American Journal of Epidemiology

Vol. 183, No. 7
DOI: 10.1093/aje/kwv232
Advance Access publication:
March 10, 2016

Practice of Epidemiology

Straight Metalworking Fluids and All-Cause and Cardiovascular Mortality Analyzed by Using G-Estimation of an Accelerated Failure Time Model With Quantitative Exposure: Methods and Interpretations

Sally Picciotto*, Petter L. Ljungman, and Ellen A. Eisen

* Correspondence to Dr. Sally Picciotto, Environmental Health Sciences Division, School of Public Health, University of California, Berkeley, 50 University Hall # 7360, Berkeley, CA 94720-7360 (e-mail: sallypicciotto@berkeley.edu).

Initially submitted February 13, 2015; accepted for publication August 25, 2015.

Straight metalworking fluids have been linked to cardiovascular mortality in analyses using binary exposure metrics, accounting for healthy worker survivor bias by using g-estimation of accelerated failure time models. A cohort of 38,666 Michigan autoworkers was followed (1941–1994) for mortality from all causes and ischemic heart disease. The structural model chosen here, using continuous exposure, assumes that increasing exposure from 0 to 1 mg/m^3 in any single year would decrease survival time by a fixed amount. Under that assumption, banning the fluids would have saved an estimated total of 8,468 (slope-based 95% confidence interval: 2,262, 28,563) person-years of life in this cohort. On average, 3.04 (slope-based 95% confidence interval: 0.02, 25.98) years of life could have been saved for each exposed worker who died from ischemic heart disease. Estimates were sensitive to both model specification for predicting exposure (multinomial or logistic regression) and characterization of exposure as binary or continuous in the structural model. Our results provide evidence supporting the hypothesis of a detrimental relationship between straight metalworking fluids and mortality, particularly from ischemic heart disease, as well as an instructive example of the challenges in obtaining and interpreting results from accelerated failure time models using a continuous exposure in the presence of competing risks.

cardiovascular outcomes; epidemiologic methods; healthy worker effect; mortality; occupational exposures; particulate matter

Abbreviations: CI, confidence interval; IHD, ischemic heart disease; $PM_{3.5}$, particulate matter of aerodynamic diameter ≤3.5 μm.

A large literature is devoted to the relationships between exposures to particulate matter from various sources, especially fossil fuel combustion, and cardiovascular outcomes (1–14). In addition to consistent associations with cardiovascular mortality (15), particulate matter is associated with hospitalizations for myocardial infarction (16), arrhythmias (17), stroke (18), and heart failure (19). Results from animal studies as well as observational studies investigating associations with intermediate endpoints such as decreased heart-rate variability, markers of inflammation, intima media thickness, and blood pressure indicate that particulate matter may affect both inflammatory/oxidative stress pathways and autonomic function (5).

Similar pathways may be responsible for cardiovascular outcomes among machinists: Particulate matter exposures at work are often at least an order of magnitude higher than ambient air pollution levels in the United States (20, 21). Made from petroleum, straight metalworking fluids are used in machining processes to reduce friction. During grinding, cutting, and drilling, the fluids heat up and aerosolize, creating particles that may include metals and polycyclic aromatic hydrocarbons and that thus share some properties with particulate matter produced by fossil fuel combustion (12, 22–25).

As with other occupational exposures, a true detrimental effect of metalworking fluids on cardiovascular outcomes may be difficult to detect because of healthy worker survivor bias. This "downward" bias (toward and sometimes across the null) arises from both an ongoing selection process and time-varying confounding affected by prior exposure, because the workers who are least susceptible to health effects accumulate the most

exposure (26–29). The bias may be avoided or reduced by analyzing the data using statistical methods designed to address this type of data structure, such as any of the g-methods developed by Robins et al. (30–36). (One of those methods, inverse probability of treatment weighting, is invalid when follow-up continues past employment termination. Only actively employed workers can be exposed, violating the required positivity assumption (35).)

Very few occupational studies have examined metalworking fluids and cardiovascular outcomes (12, 21, 37–39). A previous analysis of the autoworkers' cohort, restricted to actively employed person-time, found modest evidence for an increased risk of ischemic heart disease (IHD) with exposure to straight fluids. However, that study used traditional regression methods (12), so its results are subject to healthy worker survivor bias. Evidence using g-estimation to study these relationships has been reasonably consistent, but has relied on binary exposure variables (12, 21, 38, 39). Given the extensive exposure assessment available for this cohort (40, 41), dichotomizing the exposure represents an unsatisfying loss of data that could make a causal relationship more difficult to detect and quantify. Our most recent analysis attempted to bridge this gap partially, by running several analyses using separate binary exposure variables based on different cutoffs (21). Although the estimate for each cutoff represented the potential benefit of enforcing an exposure limit at that cutoff, the pattern over several cutoffs was difficult to interpret.

In this paper, we present analyses using g-estimation of an accelerated failure time model to assess the potential relationship of workers' histories of quantitative exposure to straight metalworking fluids with times to death from IHD and all-cause mortality.

