Cover Page

ALPHA FOUNDATION FOR THE IMPROVEMENT OF MINE SAFETY AND HEALTH

Final Technical Report

Project Title: Explore the Integration of Distributed Lag Models for Diesel Exhaust with Methods to Control Healthy Worker Survivor Bias to Assess Benefits of Exposure Interventions in Non-metal Miners
Grant Number: AFC618-63
Organization: Regents of the University of California (UC Berkeley)
Principal Investigator: Ellen A. Eisen
Contact Information : eeisen@berkeley.edu PH: (510) 643-5310 Fax: (510) 642-5815
Period of Performance: January 1, 2018 – May 31, 2019

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:

Long term exposure to diesel exhaust is known to increase the risk of lung cancer, but there is a long latency period—it takes many years from first exposure to a diagnosis of lung cancer, and recent exposure just before the diagnosis does not contribute to that risk. By contrast, little is understood about the relative importance of the intensity (concentration), duration, and timing of diesel exhaust exposure in relation to other disease endpoints, such as ischemic heart disease (IHD) and nonmalignant respiratory diseases such as chronic obstructive pulmonary disease (COPD).

We addressed this gap in an exploratory analysis of the Diesel Exhaust in Miners Study (DEMS), a cohort that includes all miners working at any of eight non-metal mines for at least 1 year after dieselization, either on the surface or underground. Follow-up started between 1947 and 1967 across the eight mines and ends on Dec. 31, 1997. The datasets have been made publicly available by NIOSH and NCI through the National Center for Health Statistics.

We explored the application of the Distributed Lag Nonlinear Model (DLNM) to tease apart the relative importance of the intensity, duration and timing of diesel exhaust, measured as respirable elemental carbon (REC), in relation to risk of COPD and IHD mortality. By comparing results between exposure scenarios with the same cumulative exposure ($\mu g/m^3$-years), but different intensity and timing (including duration), the health effects of these different elements of protracted occupational exposure can be disentangled.

G-computation (including the parametric g-formula) is a method for addressing healthy worker survivor bias, a phenomenon that arises when workers terminate employment for health-related reasons, so that the healthiest workers accumulate the most exposure. We explored a novel approach to combine the DLNM results with the parametric g-formula, thereby addressing both the key temporal aspects of exposure and the healthy worker survivor effect in the DEMS cohort.

Because we want to disentangle the exposure history—intensity and timing—relative to death due to COPD or IHD, DLNM results for each outcome are presented either as one 3-dimensional graph of the exposure-lag-response or as two 2-dimensional graphs showing the exposure-response for a selected lag and the lag-response for a selected exposure intensity. For COPD mortality, the models indicate that diesel exhaust exposures of 300 $\mu g/m^3$ occurring 10 to 40 years prior to the year at risk appear to be influential, with the lag-response peaking just over 20 years prior to death. Regarding the exposure-response, elevated risk of COPD appears to be associated with REC exposure intensities of at least 250 $\mu g/m^3$. The g-formula analysis incorporating the DLNM results suggested that eliminating exposure to REC could have reduced the risk of IHD from 14.1% to 12.9%. All models had wide confidence bands, reflecting a high degree of uncertainty, particularly for COPD.

The substantive results described above were underpowered due to relatively small numbers of cases. This exploratory work, however, has developed and tested a framework for how to incorporate a complex lag-response into analyses that use the g-formula to adjust for confounding that would otherwise cause healthy worker survivor bias. This is a major step in understanding how to account for the complex ways that timing, intensity, and duration of exposure to diesel exhaust can affect miners' health.

We plan to build on this work with our new ALPHA-funded grant (AFC820-15) and expect to produce results that are more impactful than those reported here. With follow-up extended from 1997 to 2015, the improved power provided by so many more cases should allow us to detect an effect of diesel exhaust on both IHD and COPD, while accounting for the exposure-lag-response functions estimated using the DLNM approach. Quantifying this effect will help guide interventions to reduce exposure to REC among miners, with the eventual goal of preventing COPD and IHD cases caused by diesel exhaust in this population.

2.0 Problem Statement and Objective:

This report describes the application of a nonlinear distributed lag model for teasing apart the relative importance of the intensity, duration and timing of diesel exhaust exposure in relation to risk of nonmalignant lung and heart disease mortality in a cohort of non-metal miners. It also describes the exploration of a novel approach to combine the distributed lag exposure model with g-computation to address both the key temporal aspects of exposure as well as the healthy worker survivor effect in the Diesel Exhaust in Miners Study (DEMS).

Background

It is well known that lung carcinogens, such as diesel exhaust, can increase the risk of lung cancer incidence (and mortality) years after the exposure occurred. Little is understood, however, about the relative importance of the intensity (concentration), duration, and timing of diesel exhaust exposure in relation to nonmalignant respiratory disease (NMRD) or ischemic heart disease (IHD). In our previous ALPHA-funded studies of lung cancer, IHD and NMRD mortality in relation to diesel exhaust (AFC113-8; AFC316-54), we relied on a cumulative exposure metric, which bundles duration with intensity. Exposure metrics for average daily exposure of diesel exhaust remove the influence of duration in favor of intensity. Average exposure intensity also loses information regarding the variation in intensity over time, even if average intensity is treated as a moving average. Duration of employment, another common metric usually used only in the absence of historical exposure monitoring data, fails to retain any quantitative information about concentration of exposure at all. Cumulative, average intensity, and duration are all attractive metrics in terms of implementation and interpretation, but none take full advantage of rich work- history data sets, such as the one we have available, with temporal resolution of average daily intensity of diesel exposure at the annual level.

