ALPHA FOUNDATION FOR THE IMPROVEMENT OF MINE SAFETY AND HEALTH

Final Technical Report

Project Title: Revisiting Lung Cancer Risk from Silica Exposure in Miners: Proposed Standards, Prevailing Biases, and Modern Methodology

Grant Number: AFC215-31

Organization: The Regents of the University of California

Principal Investigator: Sadie Costello, PhD

Contact Information:
Email: sadie@berkeley.edu
School of Public Health
2121 Berkeley Way #5302
Berkeley, CA 94720-7360
Phone: (510) 643-5716

Period of Performance: September 1st, 2015 – August 31st, 2018

Acknowledgement/Disclaimer: This study was sponsored by the Alpha Foundation for the Improvement of Mine Safety and Health, Inc. (ALPHA FOUNDATION). The views, opinions and recommendations expressed herein are solely those of the authors and do not imply any endorsement by the ALPHA FOUNDATION, its Directors and staff.

1.0 Executive Summary (1pg)

Silica is one of the most common occupational exposures worldwide and silicosis is the oldest known occupational disease. In 2016, Occupational Safety and Health Administration (OSHA) passed a silica standard with a permissible exposure limit (PEL) of 0.05 mg/m^3 for respirable crystalline silica, approximately half of the former PEL. The Mine Safety and Health Administration (MSHA) has indicated that they intend to follow suit with a similarly stringent silica standard, which has prompted industry to suggest that the former permissible exposure limit (PEL) is adequate to protect worker health. In addition to causing silicosis, silica has been classified as a human carcinogen by the International Agency for Research on Cancer (IARC), however the association between silica and lung cancer is still controversial. Much of the controversy stems from inconsistencies in results from occupational epidemiology studies. There are, however, several systematic biases that lead to the underestimation of effects from occupational studies, the most well-known of which is the Healthy Worker Survivor Effect (HWSE). Additionally, inadequate adjustment for competing events and the use of conventional metrics of protracted silica exposure such as cumulative exposure (which bundles duration and intensity) could also mask the true etiologic effect. We propose to reanalyze Checkoway's cohort of diatomaceous earth miners, one of the studies that influenced the IARC classification, to address these biases and further illuminate the shape of the exposure-response curve at the levels of the current and former PELs.

First, in order to disentangle the rate, duration, and timing of silica exposure, we implemented a distributed lag non-linear model (DLNM) which parses out the temporal relationships between exposure and the risk of health outcomes as the 'exposure-lag-response'. By comparing the results for exposure scenarios with the same cumulative exposure but different intensity and duration, the health effects associated with intensity, duration, and lag can be disentangled. Measures of association from DLNMs were generally higher than those from simpler (standard) models. For example, rate ratios from penalized DLNMs corresponding to average daily exposures of 0.4 mg/m^3 during lag years 31 – 50 prior to the age of observed cases indicated a 47% increase in lung cancer mortality versus 15% from a simple model of the same exposure scenario.

Second, we addressed the HWSE using two different methods, each considering competing risk by non-malignant respiratory disease. First, we used the parametric g-formula to estimate the risk of lung cancer in this population under two scenarios 1) if they had been unexposed to silica and 2) if they had been exposed to ≤0.05 mg/m^3 of silica during every year of employment and compared each scenario to the risk of lung cancer the workers actually experienced. We estimate that the risk of lung cancer in this cohort would have been 18% less if workers had been completely unexposed to silica and 14% less if they had not been exposed above the current OSHA PEL. Second we used g-estimation of a structural failure time model to estimate the years of life lost due to silica-related lung cancer mortality in this cohort. We estimated that the median number of years of life lost per worker because of lung cancer due to silica exposure was 2.21. Sensitivity analyses of exposure at the levels of the current and former PEL

supported the main results, but with larger estimates, and suggested that a strict limit would have been nearly as effective as a complete ban on silica.

This work resulted in 4 oral presentations at an international occupational epidemiology conference and 3 peer-reviewed publications in top journals. Overall, we found evidence that addressing prevailing biases in occupational epidemiology is critical to understanding the lung cancer risk from silica exposure. Evidence from our research suggests that if no worker had been exposed above the current OSHA PEL there would have been less lung cancer mortality in this cohort and these workers would have lived longer. Findings such as these are critical to informing future standards to protect the health of miners.

2.0 Problem Statement and Objective

Focus area: Injury and Disease Exposure and Risk Factors
Topical area: Respiratory Disease

2.1 Problem Statement

Silica is one of the most common occupational exposures worldwide and silicosis is the oldest known occupational disease. In 1971 the Occupational Safety and Health Administration (OSHA) promulgated the existing permissible exposure limit (PEL) for silica, which was based on a formula of crystalline silica per volume of air, and which remained in place until very recently. In addition to causing silicosis[1], exposure to silica is associated with lung cancer[2]. The International Agency for Research on Cancer (IARC) classified silica as a group 1 human carcinogen in 1997;[3] however the classification was controversial, in part due to inconsistent results from epidemiological studies. In 2001, Steenland et al[2] published a pooled analysis of 10 cohorts of silica-exposed workers which supported the IARC conclusion . In 2016, OSHA passed a new standard with an enforceable PEL of 50 $\mu g/m^3$ for respirable crystalline silica. Although the Mine Safety and Health Administration (MSHA) has indicated plans to propose a silica standard to protect miners, industry has raised opposition and no such standard has materialized.

Standards are based on results from occupational epidemiology and there are several systematic biases that lead to underestimation of effects in this field, the most well-known of which is the Healthy Worker Survivor Effect (HWSE). It is reasonable to expect HWSE to be particularly strong in studies of silica because exposure causes a debilitating chronic respiratory disease as well as cancer. Other biases that could obscure the true effect of silica on lung cancer include competing risk by silicosis and the use of conventional metrics of protracted silica exposure as cumulative exposure (which bundles duration and intensity). The application of modern methods to estimate the shape of the exposure-response curve at the low end of exposure is needed to guide exposure limits and protect worker health.

