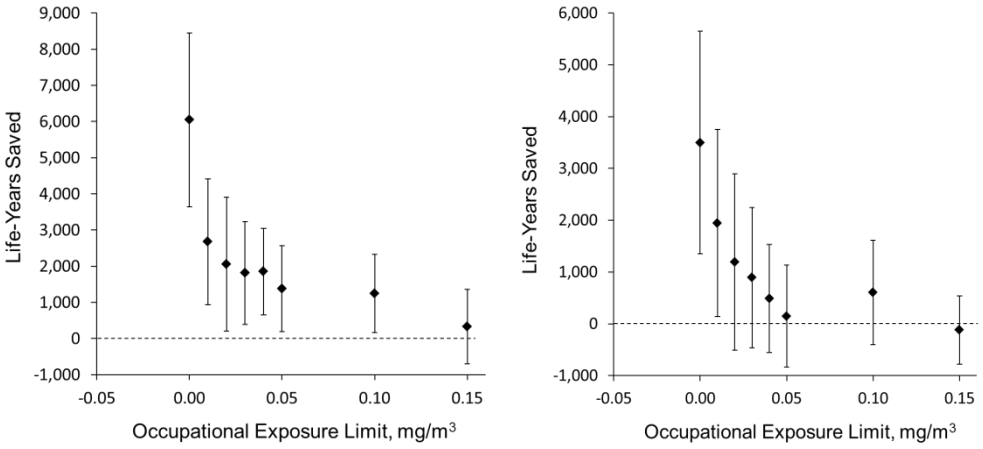
Cutoffs were selected based on the exposure distribution: closely spaced for the smaller values, where most of the data were concentrated, and with the highest cutoff near the 75th percentile of nonzero exposures. We estimated the total number of years of life that would have been saved among the cases if exposure to straight fluids had never been permitted to exceed limits of 0 (i.e., a ban), 0.01, 0.02, 0.03, 0.04, 0.05, 0.10, and 0.15mg/m3.

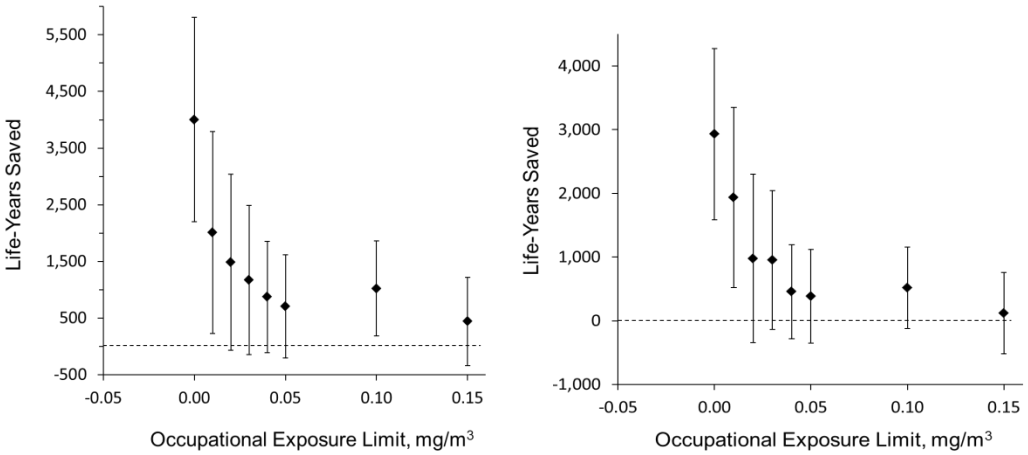
In these analyses, we modeled exposure using a separate logistic regression for each binary variable in order to adjust for confounding (variables included were identical to those used to predict exposure in Aim 2).

Results. If straight metalworking fluids had been banned, workers who were ever exposed and who died of cardiovascular disease would have lived an estimated average of 1.58 (95%CI: 0.63-2.52) years longer. On average, IHD deaths would have happened 2.77 (1.56-3.98) years later, and those who died of AMI would have lived 3.13 (1.74-4.53) years longer. Deaths from cerebrovascular disease among exposed workers would have occurred a mean of 3.19 (-0.14-6.51) years later.

Figure 2.1 shows the estimates for the population impacts on mortality for several causes under the different exposure limits. While a ban would have a reasonably strong impact on these outcomes, the magnitude of the effect quickly decreases for exposure limits only slightly above zero.

Figure 2.1. Estimated number of years of life that could have been saved by enforcing various occupational exposure limits for straight metalworking fluids, for all-cause mortality (upper left), cardiovascular mortality (upper right), IHD mortality (lower left), and AMI mortality (lower right). Note that these graphs have different scales on the y-axis.