Latency periods for cancer are usually addressed in regression models by lagging cumulative or average exposure. For example, in Steenland's pooled analysis of 10 large occupational cohorts exposed to silica, they looked at the widest possible range of possible lags for cumulative exposure, from 0 to 25 years, and found the strongest associations using a 15 year lag.[1] Lagging consists of considering *only* the exposure accumulated *prior to* the 5, 10, 15 or 25 years immediately prior to the outcome. That is, the more recent exposures are ignored in the analysis and assumed to have no effect on the outcome.

By contrast with cancer, studies of nonmalignant lung diseases do not typically consider latency periods. Results from mortality studies of diesel exhaust, radon or silica exposure for chronic heart or lung disease are generally presented based on an unlagged analysis only.[2-4] However, in two recent papers, lagged exposures were presented for NMRD mortality and

associations with silica exposure were strengthened with 10[5] and 15[6] year lags. These findings suggest that the timing of exposure may be an important consideration for non-malignant respiratory disease.

Until recently, there were far fewer workplace-based studies of chronic heart disease than of chronic lung disease, as observed in a commentary by Cullen.[7] In response to that commentary, our group has focused on occupational cohort studies of IHD incidence[8,9] and IHD mortality[10,11] in relation to respirable or fine particulate matter (PM) in cohorts of miners, autoworkers, and aluminum manufacturing workers. Results of several of these studies suggest that recent exposure may be at least as critical for heart disease risk as past cumulative exposure.[8,10] In the Nurse's Health Study, monthly exposures to both fine and coarse PM ($PM_{2.5}$ and $PM_{2.5-10}$) were estimated based on residence to develop metrics of average intensity of exposure in the 1, 3, 24, 36 and 48 months prior to the first incident non-fatal or fatal coronary heart disease event.[12] Associations were strongest for average intensity of PM2.5 over longer periods of time for fatal coronary heart disease.

In our own work, we assessed the impact of recent exposure to respirable dust (PM) and REC on risk of IHD in another ALPHA-funded study (AFC113-8) by restricting a sub-analysis to person-time that occurred while the worker was still at work or within 6 months of leaving work. The risk for IHD from cumulative exposure was stronger among those who were actively employed, compared to all workers, for both exposures, particularly for respirable dust. More refined models are needed to fully disentangle risk due to intensity, duration and timing of these common mine exposures with respect to risk of heart disease as well as lung diseases in miners.

The eight facilities selected for the DEMS study were all non-metal mines (limestone, potash, salt, or trona) that had very low exposures to known lung carcinogens, such as silica, asbestos and radon. The exposure assessment for diesel exposure in this cohort has been extensively described by Stewart et al.[13] Briefly, respirable elemental carbon (REC) values were estimated from elemental carbon measurements in personal samples collected during the 1998-2001 DEMS surveys. These measurements were used to estimate past annual REC concentrations for each job in each mine, based on mine-specific diesel exhaust-related factors.

The DEMS cohort includes all miners working at one of the mines for at least 1 year after dieselization, either on the surface or underground. Follow-up started between 1947 and 1967 across the eight mines and ends on Jan 1, 1997. The datasets have been made publically available by NIOSH and NCI through the National Center for Health Statistics. NIOSH has recently turned over the continued management of the dataset to NCI. We continue to have support from Debra Silverman (NCI) for this study

The rich historical data on diesel exhaust in DEMS can be used to address this critical research gap in assessing the risk from diesel exhaust for chronic lung and heart disease in miners. We took advantage of our previous ALPHA-funded health studies of a cohort of non-metal miners exposed to diesel exhaust, to examine the timing of exposure-related risk of chronic obstructive pulmonary disease (COPD) as well as IHD. The Diesel Exhaust in Miners Study (DEMS), conducted jointly by the national Institute of Occupational Safety and Health (NIOSH) and the National Cancer Institute (NCI), includes detailed work histories on 12,315 miners in eight non- metal mines.[14] Clarifying the relative contributions of duration, intensity and timing of exposure will improve our understanding of the risks of chronic respiratory diseases in miners

exposed to diesel exhaust, as well as better inform our ability to predict risk of disease and mortality under exposure scenarios that differ in duration, intensity and/or timing. However, we cannot ignore the Healthy Worker Survivor Effect (HWSE) in our effort to focus on the timing of exposure. If miners who experience respiratory or cardiac symptoms in response to diesel exposure tend to leave work or transfer from below ground to a surface job, the healthiest miners will accrue the most exposure – and results based on standard methods will be biased.[15]