2.2 Objective

We proposed to readdress these biases in a cohort of silica-exposed miners in the diatomaceous earth mining industry in California with follow up extended to 2011. This

reanalysis of lung cancer mortality has the potential to influence both the strength and shape of the exposure-response relation across the whole range of silica exposure, particularly the lower end which is of relevance to a future MSHA standard.

Checkoway's cohort study[4,5] of diatomaceous earth miners, one of the studies included in Steenland's pooled analysis of lung cancer and crystalline silica, has had follow up extended to 2011 (personal communication 2014, Harvey Checkoway). Diatomaceous earth is made up of the fossilized remains of certain algae whose cell walls are mostly silica. Workers at two diatomaceous earth mining and processing plants in Lompoc, California were exposed to crystalline silica, principally cristobalite. Lung cancer mortality in this cohort increased with higher levels of exposure before it plateaued and declined at the highest exposure.[6]

Occupational epidemiology studies with detailed exposure histories, such as the diatomaceous earth cohort, often incorporate information about exposure intensity and duration into a single summary metric: cumulative exposure. This metric obscures the differences in exposure rate between and within workers. A recently paper by Gasparrini[7] summarizes the temporal relationships between exposure and the risk of health outcomes as the 'exposure-lag-response'. By comparing the results for exposure scenarios with the same cumulative exposure but different intensity and duration, the health effects associated with intensity, duration, and lag can be disentangled.

In addition to the exposure metric, the strong relationship between silica exposure and silicosis (as well as other NMRD) may obscure the effects of silica on lung cancer. At least one study has found higher lung cancer risk among silicotics,[8] and a history of non-malignant respiratory disease was a risk factor for lung cancer in a recent study of diesel exhaust in non-metal miners, indicating that these diseases are not independent.[9] In mortality studies, a worker who dies of silicosis can obviously no longer die of lung cancer. Ignoring the competing risk and treating the silicosis deaths as "independent censoring" results in a biased effect estimate for lung cancer.[10] In order to correctly account for the non-independence of the censoring event and the outcome, methods such as generating subdistribution functions of risk[11] can be used. Appropriate adjustment for competing risks by silicosis and non-malignant respiratory disease may provide a more complete picture of the lung cancer survival experience of silica exposed miners.

Despite the relatively large body of literature regarding silica, silicosis and lung cancer, there remains a question as to whether excessive lung cancer risk occurs exclusively among the workers with silicosis. Clarification of this question has implications for our understanding of the mechanistic pathways involved as well as for setting a permissible exposure limit. In 1999, Checkoway et al[12] examined this question by calculating Standardized Mortality Ratios (SMRs) for lung cancer separately for workers with and without silicosis. Their results suggest that silicosis is not a necessary co-condition for silica related lung cancer. However, their results would likely have been attenuated due to using the lung cancer rates from the US population in their SMRs. By conducting internal analysis, we will estimate the direct effect of diatomaceous earth exposure on

lung cancer mortality in this population. We proposed to estimate the natural direct effect of exposure, that is, the effect of silica on lung cancer that does not travel through silicosis.

Silica differs from other lung carcinogens in that exposure also causes a chronic lung disease[13] as well as lung cancer which may exacerbate the healthy worker survivor effect. The healthy worker survivor effect arises when workers who are less susceptible to the health effects of exposure accumulate more exposure, via staying employed or keeping high exposure jobs longer. Because healthier workers are less likely to decrease their exposure, effect estimates from conventional methods may appear weak, null, or protective, even when exposure causes disease. Some of the conventional methods to adjust for HWSE include: restricting analysis to those who survive at least 15 years from their initial hire dates;[14] lagging the exposure;[15] and adjusting for current employment status.[16] Steenland et al.'s simulation[17] examined several scenarios in which the HWSE was present, and showed that conventional methods to remove this bias were unsuccessful if exposure increased the probability of leaving work or of becoming ill—in other words, when a time-varying confounder (leaving work or health status) was affected by prior exposure for some individuals. Two methods developed by Robins[18] for analyzing occupational cohort data while adjusting correctly for employment status are 1) the parametric g-formula, based on g-computation, and 2) g-estimation. The application of either method will address the bias evident in the plateauing exposure-response curve in the original publication of this cohort.

2.3 Specific Aims

We propose to readdress prevailing biases in Checkoway's updated study of silica-exposed miners in the diatomaceous earth mining industry in California, one of the cohorts used in Steenland's pooled analysis. Correction for some biases, like handling missing smoking information to control for smoking, was considered in each aim. Each objective below focuses on a different question and requires a distinct modeling approach. Our research objectives were as follows:

1) To refine the exposure-response models between silica exposure and lung cancer mortality by using each worker's detailed exposure history to disentangle the effects of exposure rate and cumulative exposure, adjusting for smoking.
2) To estimate the exposure-response between silica and lung cancer adjusting for competing risk by silicosis using inverse probability of censoring weights (IPCW) and sub-distribution hazards analysis, adjusting for smoking.
3) To determine the total effect of silica exposure on lung cancer parsing out the direct effect and the portion of the effect mediated through silicosis (indirect effect), adjusting for smoking.
4) To estimate the effect of silica exposure on lung cancer mortality adjusting for the healthy worker survivor effect by applying the g-formula, adjusting for smoking.