The relationship appeared different for cerebrovascular mortality. (See Figure 1.1.) A ban was estimated not to provide substantially more benefit for this outcome than a limit of 0.02mg/m3, while a smaller impact was estimated for a limit of 0.03mg/m3. For higher exposure limits, the impacts on cerebrovascular mortality dropped down to null.

These analyses and the method, including the public health framework, are described in an article published the *American Journal of Epidemiology*.[5] The journal also solicited a commentary to accompany this paper.[6]

Specific Aim 4: *Compare results from Aim 2 with results from corresponding conventional exposure-response methods. Interpret the difference as the magnitude of healthy worker survivor bias in the conventional analyses.*

Methodology. We ran some conventional analyses using Cox models for Aim 4. Under certain assumptions, the g-estimation parameter may be converted into a hazard ratio. However, the hazard ratio obtained compares the hazard if everyone were always exposed at 1mg/m^3 to the hazard if everyone were never exposed. This is not a particularly meaningful comparison in an occupational cohort where exposure cannot occur after employment termination. Furthermore, the conventional exposure-response analysis yields a hazard ratio that corresponds to a 1mg/m^3 increase in cumulative exposure, so a quantitative comparison of results from the two methods would make little sense. Nevertheless, we have run a conventional analysis assuming a linear relationship, in order to make a qualitative comparison. For each outcome (all-cause mortality, IHD, AMI, and cerebrovascular disease), we adjusted for cumulative exposure to soluble metalworking fluids and synthetic metalworking fluids, age, race, sex, facility, and time off work. We also ran all of those analyses adjusting additionally for employment status.

Results. Conventional Cox regression estimated null or protective associations for all-cause mortality, ischemic heart disease, acute myocardial infarction, and cerebrovascular disease, whether or not employment status was included in the model (Table 1).

Table 1. Conventional regression results: Associations between a 1mg/m^3 increase in cumulative exposure to straight metalworking fluids and various mortality outcomes. All analyses are adjusted for cumulative exposures to soluble metalworking fluids and synthetic metalworking fluids, age, race, sex, facility, and time off work.

	Not adjusted for employment status		Adjusted for employment status	
Cause of death	**Hazard ratio**	**95% CI**	**Hazard ratio**	**95%CI**
All causes	0.993	(0.987, 0.999)	0.999	(0.993, 1.005)
Ischemic heart disease	1.001	(0.990, 1.011)	1.003	(0.992, 1.013)
Acute myocardial infarction	0.992	(0.977, 1.007)	0.996	(0.981, 1.011)
Cerebrovascular disease	0.999	(0.975, 1.024)	1.003	(0.979, 1.027)

These results have not been published because of the impossibility of making a quantitative comparison between the results from Aims 2 and 3 and these conventional regression results.

Discussion

Our results from Aim 4 demonstrate that healthy worker survivor bias prevented conventional regression from detecting a harmful effect of straight metalworking fluids on these outcomes at all. By contrast, using g-estimation of accelerated failure time models (as we did in Aims 2 and 3) to examine effects of either quantitative or binary exposure history on those outcomes showed that interventions to limit or ban these fluids would have saved substantial years of life.[5]

The apparent disparity in the results between some of the categorical analyses from Aim 2 and the analyses from Aim 3 is actually consistent with a non-linear effect of exposure. In Aim 3, we noted that exposure limits above zero had much smaller impacts on IHD and acute myocardial infarction mortality than a complete ban, whereas the benefits persisted at slightly higher limits for cerebrovascular disease mortality. In that analysis, exposure was classified as several binary variables determined by whether or not annual exposure exceeded a series of cutoff values. Increases in years of life saved were observed for IHD even when limits were reduced from very low levels of exposure down to zero—so that a total ban had a dramatically stronger impact. At the same time, lowering the limit the same amount from a higher level of exposure mattered less.[5] Together, these results suggest that the slope would be attenuated if we assumed that each increase into the next decile of exposure has the same effect, as we did in Aim 2. By contrast, for cerebrovascular disease, the analysis from Aim 3 suggested that exposure limits above zero could offer substantial protection even without banning the fluids entirely, so considering exposure category as continuous resulted in less attenuation.