We provided compelling evidence that HWSE is an issue in this cohort in our previous ALPHA-funded study of IHD. To see whether the HWSE is operating in DEMS, we treated time to termination as a survival outcome and applied an accelerated failure time model to see if higher exposure to REC shortened the time to leaving work.[16] Table 1 presents hazard ratios (HRs) from the pathway analyses, which in this case are ratios comparing the employment termination rates between higher and lower REC exposures. Results suggest that an increase of 167 ug/m^3 REC, the interquartile range (75%tile -25%tile of the observed REC exposure distribution), was significantly associated with a 60% increase in leaving work (HR=1.6). When restricted to those hired after dieselization and employed for at least 3 or 5 years, the HR was slightly lower. Since higher exposure shortens the time to employment termination, g-methods need to be applied, rather than standard methods, to avoid underestimating any exposure-related health risks. This is particularly true for studying chronic diseases, such as COPD and IHD, because of the prolonged period of symptoms that precede death. Combining g-methods with a new model for identifying the most relevant temporal aspects of exposure will provide the optimal analytic approach for studying, and ultimately reducing, chronic heart and lung disease in miners.

Table 1. Hazard ratio (HR) for termination of active employment, from accelerated failure time models for an interquartile range width (IQRw=167 µg/m3) increase in REC, in the DEMS ever-underground miners.

Cohort subset	N	HR (95% confidence interval)
All ever-underground miners	8307	1.6 (1.5, 1.6)
Hired after dieselization	7750	1.5 (1.4, 1.6)
Hired after dieselization, with ≥3 years tenure	5993	1.3 (1.3, 1.4)
Hired after dieselization, with ≥5 years tenure	4971	1.2 (1.2, 1.3)

Research Objectives

Our specific aims were:

Aim 1: Estimate the relative contributions of intensity, duration and timing of diesel exhaust, measured as respirable elemental carbon (REC), to NMRD and IHD mortality using a distributed lag, non-linear model (DLNM) in the DEMS miners at eight nonmetal mines.

Aim 2: Explore how to incorporate the distributed lag models into the counterfactual framework needed to address HWSE by applying a g-method (parametric g-formula), to handle time-varying confounding by employment status (leaving work) that is affected by past exposure, a signature characteristic of healthy worker survivor bias in longitudinal studies.

Diesel exhaust is composed of vapors, gases, and fine particles emitted by diesel-fueled compression-ignition engines. There have been several studies of COPD in occupational groups exposed to diesel exhaust, as described in a recent review.[3] Most of these studies, however, have lacked specific measures of exposure. Our review of the literature concludes that though the epidemiologic literature on occupational diesel exhaust and COPD is small, it suggests that increasing exposures are associated with increasing risk. Additional research, with more advanced exposure metrics, is needed to more fully elucidate this plausible association.

We were previously funded by ALPHA to examine NMRD mortality in relation to REC and respirable PM in the DEMS cohort. We were also funded by ALPHA to explore the relative contributions of silica exposure intensity, duration and timing, on lung cancer and NMRD, in a cohort study of Californian diatomaceous earth workers[5,17] using a distributed lag nonlinear model (DLNM).[18] In this exploratory proposal we extended our work on those projects by taking fuller advantage of the elaborate historical assessment of diesel exposure[13] in DEMS to disentangle exposure intensity, duration and timing of REC exposure on two separate chronic disease outcomes.

Adjustment for smoking increased the diesel exhaust-related risk of lung cancer in the nested case-control study in this cohort.[19] This surprising result occurred because the more highly exposed miners in DEMS actually smoked less, rather than more, than the lower exposed workers. We had planned to impute smoking based on the smoking data collected for controls in the lung cancer study, however since data were missing on 95% of the cohort, we were unable to impute it with any confidence. Because of the inverse relation between smoking and exposure, however, adjustment is unnecessary: we can be confident that confounding by smoking will always increase the naïve (unadjusted) estimate.

In a recent application of distributed lag nonlinear models (DLNM) to longitudinal occupational data, Gasparrini summarized the temporal relationships between exposure and risk of a health outcome as the 'exposure-lag-response'.[20] A DLNM allows for a nonlinear exposure-response as well as a nonlinear lag- response, and uses the full work history to estimate the exposure-lag-response. By comparing the results for exposure scenarios with the same cumulative exposure but different intensity, duration, and timing, the health effects of these three moving parts of protracted occupational exposure can be disentangled.

3.0 Research Approach:

Cause of death for the DEMS cohort was obtained from NDI-plus from 1979 to 1997 and from death certificates coded by a certified nosologist for deaths occurring prior to 1997. COPD mortality is defined by the underlying cause of death, or one of the first two contributing causes, being listed as COPD or allied conditions (Bronchitis, Chronic Bronchitis, Emphysema, Asthma, Bronchiectasis, and extrinsic allergic alveolitis) in the International Classification of Disease (ICD-9-CM 490-496 and ICD-10 J40-J44, J47). IHD mortality is defined as ICD-9-CM 410-414. Due to sparsity of the outcomes, the analyses were restricted to males of white or unknown race. There were 130 deaths from COPD and 297 from IHD in this group during follow-up.

Aim 1: Estimate the relative contributions of intensity, duration and timing of diesel exhaust, measured as respirable elemental carbon (REC), to NMRD and IHD mortality using a distributed lag, non-linear model (DLNM) in the DEMS miners at eight nonmetal mines.