3.0 Research Approach

1) *To refine the exposure-response models between silica exposure and lung cancer mortality by using each workers' detailed exposure history to disentangle the effects of exposure rate and cumulative exposure, adjusting for smoking.*

In order to disentangle the rate, duration, and timing of exposure, we implemented a distributed lag non-linear model (DLNM), as recently described by Gasparrini.[7] Gasparrini summarizes the temporal relationships between exposure and the risk of health outcomes as the 'exposure-lag-response' and offers a thorough description of how to apply DLNM to longitudinal data with individual, time-varying exposures. The model allows for non-parametric estimation of both the exposure-response and the lag-response. By comparing the results for exposure scenarios with the same cumulative exposure but different intensity and duration, the health effects associated with intensity, duration, and lag can be disentangled. Cigarette smoking was available for 50% of the cohort. For this analysis, we adjusted for cigarette smoking by creating a categorical variable for smoking (ever smoking, never smoking, or missing data).

2) *To estimate the exposure-response between silica and lung cancer adjusting for competing risk by silicosis using two different methods: inverse probability of censoring weights (IPCW) and sub-distribution hazards analysis (Fine & Gray model,) adjusting for smoking.*

We first approached this aim using weights and sub-distributions in Cox proportional hazards models. We presented results at a conference and submitted a manuscript for peer review. Although that manuscript was not accepted, the review comments provided valuable insight for improving the analysis of competing risk of silica exposure on lung cancer by nonmalignant respiratory disease. The reviewers suggested using the g-formula to adjust for time-varying confounding affected by prior exposure and competing risks in the same analysis. Thus, we used the g-formula (as intended in aim 4) to handle healthy worker survivor effect and adjust for competing risk at the same time.

The parametric g-formula, based on g-computation, is a method of accurately adjusting for employment status, a time-varying confounder affected by prior exposure, when analyzing occupational cohort data. Instead of the traditional approach comparing workers with different exposure levels within strata of observed confounders, this method requires modeling workers' observed confounder and outcome values at each time as functions of prior confounder values and prior exposure history. Next, outcomes under different exposure interventions are estimated by Monte Carlo methods. The aforementioned models are used to simulate the confounder and outcome values at each time based on assigned exposure and prior simulated confounder values. Competing risks can also be simulated using the same Monte Carlo methods, in order to give a realistic sense of how the interventions on exposure might affect the outcome of interest in the presence of other outcomes also potentially affected by the interventions.

We adjusted for cigarette smoking (ever/never) in a sensitivity analysis by imputing smoking information for approximately 50% of the cohort with missing smoking data.

This imputation was performed with the SAS procedure MI using information on people who have smoking data to predict smoking status for the others. The imputation was done 50 times and we presented point estimates from the parametric g-formula on each of the 50 datasets after adjusting for the imputed smoking variable.

3) *To determine the total effect of silica exposure on lung cancer parsing out the direct effect and the portion of the effect mediated through silicosis (indirect effect), adjusting for smoking.*

We aimed to estimate the natural direct effect, as opposed to the more standard controlled direct effect. The natural direct effect describes what happens when the effect of exposure on the intermediate is blocked, but the intermediate is allowed to vary as it naturally would in the absence of exposure. Since silica exposure is necessary for a diagnosis of silicosis, the natural direct effect estimates the effect of silica exposure on lung cancer that does not occur via silicosis. By comparison with a total effect of exposure on the lung cancer outcome, we planned to estimate what percentage of any observed effect travels through each of the two pathways of interest.

However, our analysis was limited by the small number of lung cancer deaths with a positive x-ray for silicosis. In this cohort, silicosis information was obtained through surveillance x-ray records, which are an incomplete source of information. Certainly, it is possible that those with normal or unusual x-rays would have had positive x-rays if they had been followed for long enough. Thus, we tried to identify workers who would have developed silicosis had they been followed for longer to help understand the role of silicosis as a mediator between silica exposure and lung cancer.

4) *To estimate the effect of silica exposure on lung cancer mortality adjusting for the healthy worker survivor effect by applying the g-formula.*

In a slight modification of the proposed analysis, we applied g-estimation of structural nested accelerated failure time models to eliminate bias from HWSE in an examination of the relationship between exposure to crystalline silica and survival time in this cohort of diatomaceous earth workers, with a focus on lung cancer and non-malignant respiratory disease mortality as the outcomes of interest.

The structural accelerated failure time model considers counterfactual unexposed survival time (starting from hire) to be a log-linear function of observed exposures and also dependent on observed survival time. Special attention was paid to censoring by competing risks, which is considered informative if the competing event shares a cause with the outcome and is also related to the exposure. Analyses were run with and without weights to adjust for censoring by competing risks: for all natural mortality, the only potential competing risks are external causes of death; for lung cancer mortality, we considered death from NMRD a competing risk, and vice versa. The weights were

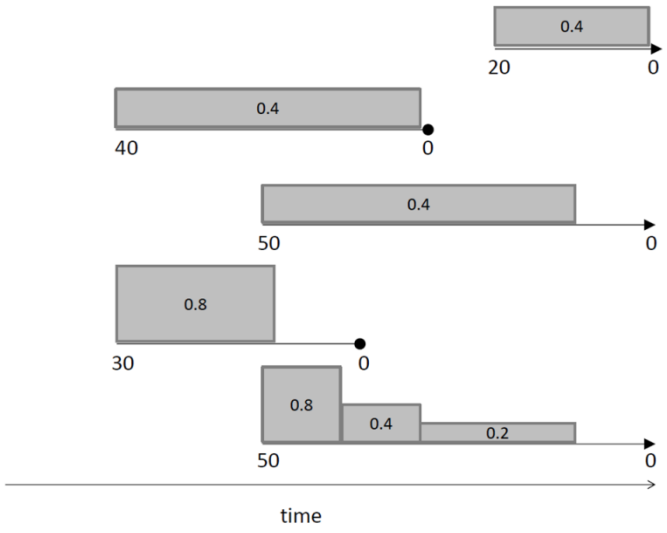
equal to the inverse of the probability of remaining uncensored (i.e., not being lost to follow-up, and not dying from specific competing risks) from that time forward. Similar to aim #1, we adjusted for cigarette smoking by creating a categorical variable for smoking (ever smoking, never smoking, or missing data).