METHODS

Cohort

The United Autoworkers–General Motors cohort has been described in detail before. Briefly, it consists of all workers at 3 plants in Michigan who were hired between 1938 and 1982 and worked for at least 3 years. Follow-up for mortality continued through 1994 (40). This analysis followed 38,666 workers from 3 years after hire to death, age 95, or the end of 1994. The outcomes studied (in separate analyses) were as follows: mortality from any cause, IHD (*International Classification of Diseases, Ninth Edition,* codes 410–414), acute myocardial infarction (code 410), or cerebrovascular disease (codes 430–438).

Job histories, including time off work, were available from company records, as were air sampling measures of total particulate matter collected over decades. These historical particulate matter data were combined with size fractions of particulate matter collected in the mid-1980s by research industrial hygienists. The particulate matter measurements were then classified as straight (oil-based), soluble, or synthetic metalworking fluid based on material safety data records for each machining and grinding operation. The resulting time-dependent job-exposure matrix was combined with job records to estimate annual averages of daily exposures to each fluid type in each calendar year for each worker (40, 41). This analysis treats complete worker history of annual average daily concentration of particulate matter of aerodynamic diameter ≤ 3.5 μm ($PM_{3.5}$) (mg/m^3) composed of straight metalworking fluids as the exposure of interest, with the other fluid types considered as confounders.

Statistical approach

Our method is very similar to that introduced by Naimi et al. (42) in a study of cumulative asbestos exposure and time to lung cancer death. However, we consider annual average daily exposure (rather than cumulative exposure) in each year and report our estimates using different metrics.

The structural accelerated failure time model considers how each year's level of exposure relates to survival time by examining the relationship between observed and counterfactual quantities. Survival time if never exposed (a counterfactual variable that is unobserved for workers who were ever exposed) is modeled as a function of the observed survival time and observed annual average daily exposure levels, using an unknown coefficient ψ:

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

where T is observed survival time, $A(t)$ is observed exposure at time t measured in mg/m^3, $T_{\bar{0}}$ is counterfactual survival time if never exposed, and ψ is the unknown coefficient to be estimated (43). This ψ could be interpreted as the log of the ratio comparing the median survival time if everyone had never been exposed with the median survival time if everyone had been exposed to 1 unit of exposure (in this case, an average daily concentration of 1 mg/m^3) during every year of follow-up.

Because our worker population can be exposed only while actively employed and follow-up extends beyond termination of work, most workers cannot experience exposure in every year of follow-up. This etiological metric for the effect measure thus requires us to extrapolate outside our data. We therefore report the cause-specific total and average number of person-years of life lost due to exposure to metalworking fluids (equivalently, the total and the average number of person-years of life that could have been saved by banning the fluid), among those who died from the cause of interest and were ever exposed (refer to the Web Appendix available at http://aje.oxfordjournals.org/). Together, these metrics help to locate our estimate within frameworks that answer both a public health question (total) and a more etiological question (average).

Note, however, that neither the total nor the average number of years of life lost is based on a comparison that is purely etiological. Both also depend on the observed distribution of exposure among the cases, because we compare the observed survival times with what would have happened if nobody had been exposed. Thus, given a certain magnitude of etiological effect, the total number of person-years of life saved under an exposure ban is greater if more workers were exposed at all (or at higher levels, or for longer durations), and the average number of years of life saved is greater if the exposed workers had higher levels or longer durations of exposure.

Furthermore, the shape of the exposure-response relationship is difficult to conceptualize on the familiar scales (e.g.,

linear, log-linear, and so on) that are often used to describe the relationship of exposure with risk of disease, because 1) survival times are different from probabilities, and 2) we do not reduce a worker's time-varying exposure to a single summary measure like cumulative exposure but instead consider his or her entire history of exposure. For example, consider 2 workers with 5 years of employment. Worker 1 experiences 5 units of exposure in year 2 (and no exposure at any other time), and Worker 2 is exposed to 1 unit in each of years 1–5 (and no exposure at any other time). Workers 1 and 2 share the same cumulative exposure but different exposure histories. Using exposure history rather than cumulative exposure permits us to avoid conflating duration and intensity of exposure.

In order to adjust for confounding in a traditional regression, measured covariates are included as predictors in the model for the outcome, so that the coefficient of exposure is conditional on confounders. However, in g-estimation, the model for the outcome is "structural" and does not depend on confounders. Adjustment for confounding is instead achieved by modeling the exposure in each year on the basis of prior values of covariates (43–45). Also included in the exposure model expressly for the purpose of g-estimation are time from cohort entry to administrative end of follow-up (i.e., the maximum follow-up time possible for each worker; this is known at hire) and an inverse-probability-of-censoring–weighted indicator variable for whether or not the worker's death would have occurred before administrative end of follow-up under all possible exposure histories. This variable is a function of counterfactual survival time; thus, under the assumption that there are no unmeasured confounders, it should be statistically independent of exposure, conditional on measured covariates. The g-estimation procedure is therefore the search for the value of the unknown coefficient ψ that produces counterfactual survival times whose β is 0 in the model predicting exposure (or equivalently, whose P value is 1) (43, 45).