We examined the impact of exposure intensity (*how much* someone was exposed to) and timing (*when* someone was exposed) on the mortality outcomes of interest in the Diesel Exhaust in Miners Study (DEMS) cohort. We focused on mortality from ischemic heart disease (IHD) and from chronic obstructive pulmonary disease (COPD), a sub-classification of nonmalignant respiratory diseases (NMRD) that excludes pneumoconiosis, rather than from all NMRD. A clearer, more specific outcome definition was more appropriate since different diseases might have different lag-exposure-response relationships. Furthermore, the few excluded NMRD deaths would not have improved power sufficiently to make up for using a less specific outcome.

A matrix of exposure values for REC exposure was created for the analysis. The matrix consisted of one row for each subject-period observation and one column for each of the lags, up to 50 years. The exposure-lag-response function was determined as described in detail by Gasparrini.[20] The cross-basis matrices for REC exposure were entered in the regression model along with other covariates, allowing for both non-linear exposure-response and non-linear lag-response functions.

We had originally intended to use Cox proportional hazards models with a natural spline, and then run a few sensitivity analyses with different assumptions and methods. But on further discussion and reading, we decided that fitting generalized additive logistic models with penalized splines would be a better choice. The generalized additive logistic model was a choice more applicable to our second aim; furthermore, penalized splines generally have better model fit when the true exposure-response is nonlinear.[21] They also offer greater flexibility and can be implemented with a more data-driven model selection process. This meant we did not need to make as many *a priori* assumptions about the model, degrees of freedom, or knots; there was thus no need to run sensitivity analyses about these assumptions.

We therefore applied an extension of Gasparrini's method[22] by fitting penalized distributed lag non-linear models (DLNM) as generalized additive logistic models, with penalized spline functions to model both the exposure-response and lag-response for each outcome. There were 297 deaths due to IHD and 130 due to COPD. Models were adjusted for age, state, calendar time and respirable dust exposure.

We did not run most of the originally planned sensitivity analyses because, as mentioned above, they would have been about the *a priori* assumptions we would have made for the Cox models but did not need for the generalized additive logistic models. However, we did refine the analysis by reconsidering some of the parameters of the penalized splines. In particular, for IHD we repeated the analysis using an even more data-driven approach that did not impose an assumption that the lag 0 effect was 0. Since IHD is known to be related to recent (even same-day) exposure to traffic-related air pollution (of which diesel exhaust is a component),(see, for example, von Klot et al.[23]) and dropping the assumption changed the results, we decided that the refined analysis was a better model for the truth, and used those results in our work on Aim 2.

These analyses were performed in R (version 3.6.1, R Development Core Team, Vienna Austria) using the 'dlnm' package.[24]

Aim 2: Explore how to incorporate the distributed lag models into the counterfactual framework needed to address HWSE by applying a g-method (parametric g-formula), to handle time-varying confounding by employment status (leaving work) that is affected by past exposure, a signature characteristic of healthy worker survivor bias in longitudinal studies.

Preliminary explorations revealed that there was no straightforward way to incorporate the distributed lag models *themselves* directly into the parametric g-formula. The flexibly nonlinear aspect of the distributed lag model simply could not be reconciled with the generalized standardization performed by the fully parametric g-formula method, and recreating the parametric g-formula to accommodate the DLNM would have resulted in a very computationally intensive and time consuming analysis tool. Nevertheless, we were able to build on the IHD results from Aim 1 by applying the exposure-lag-response estimates from the DLNM to developing a modified cumulative exposure metric to be used in the outcome model for the parametric g-formula. We thus found a different way than the one we had initially envisioned to marry the advantages of the DLNM with the adjustment for healthy worker survivor bias in the g-formula analysis.

The parametric g-formula is a method introduced by Robins[15] to control for the healthy worker survivor effect (HWSE), or in longitudinal studies more generally, for time-varying confounders affected by prior exposure. It is analogous to standardizing as a method of confounder control, generalized to be applicable to a time-varying exposure with confounding covariates that evolve over time, affected by prior exposure. Briefly, the method requires fitting parametric models for not just the outcome (in this case, IHD mortality), but also for each time-varying covariate and exposure, conditional on prior exposure and covariate histories as well as baseline covariates.[25,26] Once the coefficients are estimated for each model, counterfactual histories (of covariates and outcomes) are simulated under various exposure interventions using those coefficient estimates. This is achieved by generating a pseudo-sample based on the observed sample population. The estimated coefficients and baseline characteristics are used to generate estimates for covariate values, exposure, and outcome after the first year. This process is repeated for each time point using the previously simulated values for all variables, generating an entire history of covariates and exposures for each miner in the pseudo-sample; this provides an estimate of the IHD risk under no intervention. Then the process is repeated for each exposure intervention, except that at each time, if the simulated exposure exceeds the limit for that intervention, then the exposure is set to the limit rather than to its simulated value, and this intervened-on value is used to predict covariates at the subsequent time points. This process generates estimates of the counterfactual IHD risks under the exposure interventions. The counterfactual risks can then be compared to the risk under no intervention.