4.0 Research Findings and Accomplishments

Aim 1) In order to disentangle the rate, duration, and timing of exposure, we implemented a distributed lag non-linear model (DLNM), as recently described by Gasparrini.[1]. Gasparrini summarizes the temporal relationships between exposure and the risk of health outcomes as the 'exposure-lag-response'. The model allows for non-parametric estimation of both the exposure-response and the lag-response. By comparing the results for exposure scenarios with the same cumulative exposure but different intensity and duration, the health effects associated with intensity, duration, and lag can be disentangled.

We fitted penalized and unpenalized DLNM with time-varying covariates for lung cancer and non-malignant respiratory disease. We used a generalized additive Poisson model for the penalized models and a Cox proportional hazards model for the unpenalized models. For both types of models, the exposure-lag-response function was determined by a combination of functions for the exposure-response and lag-response. The combination of functions for both silica and asbestos exposures were entered in the regression model along with other covariates: calendar time (as a linear term), and indicator variables for Hispanic ethnicity and smoking (ever, never or missing). The exposure-response and the lag-response were modeled in the following ways: exposure-response was modeled with a linear term, categorically, with a natural spline, and with a penalized spline; lag-response was modeled as constant, categorically, with a natural spline, and with a penalized spline. Model fit for Cox models was determined by the AIC. We estimated the HR for different silica exposure scenarios and both lung cancer and NMRD.

Figure 1 is a representation of different exposure scenarios over time for different participants in a longitudinal cohort study. Each line represents a hypothetical cohort participant, with the lag dimension labeled on each line (with increasing numbers in the opposite direction of the follow-up) and the participant's exposure represented by the rectangle(s) above each line. The height of each rectangle



represents the intensity of exposure proportional to the decimal numeral inside the rectangle (e.g., in mg/m3), and the width represents the duration in years. The product (intensity x duration) represents the cumulative exposure for each rectangle. Circles represent participants who are censored after experiencing an outcome of interest, and arrows represent participants who are still at risk at the administrative end of follow-up.

Compared to a "simple model", equivalent to a simple linear term for cumulative exposure, exposure-lag-response estimates from DLNMs were generally higher for a variety of different exposure scenarios compared to estimates from models assuming constant lag-response and linear exposure-response. Overall, point estimates from the unpenalized models were higher than the estimates from the penalized models. Below are results from the "simple model", best fitting unpenalized DLNM, and penalized DLNM for silica exposure and lung cancer (Table 1.1) and silica exposure and NMRD (Table 1.2)

Table 1.1: HR (95% CI) associated with different exposure scenarios from models with varying exposure-lag-response functions for lung cancer mortality

	Exposure Scenarios			Simple model*	AIC chosen unpen. DLNM*	Pen. DLNM**
	Intensity (mg/m^3)	Timing	Cumulative (mg/m^3-years)			
1	0.2	lag 1-20	4	1.07 (0.94, 1.22)	1.40 (0.98, 2.00)	1.11 (0.94, 1.31)
2	0.2	lag 1-40	8	1.15 (0.89, 1.48)	1.95 (0.95, 4.00)	1.49 (0.98, 2.27)
3	0.2	lag 11-50	8	1.15 (0.89, 1.48)	1.95 (0.95, 4.00)	1.54 (0.99, 2.40)
4	0.4	lag 11-30	8	1.15 (0.89, 1.48)	1.81 (0.97, 3.37)	1.61 (0.93, 2.79)
5	0.4	lag 31-50	8	1.15 (0.89, 1.48)	1.81 (0.97, 3.37)	1.47 (0.92, 2.35)
6	0.8	lag 31-40	8	1.15 (0.89, 1.48)	1.49 (1.00, 2.24)	1.55 (0.94, 2.53)
7	1.0	lag 31-40	10	1.19 (0.86, 1.63)	1.46 (0.96, 2.24)	1.40 (0.93, 2.10)
8	1.0	lag 41-50	10	1.19 (0.86, 1.63)	1.46 (0.96, 2.24)	1.12 (0.87, 1.44)
9	0.2 0.4 0.8	lag 11-30, lag 31-40, lag 41-50	8	1.15 (0.89, 1.48)	1.89 (0.96, 3.71)	1.43 (0.96, 2.14)

*The simple model was based on a constant lag-response and linear exposure-response (df=1), while the unpenalized DLNM chosen based on AIC was based on a constant lag-response function and a natural cubic spline for the exposure-response (df=2).

**Effect estimates from the penalized DLNMs are rate ratios from a Poisson generalized additive model, aiming to approximate a Cox proportional hazards model.