From the etiological point of view, analysis of each binary variable in Aim 3 assumed that the level of exposure was irrelevant within each category (below or above the cutoff).[5] In Aim 2, either exposure or exposure category was instead treated as a continuous variable in the structural model. We were therefore making different assumptions about etiologically relevant levels of exposure: in the continuous exposure analysis, a one-unit change in a constant annual exposure was assumed to be equally important regardless of the level of exposure, while in the categorical exposure analysis, an increase to the next decile category was assumed to be equally important regardless of the level of exposure, and these assumptions may be incorrect. For example, workers who can already tolerate a given high level of exposure may be likely to tolerate an exposure 10% higher, whereas the same absolute increase would represent a proportionally much higher increase for a low exposure. Thus, we do not conclude from the Aim 2 categorical results that banning straight metalworking fluids would have had little or no effect on survival times for IHD. Rather, we believe that for IHD, and to a lesser extent for acute myocardial infarction, the damage caused by lower levels of exposure may be insufficiently captured by the categorical exposure metric and parametric model.[7]

For this cohort, the most effective way to incorporate quantitative exposure into a g-estimation analysis appeared to be to create a series of dichotomous exposure variables using cutoff values and then run a separate analysis for each one. For some outcomes, using a continuous exposure variable in the structural model was not possible, and for some, using a categorical exposure variable in the structural model led to severe attenuation in the results.

An important lesson from our work is that this method is well-suited to studying all-cause mortality and other non-rare outcomes, but that other methods are preferable when the outcome is rare. The more common the outcome, the more suitable the accelerated failure time model is.

Our work on this grant contributes to the occupational literature in two important methodological ways: we minimize healthy worker survivor bias and also approach the research question from a public health perspective. G-estimation provided better control for the healthy worker survivor effect than is generally possible using traditional analytic methods in occupational studies where follow-up continues past employment termination. By estimating the number of years of life that would have been saved by enforcing exposure limits, we obtain results with a concrete interpretation.

From the substantive perspective, our results provide evidence supporting the hypothesis of cardiovascular effects of exposure to aerosolized straight metalworking fluids. The evidence is particularly strong for ischemic heart disease and cerebrovascular mortality, but further research using different models is needed, especially for the latter outcome.

References

1. Picciotto S, Brown DM, Chevrier J, Eisen EA. Healthy worker survivor bias: implications of truncating follow-up at employment termination. Occup Environ Med 2013;70:736-42
2. Naimi AI, Cole SR, Hudgens MG, Richardson DB. Estimating the effect of cumulative occupational asbestos exposure on time to lung cancer mortality: using structural nested failure-time models to account for healthy-worker survivor bias. Epidemiology 2014;25:246-54.
3. 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.
4. 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.
5. Picciotto S, Peters A, Eisen EA. Hypothetical exposure limits for straight metalworking fluids and cardiovascular mortality in a cohort of autoworkers: structural accelerated failure time models in a public health framework. Am J Epidemiol 2015. doi: 10.1093/aje/kwu484
6. Naimi AI, Tchetgen EJT. Estimating population impact in the presence of competing risks. Am J Epidemiol 2015. doi:10.1093/aje/kwu486
7. Pope CA, 3rd, Burnett RT, Krewski D, et al. Cardiovascular mortality and exposure to airborne fine particulate matter and cigarette smoke: shape of the exposure-response relationship. Circulation 2009;120:941-8.

Publications

Picciotto S, Brown DM, Chevrier J, Eisen EA: [2013] Healthy worker survivor bias: implications of truncating follow-up at employment termination. Occupational and Environmental Medicine. 70(10):736-742

Addressing Aim 3:

Picciotto S, Peters A, Eisen EA: [2015] Hypothetical exposure limits for straight metalworking fluids and cardiovascular mortality in a cohort of autoworkers: structural accelerated failure time models in a public health framework. American Journal of Epidemiology; doi: 10.1093/aje/kwu484.

Addressing Aims 1 and 2:

Picciotto S, Ljungman PL, Eisen EA: Straight metalworking fluids and all-cause and cardiovascular mortality analyzed using g-estimation of an accelerated failure time model with quantitative exposure: methods and interpretations. Under review at American Journal of Epidemiology.

Inclusion of gender and minority study subjects

No new study subjects were recruited in this re-analysis of existing data. The workers included in this analysis were the 33,907 men and 4759 women (88% male), including 7144 black and 31,522 white workers (82% white), who had complete data. The numbers in the enrollment table below include those excluded from analysis due to missing data.

ENROLLMENT: Number of Subjects			
Ethnic Category	**Sex/Gender**		
	Females	**Males**	**Total**
Hispanic or Latino			
Not Hispanic or Latino	4680	41,717	46,397
Ethnic Category: Total of All Subjects *	4680	41,717	46,397
Racial Categories			
American Indian/Alaska Native			
Asian			
Native Hawaiian or Other Pacific Islander			
Black or African American	1265	6486	7751
White	3415	35,231	38,646
Racial Categories: Total of All Subjects *	4680	41,717	46,397

* The “Ethnic Category: Total of All Subjects” must be equal to the “Racial Categories: Total of All Subjects.”

Materials available for other investigators

Code for g-estimation will be made available on our website upon publication of the final paper.