Parametric models were fitted for IHD death (treating COPD as a competing risk), exposure (annual average daily REC concentrations), and time-varying covariates for active employment status, cumulative exposure to respirable dust, and job location (with different levels for underground, surface, and inactive). These models were all conditional on prior time-varying covariate and exposure histories and on baseline covariates (age,

calendar year, sex, race, state, and an indicator for whether the subject was hired before or after start of follow-up (dieselization)). Person-time used in these models was restricted to age less than or equal to 80 because there were so few cases after that age and follow-up was long enough to allow estimates of cumulative risk to be interpreted as lifetime risk.

For our analysis, the *key innovation* was to model IHD mortality as a function of exposure in a way that took into account the exposure-lag-response we had estimated in Aim 1. This was done by computing a modified cumulative exposure variable in which REC exposures in each year were weighted before being added together, with weights chosen based on the lag-response curve estimated in Aim 1 (Figure 2). Essentially, the weights decrease approximately exponentially as the lag increases (that is, the longer ago an exposure occurred, the lower the weight). This choice of weights seemed to most closely approximate the lag-response curve while also reproducing the observed outcomes when simulating the scenario under no exposure intervention.

We considered separate interventions on both REC and respirable dust, since both are potentially involved in the disease process; we also considered joint interventions in which both REC and respirable dust are simultaneously limited to certain levels.

4.0 Research Findings and Accomplishments:

Aim 1: Estimate the relative contributions of intensity, duration and timing of diesel exhaust, measured as respirable elemental carbon (REC), to NMRD and IHD mortality using a distributed lag, non-linear model (DLNM) in the DEMS miners at eight nonmetal mines.

Figure 1 below depicts results presented as two 2-dimensional representations of the exposure-lag-response for IHD mortality, with gray shaded areas representing confidence intervals. The cross-section of the original lag-response (Figure 1A), set at an exposure intensity of 300 $\mu g/m^3$, suggests that the RR rises to a maximum at lag 10-20 years. Figure 1B suggests that with a 10 year exposure lag, the RR for IHD mortality rises monotonically to 1.05 per year exposed, after an initial dip in RR below 1.0 for exposures less than 100 $\mu g/m^3$.

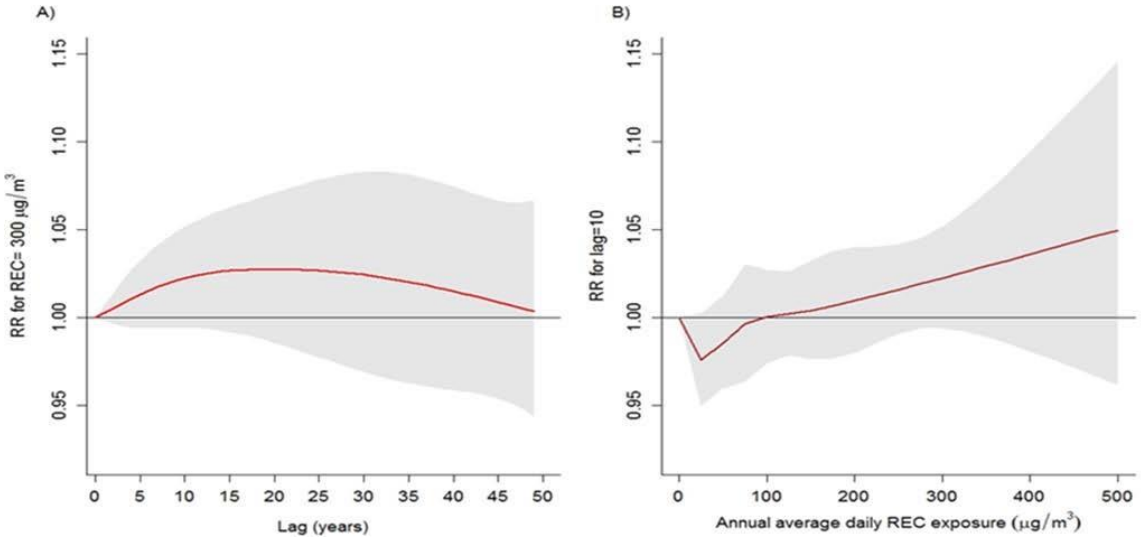

Figure 1: Rate ratios for IHD mortality according to REC exposure, at lags of 0–50 years, among DEMS workers. A) Lag-response at 300 µg/m^3 REC; B) exposure-response of annual exposure, with a 10 year lag, based on penalized distributed-lag nonlinear models.

The shape of the lag-response, however, depended on assuming that the effect at lag 0 was 0. Figure 2 shows the refined lag-response at 400µg/m^3 REC when analysis was repeated *without* forcing the lag 0 effect to be 0.