Table 1.2: HR (95% CI) associated with different exposure scenarios from models with varying exposure-lag-response functions for NMRD mortality

	Exposure Scenarios			Simple Model*	AIC chosen unpen. DLNM*	Pen. DLNM**
	Intensity (mg/m^3)	Timing	Cumulative (mg/m^3-years)			
1	0.2	lag 1-20	4	1.11 (1.02, 1.21)	1.11 (0.77, 1.60)	1.03 (0.80, 1.34)
2	0.2	lag 1-40	8	1.23 (1.03, 1.46)	1.73 (0.97, 3.07)	1.18 (0.83, 1.68)
3	0.2	lag 11-50	8	1.23 (1.03, 1.46)	2.09 (1.18, 3.69)	1.46 (1.04, 2.06)
4	0.4	lag 11-30	8	1.23 (1.03, 1.46)	1.63 (0.76, 3.48)	1.19 (0.72, 1.97)
5	0.4	lag 31-50	8	1.23 (1.03, 1.46)	2.42 (1.27, 4.60)	1.80 (1.14, 2.85)
6	0.8	lag 31-40	8	1.23 (1.03, 1.46)	1.59 (0.83, 3.04)	1.39 (0.84, 2.30)
7	1.0	lag 31-40	10	1.29 (1.04, 1.61)	1.39 (0.70, 2.77)	1.48 (0.86, 2.55)
8	1.0	lag 41-50	10	1.29 (1.04, 1.61)	3.61 (1.63, 7.98)	2.51 (1.31, 4.81)
9	0.2 0.4 0.8	lag 11-30, lag 31-40, lag 41-50	8	1.23 (1.03, 1.46)	2.31 (1.26, 4.23)	1.73 (1.17, 2.55)

* The simple model was based on a constant lag-response and linear exposure-response (df=1), while the unpenalized DLNM chosen based on AIC was based on natural cubic spline functions for both the lag-response and exposure-response (df=6).

** Effect estimates from the penalized DLNMs are rate ratios from a Poisson generalized additive model, aiming to approximate a Cox proportional hazards model.

Our findings indicate that intensity, timing, and duration are all potentially relevant aspects of exposure, and approaches relying on cumulative exposure likely underestimated effects for various exposure scenarios compared to more flexible DLNM approaches. Different lag-response shapes were observed for malignant compared to non-malignant respiratory disease mortality, but our findings were suggestive of delayed exposure effects (latency) for both outcomes.

None of the results for lung cancer were statistically significant, whereas many of the exposure-lag scenarios for non-malignant respiratory disease were statistically significant. Distributed lag models rely on *a priori* determinations of various lag-and-exposure functions and there are no well-established criteria for comparing model fit or selecting the “best” model. This problem is alleviated when using a penalized spline distributed lag model, but remains a limitation of the approach.

2) We used the parametric g-formula to assess the cumulative risk of lung cancer and non-malignant respiratory disease mortality under hypothetical interventions on crystalline silica exposures. This method allows us to assessed counterfactual risk under different exposure scenarios while accounting for competing causes of death and addressing the healthy worker survivor effect. A counterfactual risk means the risk that these same workers would have had if they had a different exposure than the one they actually did.

We considered two hypothetical interventions on silica exposure, one setting a hypothetical maximum exposure limit on average daily crystalline silica exposures equivalent to the current OSHA PEL of 50 μg/m^3 for the duration of follow-up, and one setting silica exposures to zero. For the intervention setting a hypothetical exposure limit of 50 μg/m^3, all predicted silica exposure values above 50 μg/m^3 were replaced with 50 μg/m^3 and otherwise remained unchanged. We compared both interventions to no intervention, i.e., the observed natural course (what actually happened).

Estimates of risk of lung cancer and non-malignant respiratory disease mortality under hypothetical interventions on crystalline silica exposures, along with risk ratios (RR) and risk differences (RD) compared to the natural course are presented in Table 2.1 below. Since silica exposure is associated with increased risk of disease, reducing exposure would reduce the amount of disease – therefore, we expect to see point estimates below 1.

Table 2: Cumulative risk of lung cancer and non-malignant respiratory disease mortality at age 90 under the natural course and under hypothetical interventions on crystalline silica, along with RR and RD estimated with corresponding 95% CI comparing each intervention to the natural course.

Intervention	Cumulative risk	Range[a]	RR	95% CI	RD	95% CI
Lung cancer mortality						
Simulated natural course	7.2	5.1, 15.4	1.00	.	0.0	.
Annual average Silica ≤0.05 mg/m^3	6.2	3.8, 10.9	0.86	0.63, 1.22	-1.0	-3.4, 1.4
Annual average Silica = 0 mg/m^3	5.9	3.6, 11.8	0.82	0.53, 1.26	-1.3	-4.0, 1.4
Non-malignant respiratory disease mortality						
Simulated natural course	12.0	6.9, 15.1	1.00	.	0.0	.
Annual average Silica ≤0.05 mg/m^3	8.3	5.0, 13.1	0.69	0.52, 0.93	-3.7	-7.2, -0.2

Annual average Silica = 0 mg/m^3	7.5	4.9, 13.8	0.63	0.43, 0.91	-4.5	-8.3, -0.7

Abbreviations: CI, confidence interval; RD, risk difference; RR, risk ratio.
[a]Range of risk estimates from 200 bootstrap samples.

The RR for lung cancer mortality risk under an intervention setting a hypothetical exposure limit equivalent to the current OSHA PEL compared to the observed risk was 0.86 (95% CI: 0.63, 1.22). In other words, there would have been 14% less lung cancer mortality in this cohort if no one had been exposed above the OSHA PEL. The corresponding RR for non-malignant respiratory disease under the same intervention was 0.69 (95% CI: 0.52, 0.93). Interventions setting exposure to zero resulted in RR = 0.82 (95% CI: 0.53, 1.26) for lung cancer and 0.63 (95% CI: 0.43, 0.91) for non-malignant respiratory disease mortality, respectively.

It should be noted that the parametric g-formula relies on a large number of parametric models. Increased likelihood of model misspecification is a recognized limitation of this approach. Our findings suggest that risks from both outcomes would have been considerably lower if historical silica exposures in this cohort had not exceeded current regulatory limits, although only the findings for non-malignant respiratory disease were statistically significant.