For this analysis, we used a multinomial logistic regression to predict level of annual average daily $PM_{3.5}$ exposure in 11 categories: unexposed in that year and categories corresponding to deciles of exposure (42). We chose this model because exposure is determined by job, which is inherently more of a qualitative category than a ranking by exposure level (refer to the Web Appendix). The covariates adjusted for in this model were age, race (black or white), sex, plant, an indicator variable for calendar year being before/after 1970 (when exposure levels dramatically decreased at the 3 plants), previous time off work, and exposures to all 3 types of metalworking fluids in previous years. A sensitivity analysis used logistic regression to predict exposure being nonzero in each year, using the same predictors. In both cases, the analysis controls for employment status by running the exposure model only on actively employed person-time (46), because exposure is necessarily 0 when a worker is not employed.

Censoring by death from causes other than the one of interest (i.e., competing risks) may cause bias if unmeasured shared risk factors for the cause of interest and the other causes exist. For example, (unmeasured) risk factors are the same for acute myocardial infarction and IHD because they have the same root cause, atherosclerosis (47). Thus, in the analysis examining the relationship between metalworking fluids and acute myocardial infarction mortality, censoring

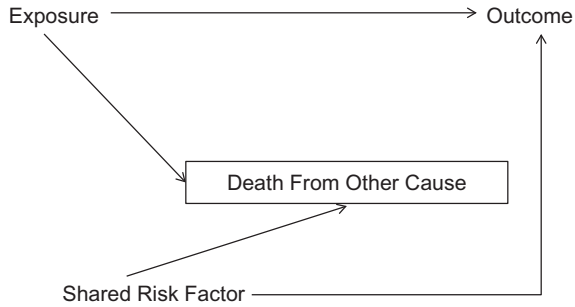

Figure 1. This diagram shows an exposure, an outcome, a competing risk, and an unmeasured shared risk factor of the outcome and the competing risk. Because the outcome cannot be observed in those who die from another cause, we necessarily restrict our analysis to those who were not censored by death (this restriction is represented by a box around "death from other cause"). This means that an unblocked, noncausal path exists connecting the exposure to the outcome through censoring (which is a collider) and the unmeasured shared risk factor.

workers who died from other IHD means that we may introduce bias in the relationship between the exposure and acute myocardial infarction mortality (Figure 1). To avoid this bias, we estimated the relationship in a pseudo population in which death from other causes does not occur. For each outcome, we used inverse probability weighting to adjust for censoring by all other causes of death (17% of workers in the analysis of IHD, 20% for acute myocardial infarction, and 23% for cerebrovascular disease) and by loss to follow-up (<4%) (21, 43). A sensitivity analysis without weights did not adjust for competing

Table 1. Demographic and Employment Characteristics of the United Autoworkers–General Motors Cohort, Michigan, 1941–1994

Characteristic	No.	%	Person-Years
Total	38,666	100	972,476
Race			
Black	7,144	18	169,065
White	31,522	82	803,411
Sex			
Male	33,907	88	869,585
Female	4,759	12	102,891
Plant			
1		24	151,901
2		40	248,722
3		36	224,545
Active employment years with no time taken off		72	447,244
Ischemic heart disease deaths	2,612	7	
Acute myocardial infarction deaths	1,699	4	
Cerebrovascular disease deaths	501	1	
All-cause mortality	9,539	25	

Table 2. Further Demographic and Employment Characteristics of the United Autoworkers–General Motors Cohort of 38,666 Workers, Michigan, 1941–1994

Characteristic	Mean (SD)
Age at baseline, years	30.8 (9.1)
Years worked	16.2 (9.5)
Length of follow-up, years	24.5 (11.2)
Age at death from ischemic heart disease, years	64.4 (11.7)
Age at death from acute myocardial infarction, years	62.7 (11.6)
Age at death from cerebrovascular disease, years	67.0 (12.9)
Age at death from any cause, years	62.5 (13.5)
Proportion of year taken off (among actively employed person-years)	0.097 (0.244)
Proportion of year taken off if >0 (among actively employed person-years)	0.341 (0.354)

Abbreviation: SD, standard deviation.

risks and instead assumed either that the other causes of death did not share risk factors with the outcome of interest or that exposure did not cause mortality other than for the outcome of interest.

Confidence intervals were obtained 2 ways. The estimation process was computationally intensive, and our data set was large, rendering bootstrapping impractical. We therefore computed a conservative Wald test-based 95% confidence interval by searching for the values of ψ that produced P values as close as possible to 0.05 without exceeding 0.05. We also computed a symmetrical slope-based confidence interval (42). A linear regression of the z statistic (β divided by its standard error) using ψ as the only predictor yielded a slope m; the estimated standard error of the estimate of ψ is $|m^{-1}|$ (44).

For further details on g-estimation and the models, refer to the Web Appendix and Web Table 1. All analyses were carried out in SAS, version 9.4, software (SAS Institute, Inc., Cary, North Carolina). The study was approved by the University of California, Berkeley, Committee for the Protection of Human Subjects.

RESULTS

Tables 1 and 2 describe the cohort. Table 3 provides information about exposure.