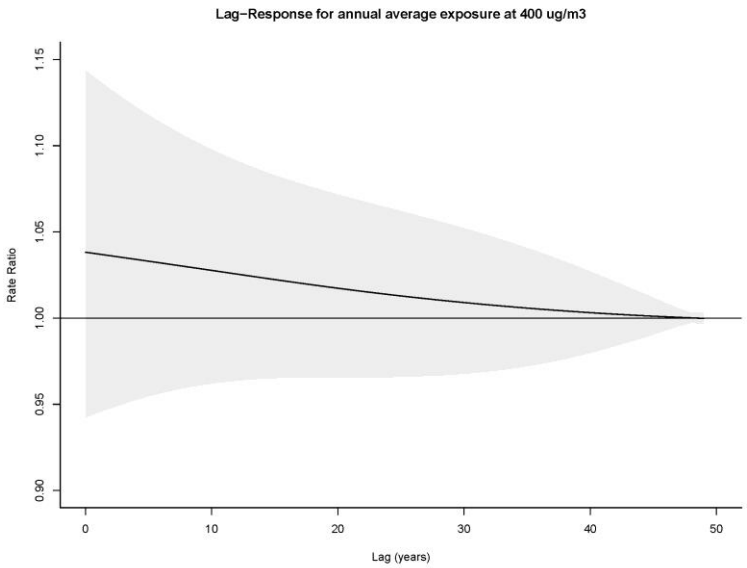

Figure 2: Lag-response curve for IHD, for annual average REC exposure at 400 µg/m^3, when intercept for lag is not forced to be at 0.

Similar to the results for IHD, Figure 3 below depicts two 2-dimensional representations of the exposure-lag-response for COPD mortality (for which a lag 0 effect of 0 makes sense). The RR for REC exposure intensity of 300 µg/m^3 peaks for annual exposure 20 years prior to death (Figure 3A). The RR for REC exposure lag 20 (Figure 3B) rises to 1.8, though the confidence bands are very wide across the exposure range.

These models indicate that exposures occurring during years between lags 10 and 40 (i.e. exposures occurring between 10 and 40 years prior to the year at risk) appear to be more influential with respect to risk of COPD mortality, while in terms of the exposure-response, elevated risk appears to be associated with exposures at intensities of 250 µg/m^3 and higher.

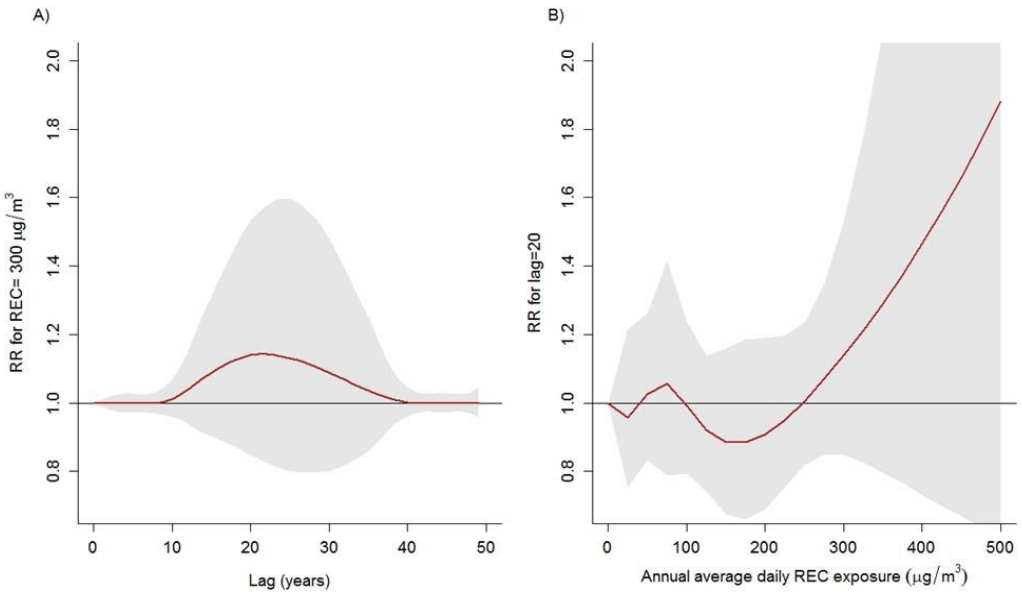

Figure 3: Rate ratios for COPD mortality according to REC exposure, at lags of 0–50 years, among DEMS workers. A) Lag-response at 300 µg/m^3 REC; B) exposure-response of annual exposure, with a 20 year lag, based on penalized distributed-lag nonlinear models.

Although harder to interpret, Figure 4 presents the 3-dimensional figure for COPD and REC that gave rise to the two 2-dimensional Figures 3A and 3B presented above.