3) Initially, we found little evidence of an indirect effect of silica (via silicosis) on lung cancer. There are very few workers in this cohort who had both silicosis, as determined by surveillance x-ray, and mortality from lung cancer. Many people in this cohort don't have a surveillance x-ray after leaving work, thus our last measure of their silicosis status could be years before they would be expected to develop silicosis. We created a tiered system identifying the degree to which we are confident that the silicosis status at last x-ray represents a reasonable final status for each worker. We investigated the effect of multiply imputing silicosis status after setting to missing silicosis statuses with lower confidence. Ultimately, however, we settled on identifying workers who were at risk for developing silicosis after their last x-ray as those with a 0/1 status at the final observed chest x-ray, referred to below as "borderline silicosis". Our results change if we include these workers.

Table 3: Estimation of the Natural Direct and Indirect Effects of Silica Dust Exposure on Lung Cancer as Mediated by Silicosis Diagnosis Under Different Classifications of Borderline Silicosis Status

Removing all borderline silicosis diagnoses				
Natural Direct Effect		Indirect Effect		
Silica metric	HR	Silica metric	HR	% total effect mediated by silicosis
mg/m^3	1.024	mg/m^3	1.002	8
highest vs lowest quantile	1.41	highest vs lowest quantile	1.02	4

Classifying borderline silicosis diagnoses as silicosis				
Natural Direct Effect		Indirect Effect		

Silica metric	HR	Silica metric	HR	% total effect mediated by silicosis
mg/m^3	1.022	mg/m^3	1.021	49
highest vs lowest quantile	1.26	highest vs lowest quantile	1.09	26

Table 3 contains estimates of the Natural Direct Effect (NDE) and the Indirect Effect (IDE) of silica dust exposure on lung cancer, as mediated by silicosis. The NDE represents the amount of the exposure effect that does not travel through silicosis, while the IDE represents the remaining portion of the effect that does travel through silicosis. The reported percent represents the percentage of the total effect that travels through the indirect pathway. We present results from two different cohorts, one in which borderline silicosis cases are removed from the analysis and the other in which borderline silicosis cases are included and classified as silicosis.

We observed that when borderline cases are classified as silicotic, there appears to be a significant indirect effect (49% when exposure is modeled continuously, 26% when modeled categorically). However, when these cases are removed from the cohort, this indirect effect appears much smaller (8% when exposure is modeled continuously, 4% when modeled categorically). We used a non-parametric simulation to determine the likelihood of the effect occurring by chance, and found that the p-values associated with both findings were less than 0.01.

Certainly, surveillance data are incomplete and it is possible that those with unusual x-rays would have had positive x-rays if they had been followed for long enough. However, 0/1 x-rays do not have a clinically meaningful interpretation and drawing conclusions based on these x-rays is beyond our expertise.

4) We applied g-estimation of structural nested accelerated failure time models to eliminate bias from HWSE in an examination of the relationship between exposure to crystalline silica and survival time in this cohort of diatomaceous earth workers, with a focus on lung cancer and non-malignant respiratory disease mortality as the outcomes of interest.

First, we conducted a path analysis to determine whether HWSE could cause bias in this cohort: was employment status associated with future exposure and outcomes, and also affected by prior exposure? The path analysis showed that employment duration was significantly associated with natural-cause mortality, lung cancer, and NMRD, confirming that it is a confounder. In addition, exposure shortened time to leaving work: workers would have terminated employment a median of 0.47 (95%CI: 0.11,0.93) years later if they had never been exposed. Thus, time-varying confounding affected by prior exposure was present, necessitating the use of g-methods to avoid bias from HWSE.

Once we determined that employment was a time-varying confounder affected by prior exposure, we proceeded to assess the relationship between exposure and the outcomes of interest using g-estimation in order to adjust correctly for this confounding. The structural accelerated failure time model considers counterfactual unexposed survival time (starting from hire) to be a log-linear function of observed exposures and also dependent on observed survival time. Special attention was paid to censoring by

competing risks, which is considered informative if the competing event shares a cause with the outcome and is also related to the exposure. Analyses were run with and without weights to adjust for censoring by competing risks: for all natural mortality, the only potential competing risks are external causes of death; for lung cancer mortality, we considered death from NMRD a competing risk, and vice versa. The weights were equal to the inverse of the probability of remaining uncensored (i.e., not being lost to follow-up, and not dying from specific competing risks) from that time forward.

For mortality from all natural causes, the ratio of median survival times comparing what would have happened if everyone had been exposed to crystalline silica at an average daily intensity of 0.1mg/m^3 (approximately the 32nd percentile of nonzero exposures), every year from the date they entered the cohort to the end of follow-up, with what would have happened if no one had been exposed during follow-up was 0.970 (95%CI: 0.943,0.999). That is, exposing everyone to this level of crystalline silica in every year from their entry into the cohort until the end of follow-up without regard for employment status would have shortened survival time from cohort entry by at least 3% for half of the workers, compared to everyone being unexposed, a statistically significant result. The corresponding estimate for lung cancer (0.902 [0.859,0.947]) was considerably stronger and statistically significant. The results for NMRD were sensitive to the use of weights to adjust for censoring by competing deaths from lung cancer. With adjustment, the relationship of survival time with silica exposure was also strong, with a statistically significant ratio of median survival times estimated at 0.900 (0.837,0.968).
The median numbers of years of life that could have been saved per worker if the worker had not had any silica exposure are presented in Figure 2.