For acute myocardial infarction and cerebrovascular mortality, the test statistic crossed 0 more than once within the search interval of plausible values for ψ, so that a unique

Table 3. Distribution of Exposure to Straight Metalworking Fluids Among Ever-Exposed Workers in the United Autoworkers–General Motors Cohort (n = 38,666), Michigan, 1941–1994[a]

	No. of Workers	Person-Years	%	Mean (SD)
Workers ever exposed	20,202		52	
Exposed		157,512	16	
Exposure category, among active employment time				
Unexposed (reference category for exposure model)		467,938	75	
Exposure >0 and ≤0.00808 mg/m^3 (category 1 for exposure model)		15,719	3	
Exposure >0.00808 and ≤0.00998 mg/m^3 (category 2 for exposure model)		17,574	3	
Exposure >0.00998 and ≤0.01876 mg/m^3 (category 3 for exposure model)		13,890	2	
Exposure >0.01876 and ≤0.02247 mg/m^3 (category 4 for exposure model)		22,126	4	
Exposure >0.02247 and ≤0.03110 mg/m^3 (category 5 for exposure model)		9,416	2	
Exposure >0.03110 and ≤0.05649 mg/m^3 (category 6 for exposure model)		15,640	3	
Exposure >0.05649 and ≤0.10365 mg/m^3 (category 7 for exposure model)		15,692	3	
Exposure >0.10365 and ≤0.16200 mg/m^3 (category 8 for exposure model)		15,910	3	
Exposure >0.16200 and ≤0.38485 mg/m^3 (category 9 for exposure model)		15,635	3	
Exposure >0.38485 mg/m^3 (maximum = 2.98560) (category 10 for exposure model)		15,628	2	
Exposure duration, years				4.1 (6.6)
Exposure duration of ever-exposed workers, years				
Among those who died from ischemic heart disease				4.5 (7.1)
Among those who died from acute myocardial infarction				4.4 (7.0)
Among those who died from cerebrovascular disease				4.8 (7.1)
Among those who died from any cause				4.2 (6.8)
Annual average daily exposure among exposed person-years, mg/m^3				0.2 (0.4)
Annual average daily exposure among exposed person-years, mg/m^3			0.03 (0.01–0.12)[b]	

Abbreviation: SD, standard deviation.
[a] The total number of person-years was 972,476.
[b] Median and interquartile range (25th–75th percentiles).

estimate could not be determined. Results are thus presented only for IHD and all-cause mortality.

The etiological relationship is expressed as the ratio of median survival times if always exposed to straight metalworking fluids at a $PM_{3.5}$ concentration of 1 mg/m^3 versus never exposed. (This level of exposure is very high in this population, representing approximately the 97th percentile among the person-time observations with nonzero exposure; the 2006 Environmental Protection Agency daily standard for ambient concentrations of particulate matter of aerodynamic diameter ≤2.5 μm ($PM_{2.5}$) was 0.035 mg/m^3 (48).) The ratio is less than 1 if the exposure is harmful. For IHD mortality, this ratio was 0.41 (slope-based 95% confidence interval (CI): 0.17, 0.99; test-based 95% CI: 0.17, 1.03); for all-cause mortality it was 0.50 (slope-based 95% CI: 0.31, 0.80; test-based 95% CI: 0.40, 0.77).

We estimated that banning straight metalworking fluids could have saved a total of 8,468 (slope-based 95% CI: 2,262, 28,563; test-based 95% CI: 2,628, 16,380) person-years of life among those who died during follow-up and 4,334 (slope-based 95% CI: 23, 37,953; test-based 95% CI: −88, 34,322) person-years of life among those who died from IHD during follow-up. For a hypothetical ban on straight metalworking fluids, Figure 2 shows the average number of years of life that would have been saved per worker who died during follow-up *and* had ever been exposed for these 2 outcomes. For comparison, results from the sensitivity analyses of the same structural model but using a logistic model to predict probability of nonzero exposure, as well as results from our previous analyses using a binary exposure (21), are also presented in this figure.

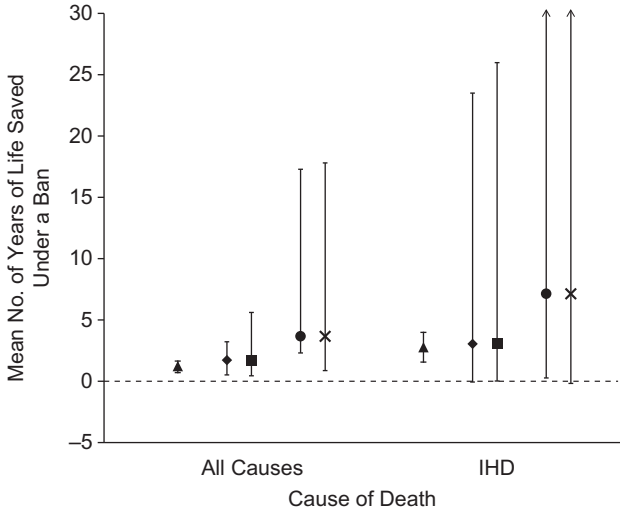

Figure 2. Estimated average number of years of life saved per exposed worker for all-cause mortality and ischemic heart disease (IHD) under a hypothetical ban on exposure to straight metalworking fluids in the United Autoworkers–General Motors cohort ($n = 38,666$), Michigan, 1941–1994. Results are shown for analyses using, from left to right for each outcome, a binary exposure variable in the structural model for the outcome and a logistic model predicting probability of nonzero exposure (with confidence intervals calculated by using bootstraps (triangles)) (21); a continuous exposure variable in the structural model for the outcome and a multinomial logistic model predicting exposure category (with confidence intervals computed by a Wald test (diamonds) and using the slope (squares)); and a continuous exposure variable in the structural model for the outcome and a logistic model predicting probability of nonzero exposure (with confidence intervals computed by a Wald test (circles) and using the slope (X's)). For IHD, the upper limits for the confidence intervals for the circle point and the X point are 280 and 262, respectively.