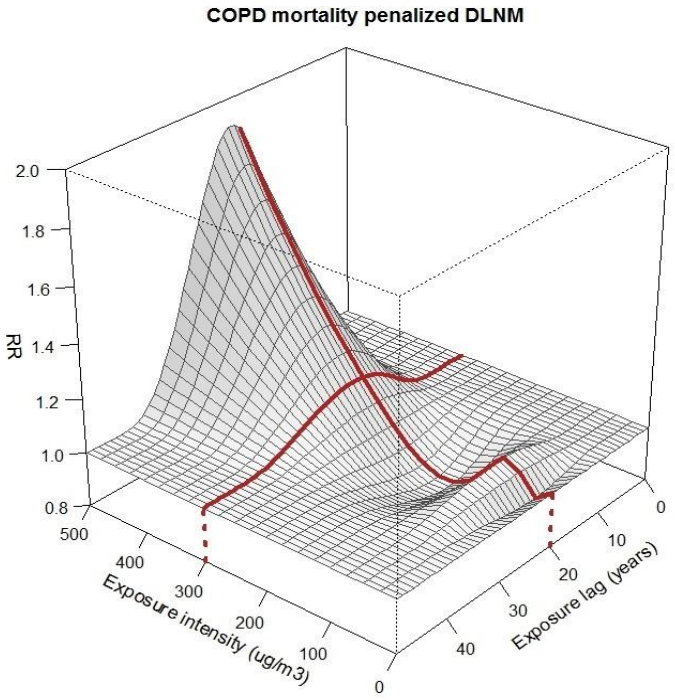

Figure 4: The 3-D graph presents both the exposure-lag and the exposure-response at once. The bolded red curves indicate the two cross-sections (at 300 µg/m^3 intensity and at exposure lag 20) highlighted in Figures 3A and 3B.

Table 2 below summarizes rate ratios (with corresponding 95% CIs) for selected hypothetical exposure scenarios over a possible 50 year window of exposure. An exposure of 300 μg/m^3 for 10 years (for a total of 3000 μg/m^3-years), during lags 10-19, resulted in a RR of 2.09 (95% CI: 0.43, 10.19) for COPD mortality. The same exposure intensity at double the duration (20 years) during lags 0-19 resulted in a RR of 1.44 (95% CI: 0.92, 2.28) for IHD mortality. The third scenario presented corresponds to considering people who have had no occupational diesel exposure in the most recent 19 years (for example, miners who retired 20 years ago), and who were exposed during employment at a high level of intensity (400 μg/m^3) for 10 years, at a somewhat lower level (300 μg/m^3) for 10 years, and then at a much lower level (100 μg/m^3) in their final 10 years of employment (lags 20-29). Although this corresponds to a higher cumulative exposure than the other two scenarios presented, the RR is lower for both outcomes because the longer lag means exposure has less of an effect, as seen in Figure 2.

Table 2: Mortality Rate Ratios (RR) for IHD and COPD mortality in relation to REC exposure.

Exposure Intensity (μg/m^3)	**Timing of Lag (years)**	**Cumulative Exposure (μg/m^3-years)**	**RR (95% CI)**	
			IHD mortality	**COPD mortality**
300	0-19	6000	1.44 (0.92, 2.28)	2.08 (0.42, 10.34)
300	10-19	3000	1.29 (0.93, 1.79)	2.09 (0.43, 10.19)
100 300 400	20-29 30-39 40-49	8000	1.28 (0.22, 7.39)	1.42 (0.07, 29.29)

Confidence intervals for effect estimates for COPD mortality were consistently very wide, as also seen in the figures. Results were underpowered due to the small number of cases; including the small number of NMRD deaths that were not COPD was not sufficient to affect the results.

Aim 2: Explore how to incorporate the distributed lag models into the counterfactual framework needed to address HWSE by applying a g-method (parametric g-formula), to handle time-varying confounding by employment status (leaving work) that is affected by past exposure, a signature characteristic of healthy worker survivor bias in longitudinal studies.

Table 3 shows results from the g-formula analysis. The estimated cumulative risks of IHD mortality at age 80 under various exposure scenarios are compared to the estimated cumulative risk under no intervention. Interventions on REC corresponded to progressively lower risks as the interventions grew more stringent—lower limits for exposure resulted in lower risks—but the confidence intervals still included the null. Results are presented in similar fashion to published results from Neophytou et al.[27] that did not take into account a varying lag-response for REC effects.

Effect estimates for respirable dust were very small, but the joint intervention setting both REC and respirable dust to 0 reached statistical significance. We estimated that preventing all exposure to REC and respirable dust would have resulted in cumulative risk of IHD mortality at

age 80 of 12.7% instead of 14.1% (a risk ratio of 0.9 [95% CI: 0.82, 0.99]). This suggests that although the analyses for interventions only on REC were slightly underpowered, the estimated decreases in risk under interventions to limit exposure to REC are probably meaningful. Furthermore, although the effect estimates presented here are smaller than those reported in Neophytou et al.[27], these estimates give a monotonic exposure-response for REC whereas the published ones did not. This could indicate that using the flexible lag-response in the new g-formula analysis resulted in a more correct model specification than in the earlier analysis,[27] allowing us to do a better job of removing healthy worker survivor bias.

Table 3: Cumulative risk of IHD mortality at age 80 under the natural course and under hypothetical exposure scenarios for REC and respirable dust exposures in US male miners in the Diesel Exhaust in Miners Study hired after the beginning of dieselization (n=10,779), 1946-1997, taking into account the non-linear lag-response.