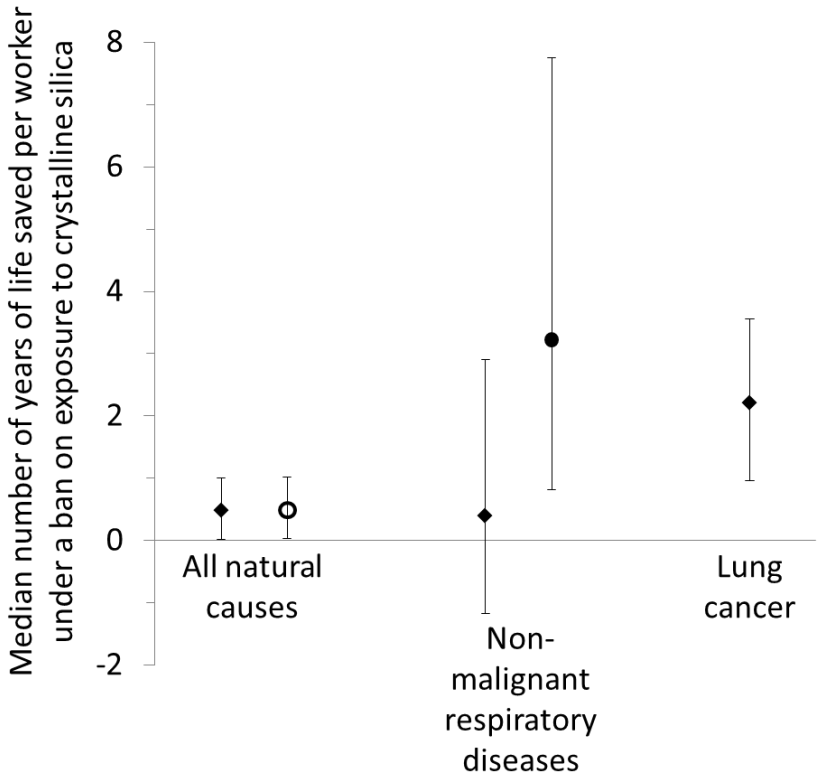


Figure 2: Median number of years of life that could have been saved per worker if exposure to crystalline silica had been eliminated starting in 1925, among workers who died of various causes during follow-up. All analyses adjusted for the following confounders: Latino ethnicity, age, calendar year, smoking (ever/never/missing), time taken off work in the previous year, exposure to silica in the previous year, previous cumulative exposures to dust and asbestos, and employment duration prior to start of follow-up. Diamonds represent estimates from analyses that were adjusted

for censoring by loss to follow-up. The hollow circle represents the estimate from an analysis additionally adjusted for censoring by deaths from external causes. The solid circle represents the estimate from an analysis additionally adjusted for censoring by deaths from lung cancer.

These results support the hypothesis that lung cancer death acts as an informative censoring event in the analysis of NMRD: without adjustment for competing risks, the effect of exposure to crystalline silica on NMRD mortality appears to be null. By contrast, no estimate was obtained at all when using weights to adjust for censoring by NMRD death in the analysis of lung cancer, though the estimate without weights was convincing. The lung cancer deaths generally occurred at younger ages than deaths from NMRD, perhaps indicating that the people at risk of lung cancer did not live long enough for life-threatening NMRD to develop. In that case, the analysis without weights makes sense.

The g-estimation approach was sensitive to the adjustment for competing risks and failed to produce any estimate for lung cancer adjusting for NMRD death. This limitation of g-estimation has not been discussed in the literature before, thus we have no established framework in which to interpret this problem.

Overall, using a method that controls bias due to the HWSE, we estimated that survival times would have been significantly longer under a hypothetical intervention banning exposure. While it has been known for some time that workplace silica exposure increases the *risks* of diseases that shorten life, our results quantify the *amount* of life lost due to exposure to crystalline silica.

Lack of smoking data as a limitation

The lack of detailed information on smoking for every member of our cohort is a limitation for each of our analyses. In this study, only crude smoking information was available for half of the cohort, so there may be have been residual confounding by smoking or bias due to the methods we used to handle missing smoking information (either using a category for missing data or multiple imputation techniques).

5.0 Publication Record and Dissemination Efforts

We presented these results in 4 oral presentations at an international occupational epidemiology conference, EPICOH, and in 3 published peer-review manuscripts.

The following two abstracts were presented in Barcelona in 2016:

> Brown D, Eisen E, Neophytou A, Picciotto S, Costello S. O05-4 Mediation analysis of the role of silicosis in the relationship between silica exposure and lung cancer. In: *Oral Session 5 – Respiratory 1*. Vol 73. BMJ Publishing Group Ltd; 2016:A10.2-A10. doi:10.1136/oemed-2016-103951.26.
>
> Neophytou A, Eisen E, Brown D, Picciotto S, Costello S. O15-4 Estimating

absolute risk in the presence of confounders and competing risks: combining inverse probability weights and a cumulative incidence function in an occupational study of crystalline silica and lung cancer. In: *Oral Session 15 – Statistical Methods*. Vol 73. BMJ Publishing Group Ltd; 2016:A29.1-A29. doi:10.1136/oemed-2016-103951.77.

The following two abstracts were presented in Edinburgh in 2017:

Neophytou A, Picciotto S, Brown D, Eisen E, Checkoway H, Costello S. 0137 Exposure-lag-response in occupational epidemiology: application of distributed non-linear lag models in a cohort of diatomaceous earth workers exposed to crystalline silica. In: *Oral Presentation*. Vol 74. BMJ Publishing Group Ltd; 2017:A40.1-A40. doi:10.1136/oemed-2017-104636.108.

Picciotto S, Brown DM, Neophytou AM, et al. 0190 Occupational exposure to crystalline silica and death from lung cancer: g-estimation of structural accelerated failure time models. In: *Oral Presentation*. Vol 74. BMJ Publishing Group Ltd; 2017:A57.3-A58. doi:10.1136/oemed-2017-104636.154.

Work on this project has resulted in the following three peer-reviewed publications:

Neophytou AM, Picciotto S, Brown DM, et al. Exposure-Lag-Response in Longitudinal Studies: Application of Distributed-Lag Nonlinear Models in an Occupational Cohort. *Am J Epidemiol*. 2018;187(7):1539-1548. doi:10.1093/aje/kwy019.