The sensitivity analyses using the pooled logistic (rather than multinomial logistic) model to predict probability of nonzero exposure produced estimates of ψ that were about 45% higher for both IHD and all-cause mortality, with much wider confidence intervals. For IHD, the sensitivity analysis with no weights for competing risks resulted in a similar estimate that was slightly farther from the null. None of the sensitivity analyses produced estimates for acute myocardial infarction or cerebrovascular mortality.

DISCUSSION

In a cohort of 38,666 autoworkers, we investigated the potential benefits of banning straight metalworking fluids on survival times for all-cause mortality and IHD. G-estimation yielded estimates for the number of years of life saved as a function of quantitative exposure history to $PM_{3.5}$ composed of metalworking fluids while minimizing healthy worker bias. Our analyses suggested that nearly 8,500 years of life could have been saved in this cohort by banning straight metalworking fluids, and that exposed workers who died from IHD would have lived, on average, 3 years longer if they had never been exposed to straight metalworking fluids.

Note that the results from these analyses are not easily compared with those from traditional regression analyses of cumulative exposure. 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 composite effect of the cumulative exposure. This disentanglement of exposure over time is how a structural nested model avoids the bias that occurs when stratifying on a time-varying confounder that is affected by prior exposure.

As was noted above, particles of aerosolized straight metalworking fluids contain polycyclic aromatic hydrocarbons and metals and are derived from petroleum; they thus share some properties with particulate matter from fossil fuel combustion (12, 22–24). Because ambient traffic-related air pollution has been strongly linked to IHD, acute myocardial infarction, and stroke, as well as to intermediate outcomes such as inflammation, occupational exposure to metalworking fluids may cause these outcomes through similar mechanisms.

However, interpreting the results of these analyses causally would require a number of assumptions, including conditional exchangeability, consistency, and correct model specification, most of which are discussed at length in our previous paper (21). Below, we draw attention to the specific assumptions that differ between the approaches used for binary versus quantitative exposure.

Compared with results observed in a previous analysis in which the exposure was treated as a binary indicator of whether or not a worker was exposed at all during that year (21), results from this analysis using continuous exposure suggest that the public health benefits for IHD and all-cause mortality

from banning straight metalworking fluids are about 10% and 40% greater, respectively. Furthermore, the previous analysis produced estimates for all 4 outcomes, while the present analysis was unsuccessful for 2. Possible explanations for these disparities in results include any of the 3 ways in which the present analysis differs from the previous one: 1) The exposure level was predicted in 11 categories by using a multinomial logistic regression rather than in 2 categories by using a logistic regression; 2) the g-estimation algorithm itself was different; 3) exposure was treated as a continuous (rather than binary) variable in the structural model for the outcome. We explore these 3 below.

Like the main analyses, the sensitivity analyses in which we predicted exposure using a logistic (rather than multinomial logistic) regression produced no estimates for acute myocardial infarction and cerebrovascular mortality. They also produced considerably stronger results for IHD and all-cause mortality. The results from the main analyses that use multinomial logistic regression to predict exposure fall between the results from the 2 analyses using logistic regression to predict probability of nonzero exposure (one using binary exposure in the structural model; one using continuous), as seen in Figure 2. Thus, the change in model type for predicting exposure cannot explain the disparity in estimates between the 2 methods. Furthermore, confidence intervals were much wider in the sensitivity analysis. This, as well as the stronger estimates in the sensitivity analysis, may be at least partially due to uncontrolled confounding. In an analysis where we assume that exposure *intensity* affects the outcome, ensuring that counterfactual survival time is unrelated to the *probability* of exposure (as the logistic regression does) may be insufficient.

The estimating procedure itself differed for the 2 models. When exposure was binary, the logistic model predicting exposure was run only once, and its residuals were used to calculate the estimating function directly (21). With the other approach, we instead ran the multinomial logistic model to predict exposure separately for each candidate value of ψ to find the one with coefficient $\hat{\beta} = 0$. However, the sensitivity analyses with continuous exposure in the structural model and a logistic model predicting exposure probability were performed using *both* of these g-estimation procedures, and these produced identical estimates.

We therefore conclude that the stronger results observed for IHD and all-cause mortality using quantitative rather than binary exposure in the structural model are probably due to the different assumptions required by the 2 models. The binary analysis treats all levels of exposure above 0 as equivalent (21). In our main analysis, exposure (in mg/m^3) is instead treated as a continuous variable in the structural model. We are therefore making a different assumption about etiologically relevant levels of exposure: A 1-unit increase in a constant annual exposure is assumed to be equally important regardless of the level of exposure, and this may be biologically 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 caution readers not to conclude that the stronger results obtained in this analysis represent a conflict with the results from our earlier analysis. The truth may be somewhere in between, or the strong assumption of linearity in the present analysis may have exaggerated the strength of the relationship.

We assume consistency, that is, that each worker's counterfactual survival time under his or her observed exposure is equal to that worker's observed survival time; this assumption holds if the intervention on the exposure is well-defined. Here, the company cannot realistically ban the use of *all* metalworking fluids, because metal cutting requires lubrication, and high-speed grinding requires cooling. However, it could eliminate exposure to straight metalworking fluids by enforcing the use of 100% effective protective equipment (if it exists) or by replacing the oil-based straight fluids with soluble or synthetic fluids. (Indeed, straight fluids have now been mostly phased out in this industry (49).) The analysis assumes that health effects of these interventions are equivalent if they result in everyone being unexposed to straight fluids.

Estimation of population impact in the presence of competing events is complex. We adjusted for censoring by competing risks using inverse probability weighting. One way to conceptualize this is that we are imagining an intervention in which we prevent both exposure and death from other causes; preventing such deaths is impossible and therefore not a well-defined intervention. Thus, counterfactual consistency may be violated (50–53); this may help to account for the failure of this model to produce estimates for the 2 outcomes with the most competing events.

IHD is a fairly common outcome with many risk factors that were not intervened on, so the bias due to violating the assumption that these workers would have died from IHD regardless of their exposure level may not be too severe. All-cause mortality is not subject to competing risks, so results for that outcome are unaffected by this issue. Another advantage of studying all-cause mortality in a public health framework is that, for interventions on workplace exposures, the total number of person-years of life saved may be more relevant than those from only 1 particular cause.

For cause-specific mortality, alternative methods for analyzing the data can avoid this problem. One option is to use a different set of censoring weights to calculate the survivor average causal effect, which considers a subgroup consisting of individuals who would not have died during follow-up from any competing cause under any possible exposure history (52, 54, 55). However, this method yields results that are also difficult to interpret, requires some assumptions that may be problematic, and is impracticable for this particular analysis because of the extensive bootstrapping required (52, 54, 55). Other options include using a different analytical method, such as the parametric g-formula (56, 57) or targeted maximum likelihood estimation (58). Future work will compare these results with estimates from those methods. Refer to the Web Appendix for more on censoring.

One challenge of g-estimation of an accelerated failure time model seems to be how it handles a continuous exposure. A study of different doses of a pharmaceutical intervention for subjects on hemodialysis previously pointed out difficulties with g-estimation because of artificial censoring (59). Our experience in this analysis, compared with the previous one in the same cohort (21), suggests that g-estimation is more successful for a binary exposure variable. The instability of the test statistic, which led to failure to find a unique estimate for 2 outcomes,

seems to be more of an issue when using a continuous exposure variable in the structural model. Thus, in many cases, g-estimation of an accelerated failure time model studying a quantitative exposure may be better achieved by considering separate analyses incorporating a series of binary exposure variables, each indicating whether exposure does or does not exceed a different cutoff level (21, 60).

To our knowledge, only 1 prior occupational study has used g-estimation of an accelerated failure time model with a quantitative exposure metric (42). Its use of history of cumulative exposure rather than history of annual exposure assumes that the effect of exposure in a given year persists over the course of subsequent follow-up rather than decreasing survival time by a fixed amount (Web Appendix and Web Table 2), so that exposures earlier in a long follow-up have greater impact than those occurring later. Results in that paper were interpreted as the ratio of survival times corresponding to a 1-unit increase in cumulative exposure; however, the timing of that increase in cumulative exposure matters. We avoid those issues and assumptions by comparing counterfactual unexposed survival times with observed ones and by considering history of annual exposure.

This study is the first to use g-methods to consider the relationship between a quantitative measure of exposure to metalworking fluids and survival time for all-cause and cardiovascular mortality. It provides evidence supporting the hypothesis of a detrimental relationship between $PM_{3.5}$ composed of straight metalworking fluids and mortality, particularly from IHD, as well as an instructive example of the challenges in obtaining and interpreting results from accelerated failure time models using a continuous exposure in the presence of competing risks.

ACKNOWLEDGMENTS

Author affiliations: Environmental Health Sciences Division, School of Public Health, University of California, Berkeley, Berkeley, California (Sally Picciotto, Ellen A. Eisen); Unit of Environmental Epidemiology, Institute of Environmental Medicine, Karolinska Institute, Stockholm, Sweden (Petter L. Ljungman); Unit of Environmental Health, Center for Occupational and Environmental Health, Stockholm City Council, Stockholm, Sweden (Petter L. Ljungman); Department of Cardiology, Danderyds Hospital, Stockholm, Sweden (Petter L. Ljungman); and Cardiovascular Epidemiology Research Unit, Department of Cardiology, Beth Israel Deaconess Medical Center, Boston, Massachusetts (Petter L. Ljungman).

This work was financially supported by the National Institute for Occupational Safety and Health at the Centers for Disease Control and Prevention (grant R03 OH010202).

Conflict of interest: none declared.