Neophytou AM, Picciotto S, Brown DM, et al. Estimating Counterfactual Risk Under Hypothetical Interventions in the Presence of Competing Events: Crystalline Silica Exposure and Mortality From 2 Causes of Death. *Am J Epidemiol*. 2018;187(9):1942-1950. doi:10.1093/aje/kwy077.

Picciotto S, Neophytou AM, Brown DM, Checkoway H, Eisen EA, Costello S. Occupational silica exposure and mortality from lung cancer and nonmalignant respiratory disease. *Environ Epidemiol*. August 2018:1. doi:10.1097/EE9.0000000000000029.

6.0 Conclusions and Impact Assessment

Evidence from our research suggests that reducing respirable silica exposures in mines to the 2016 OSHA PEL of 0.05 mg/m^3 would protect workers from death from non-malignant respiratory disease, and most likely from lung cancer as well. We demonstrated that if no worker had been exposed above the current OSHA PEL there would have been 31% less mortality from non-malignant respiratory disease and most likely 14% less mortality from lung cancer in this cohort, and these workers would have lived longer. While our research indicated that reducing silica exposures even further

would protect more workers, the vast majority of the protection would be achieved by reducing annual exposure to respirable silica to the current OSHA PEL. We used a new class of statistical methods that allows for the estimation of risk if no worker was exposed above a certain level. These methods were developed for use in occupational cohorts; however, until recently, they have mostly been used to analyze data from clinical trials. Although we were somewhat limited by available data regarding smoking and the relatively small size of the cohort, results from this study are a big step forward in illuminating the risk of mortality from malignant and nonmalignant respiratory disease from silica exposure at the current OSHA PEL.

7.0 Recommendations for Future Work

We have been funded by the Alpha Foundation to study risk of chronic disease due to diesel exhaust in the DEMS cohort of non-metal miners. We have previously found evidence of HWSE in DEMS and will continue to refine these methods to address HWSE as well as unpack cumulative diesel exposure to understand the relative contributions of exposure lag, intensity and duration. Future work regarding the shape of the exposure-response curve for silica exposure and lung cancer should address these biases and be conducted in larger cohorts for better statistical power.

8.0 References

1. Hughes J, Jones R, Gilson J, et al. Determinants of progression in sandblasters' silicosis. *Ann Occup Hyg*. 1982;26(1-4):701-712.
2. Steenland K, Mannetje A, Boffetta P, et al. Pooled exposure - response analyses and risk assessment for lung cancer in 10 cohorts of silica-exposed workers: an IARC multicentre study. *Cancer Causes Control*. 2001;12(9):773-784.
3. International Agency for Research on Cancer (IARC). *Monographs on the Evaluation of Carcinogentic Risk to Humans, Vol. 68: Silica, Some Silicates, Coal Dust and Para-Aramid Fibrils.* Lyon; 1997.
4. Checkoway H, Heyer N, Demers P, Breslow N. Mortality among workers in the diatomaceous earth industry. *Br J Ind Med*. 1993;50(7):586-597.
5. Checkoway H, Heyer N, Seixas N, et al. Dose-response associations of silica with nonmalignant respiratory disease and lung cancer mortality in the diatomaceous earth industry. *Am J Epidemiol*. 1997;145(8):680-688.
6. Eisen E, Agalliu I, Thurston S, Coull B, Checkoway H. Smoothing in occupational cohort studies: an illustration based on penalised splines. *Occup Env Med*. 2004;61(10):854–860.
7. Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. *Stat Med*. 2014;33(5):881-899. doi:10.1002/sim.5963.
8. de Klerk N, Musk A. Silica, compensated silicosis, and lung cancer in Western Australian goldminers. *Occup Env Med*. 1998;55(4):243-248.
9. Silverman DT, Samanic CM, Lubin JH, et al. The Diesel Exhaust in Miners study: a nested case-control study of lung cancer and diesel exhaust. *J Natl Cancer Inst*. 2012;104(11):855-868. doi:10.1093/jnci/djs034.
10. Andersen PK, Geskus RB, De witte T, Putter H. Competing risks in epidemiology:

Possibilities and pitfalls. *Int J Epidemiol*. 2012;41(January):861-870.
11. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol*. 2009;170(2):244-256.
12. Checkoway H, Hughes JM, Weill H, Seixas NS, Demers P. Crystalline silica exposure, radiological silicosis, and lung cancer mortality in diatomaceous earth industry workers. *Thorax*. 1999;54(1):56-59.
13. Park R, Rice F, Stayner L, Smith R, Gilbert S, Checkoway H. Exposure to crystalline silica, silicosis, and lung disease other than cancer in diatomaceous earth industry workers: a quantitative risk assessment. *Occup Env Med*. 2002;59(1):36-43.
14. Fox A, Collier P. Low mortality rates in industrial cohort studies due to selection for work and survival in the industry. *Br J Prev Soc Med*. 1976;30(4):225-230.
15. Gilbert E. Some confounding factors in the study of mortality and occupational exposures. *Am J Epidemiol*. 1982;116(1):177-88.
16. Gilbert E. Analysis of the mortality of workers in a nuclear facility. *Radiat Res*. 1979;79(1):122-48.
17. Steenland K, Deddens J, Salvan A, Stayner L. Negative bias in exposure-response trends in occupational studies: modeling the healthy workers survivor effect. *Am J Epidemiol*. 1996;143(2):202-210.
18. Robins JM. A new approach to causal inference in mortality studies with sustained exposure periods - application to control of the healthy worker survivor effect. *Math Model*. 1986;7:1393-1512.

9.0 Appendices

None