REFERENCES

1. Burnett RT, Pope CA 3rd, Ezzati M, et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ Health Perspect*. 2014;122(4):397–403.
2. Pope CA 3rd, Burnett RT, Turner MC, et al. Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure-response relationships. *Environ Health Perspect*. 2011;119(11):1616–1621.
3. Peters A. Particulate matter and heart disease: evidence from epidemiological studies. *Toxicol Appl Pharmacol*. 2005;207(2 suppl):477–482.
4. Peters A. Ambient particulate matter and the risk for cardiovascular disease. *Prog Cardiovasc Dis*. 2011;53(5):327–333.
5. Brook RD, Rajagopalan S, Pope CA 3rd, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation*. 2010;121(21):2331–2378.
6. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*. 2006;295(10):1127–1134.
7. Fang SC, Cassidy A, Christiani DC. A systematic review of occupational exposure to particulate matter and cardiovascular disease. *Int J Environ Res Public Health*. 2010;7(4):1773–1806.
8. Faught BE, Flouris AD, Cairney J. Epidemiological evidence associating secondhand smoke exposure with cardiovascular disease. *Inflamm Allergy Drug Targets*. 2009;8(5):321–327.
9. Gold DR, Mittleman MA. New insights into pollution and the cardiovascular system: 2010 to 2012. *Circulation*. 2013;127(18):1903–1913.
10. McGill HC Jr. The cardiovascular pathology of smoking. *Am Heart J*. 1988;115(1 pt 2):250–257.
11. Costello S, Brown DM, Noth EM, et al. Incident ischemic heart disease and recent occupational exposure to particulate matter in an aluminum cohort. *J Expo Sci Environ Epidemiol*. 2014;24(1):82–88.
12. Costello S, Garcia E, Hammond SK, et al. Ischemic heart disease mortality and $PM_{3.5}$ in a cohort of autoworkers. *Am J Ind Med*. 2013;56(3):317–325.
13. Neophytou AM, Costello S, Brown DM, et al. Marginal structural models in occupational epidemiology: application in a study of ischemic heart disease incidence and $PM_{2.5}$ in the US aluminum industry. *Am J Epidemiol*. 2014;180(6):608–615.
14. O'Donnell MJ, Fang J, Mittleman MA, et al. Fine particulate air pollution $PM_{2.5}$ and the risk of acute ischemic stroke. *Epidemiology*. 2011;22(3):422–431.
15. Lepeule J, Laden F, Dockery D, et al. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities Study from 1974 to 2009. *Environ Health Perspect*. 2012;120(7):965–970.
16. Peters A, von Klot S, Heier M, et al. Particulate air pollution and nonfatal cardiac events. Part I. Air pollution, personal activities, and onset of myocardial infarction in a case-crossover study. *Res Rep Health Eff Inst*. 2005;(124):1–66; discussion 67–82, 141–148.
17. Ljungman PL, Berglind N, Holmgren C, et al. Rapid effects of air pollution on ventricular arrhythmias. *Eur Heart J*. 2008;29(23):2894–2901.
18. Ljungman PL, Mittleman MA. Ambient air pollution and stroke. *Stroke*. 2014;45(12):3734–3741.
19. Shah AS, Langrish JP, Nair H, et al. Global association of air pollution and heart failure: a systematic review and meta-analysis. *Lancet*. 2013;382(9897):1039–1048.

20. Cullen MR. Invited commentary: The search for preventable causes of cardiovascular disease—whither work? *Am J Epidemiol*. 2009;169(12):1422–1425.
21. Picciotto S, Peters A, Eisen EA. Hypothetical exposure limits for oil-based metalworking fluids and cardiovascular mortality in a cohort of autoworkers: structural accelerated failure time models in a public health framework. *Am J Epidemiol*. 2015; 181(8):563–570.
22. World Health Organization. *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*. Vol 33. Polynuclear aromatic hydrocarbons, part 2, carbon blacks, mineral oils, and some nitroarenes. Lyon, France: IARC; 1984.
23. Stebbins A. Cobalt exposures enhanced by synthetic coolants. *Environ Health News*. 1989;3:3–7.
24. Oxhøj H, Andreasen H, Henius UM. Respiratory symptoms and ventilatory lung function in machine shop workers exposed to coolant-lubricants. *Eur J Respir Dis Suppl*. 1982;118:85–89.
25. Mosher E, Peterson L, Skold R. The chemical control of cobalt leaching from cemented carbide tooling. *Mater Perform*. 1986; 25(10):38–43.
26. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology*. 1994;5(2): 189–196.
27. Eisen EA, Picciotto S, Robins JM. Healthy worker effect in occupational studies. In: El-Shaarawi AH, Piegorsch WW, eds. *Encyclopedia of Environmetrics, Second Edition*. Vol 2. Chichester, United Kingdom: John Wiley & Sons; 2012: 1269–1272.
28. Buckley JP, Keil AP, McGrath LJ, et al. Evolving methods for inference in the presence of healthy worker survivor bias. *Epidemiology*. 2015;26(2):204–212.
29. Picciotto S, Hertz-Picciotto I. Commentary: Healthy worker survivor bias: a still-evolving concept. *Epidemiology*. 2015; 26(2):213–215.
30. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Modell*. 1986;7(9-12):1393–1512.
31. Robins J. Estimation of the time-dependent accelerated failure time model in the presence of confounding factors. *Biometrika*. 1992;79(2):321–334.
32. Robins J, Tsiatis AA. Semiparametric estimation of an accelerated failure time model with time-dependent covariates. *Biometrika*. 1992;79(2):311–319.
33. Robins JM. Structural nested failure time models. In: Andersen PK, Keiding N, eds. *Survival Analysis*. Chichester, United Kingdom: John Wiley & Sons; 1998:4372–4389.
34. Robins JM, Blevins D, Ritter G, et al. G-estimation of the effect of prophylaxis therapy for *Pneumocystis carinii* pneumonia on the survival of AIDS patients. *Epidemiology*. 1992;3(4): 319–336.
35. Robins JM, Hernán MA. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, et al, eds. *Longitudinal Data Analysis*. New York, NY: Chapman & Hall/CRC; 2009:553–599.
36. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550–560.
37. Park RM. Mortality at an automotive engine foundry and machining complex. *J Occup Environ Med*. 2001;43(5): 483–493.
38. 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(2):212–219.
39. Costello S, Picciotto S, Rehkopf DH, et al. Social disparities in heart disease risk and survivor bias among autoworkers: an examination based on survival models and g-estimation. *Occup Environ Med*. 2015;72(2):138–144.
40. Hallock MF, Smith TJ, Woskie SR, et al. Estimation of historical exposures to machining fluids in the automotive industry. *Am J Ind Med*. 1994;26(5):621–634.
41. Woskie SR, Virji MA, Hallock M, et al. Summary of the findings from the exposure assessments for metalworking fluid mortality and morbidity studies. *Appl Occup Environ Hyg*. 2003;18(11):855–864.
42. Naimi AI, Cole SR, Hudgens MG, et al. 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(2):246–254.
43. Hernán MA, Cole SR, Margolick J, et al. Structural accelerated failure time models for survival analysis in studies with time-varying treatments. *Pharmacoepidemiol Drug Saf*. 2005; 14(7):477–491.
44. Mark SD, Robins JM. Estimating the causal effect of smoking cessation in the presence of confounding factors using a rank preserving structural failure time model. *Stat Med*. 1993; 12(17):1605–1628.
45. Naimi AI, Richardson DB, Cole SR. Causal inference in occupational epidemiology: accounting for the healthy worker effect by using structural nested models. *Am J Epidemiol*. 2013; 178(12):1681–1686.
46. Picciotto S, Brown DM, Chevrier J, et al. Healthy worker survivor bias: implications of truncating follow-up at employment termination. *Occup Environ Med*. 2013;70(10): 736–742.
47. Marzilli M, Merz CN, Boden WE, et al. Obstructive coronary atherosclerosis and ischemic heart disease: an elusive link! *J Am Coll Cardiol*. 2012;60(11):951–956.
48. US Environmental Protection Agency. Final rule. National ambient air quality standards for particulate matter. https://www.federalregister.gov/articles/2013/01/15/2012-30946/national-ambient-air-quality-standards-for-particulate-matter. Published January 15, 2013. Accessed May 22, 2015.
49. Mirer FE. New evidence on the health hazards and control of metalworking fluids since completion of the OSHA advisory committee report. *Am J Ind Med*. 2010;53(8):792–801.
50. Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? *Epidemiology*. 2009; 20(1):3–5.
51. VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology*. 2009;20(6):880–883.
52. Naimi AI, Tchetgen Tchetgen EJ. Invited commentary: Estimating population impact in the presence of competing events. *Am J Epidemiol*. 2015;181(8):571–574.
53. Hernán MA, VanderWeele TJ. Compound treatments and transportability of causal inference. *Epidemiology*. 2011;22(3): 368–377.
54. Tchetgen Tchetgen EJ. Identification and estimation of survivor average causal effects. *Stat Med*. 2014;33(21): 3601–3628.
55. Shardell M, Hicks GE, Ferrucci L. Doubly robust estimation and causal inference in longitudinal studies with dropout and truncation by death. *Biostatistics*. 2015;16(1):155–168.
56. Cole SR, Richardson DB, Chu H, et al. Analysis of occupational asbestos exposure and lung cancer mortality using the g formula. *Am J Epidemiol*. 2013;177(9):989–996.

57. Keil AP, Edwards JK, Richardson DB, et al. The parametric g-formula for time-to-event data: intuition and a worked example. *Epidemiology*. 2014;25(6):889–897.
58. Brown DM, Petersen M, Costello S, et al. Occupational exposure to $PM_{2.5}$ and incidence of ischemic heart disease: longitudinal targeted minimum loss-based estimation. *Epidemiology*. 2015;26(6):806–814.
59. Joffe MM, Yang WP, Feldman H. G-estimation and artificial censoring: problems, challenges, and applications. *Biometrics*. 2012;68(1):275–286.
60. Picciotto S, Chevrier J, Balmes J, et al. Hypothetical interventions to limit metalworking fluid exposures and their effects on COPD mortality: G-estimation within a public health framework. *Epidemiology*. 2014;25(3):436–443.