Intervention	IHD risk, %	Risk Ratio	95% CI	Risk Difference, %	95% CI
Natural course (no intervention)	14.1	1.00		0.0	
REC scenarios					
REC≤106 μg/m^3	14.0	1.00	0.94, 1.1	-0.1	-1.0, 0.8
REC≤25 μg/m^3	13.2	0.94	0.86, 1.0	-0.9	-2.1, 0.4
REC=0 μg/m^3	12.9	0.92	0.83, 1.0	-1.1	-2.5, 0.3
Respirable dust scenarios					
Resp. dust≤1.69 mg/m^3	13.9	0.99	0.97, 1.01	-0.1	-0.4, 0.2
Resp. dust ≤1.15 mg/m^3	13.9	0.99	0.96, 1.02	-0.1	-0.5, 0.2
Resp. dust=0 mg/m^3	13.8	0.98	0.91, 1.05	-0.3	-1.3, 0.7
Joint scenarios for REC and resp. dust					
REC≤106 μg/m^3 & Resp. dust ≤1.69 mg/m^3	13.9	0.99	0.94, 1.03	-0.2	-0.9, 0.5
REC≤25 μg/m^3 & Resp. dust ≤1.15 mg/m^3	13.2	0.94	0.86, 1.02	-0.9	-2.0, 0.3
REC=0 & Resp. dust=0 mg/m^3	12.7	0.90	0.82, 0.99	-1.3	-2.5, -0.1

Abbreviations: CI: confidence interval; IHD: ischemic heart disease; REC: respirable elemental carbon.

5.0 Publication Record and Dissemination Efforts: No publications have yet been produced, as the analyses will be repeated with extended follow-up and better power. However, the

following manuscript from an earlier ALPHA-funded grant (AFC316-54) is now under review at *Environmental Research*:

> Jacqueline M Ferguson, Sadie Costello, Holly Elser, Andreas M. Neophytou, Sally Picciotto, Debra Silverman, Ellen A. Eisen: Chronic Obstructive Pulmonary Disease Mortality: The Diesel Exhaust in Miners Study (DEMS). (Submitted July 1, 2019)

Partial results of related ALPHA-funded projects have also been presented at conferences:

1. Ferguson J, Stewart H, Costello S, Neophytou A, Silverman, Debra, Eisen E: Chronic Obstructive Pulmonary Disease Mortality in the Diesel Exhaust in Miners Study (DEMS), International Society of Exposure Science-International Society for Environmental Epidemiology (ISES-ISEE) 2018 Joint Annual Meeting, August 26th-30th, 2018, Ottawa, Canada
2. Neophytou AM, Costello S, Eisen EA. Occupational diesel exhaust exposure in relation to lung cancer and ischemic heart disease mortality. Occup Environ Med. 2018;75(Suppl. 2)A404. Presented at International Commission on Occupational Health (ICOH) conference, April 29th-May 4th 2018, Dublin, Ireland

6.0 Conclusions and Impact Assessment: Although the substantive results were underpowered, this exploratory work has developed and tested a framework for how to incorporate a complex lag-response into analyses that use the parametric g-formula to adjust for confounding that would otherwise cause healthy worker survivor bias. This is a major step in understanding how to account for the complex ways that timing, intensity, and duration of exposure to diesel exhaust can affect miners' health.

We plan to build on this work in analyses with extended follow-up for our new ALPHA-funded grant (AFC820-15), which will produce substantive results that are more impactful than the results reported here. In particular, we expect that the improved power provided by having so many more cases will allow us to detect an effect of diesel exhaust on COPD, while taking into account the exposure-lag-response functions estimated using the DLNM approach. Quantifying this effect will help guide interventions to reduce exposure to REC among miners, with the eventual goal of preventing COPD and IHD cases caused by diesel exhaust in this population.

7.0 Recommendations for Future Work: We have been funded by ALPHA to extend follow-up on this study. An additional 18 years will provide many more cases, enabling us to apply what we have learned analyzing IHD in this exploratory project to the outcome of COPD. Because the analyses of COPD were so underpowered, we will repeat the Aim 1 analyses with extended follow-up for Aim 3a of the new grant and then use the results to inform our modeling for Aim 2 (Aim 3b of the new grant). The precise aims in the new grant that build on this exploratory work are:

> **Aim 3.** Identify the relative contributions of intensity, duration and timing of REC and dust exposure on risk of COPD mortality to clarify which features of cumulative exposures are driving the risk.

Aim 3a: Fit non-linear distributed lag models for diesel exhaust and dust to characterize the relative weights of different exposure periods.
Aim 3b: For each of the two exposures, use results of Aim 3a to partition cumulative exposures and incorporate the most biologically relevant time windows of exposure into a refined g-formula.

8.0 References:

1. Steenland K, Mannetje A, Boffetta P, Stayner L, Attfield M, Chen J, Dosemeci M, DeKlerk N, Hnizdo E, Koskela R, Checkoway H. 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):773-84.
2. Kreuzer M, Sogl M, Bruske I, Mohner M, Nowak D, Schnelzer M, Walsh L. Silica dust, radon and death from non-malignant respiratory diseases in German uranium miners. *Occup Environ Med* 2013;**70**(12):869-75.
3. Hart JE, Eisen EA, Laden F. Occupational diesel exhaust exposure as a risk factor for chronic obstructive pulmonary disease. *Curr Opin Pulm Med* 2012;**18**(2):151-4.
4. Cherry N, Harris J, McDonald C, Turner S, Taylor TN, Cullinan P. Mortality in a cohort of Staffordshire pottery workers: follow-up to December 2008. *Occup Environ Med* 2013;**70**(3):149-55.
5. Gallagher LG, Park RM, Checkoway H. Extended follow-up of lung cancer and non-malignant respiratory disease mortality among California diatomaceous earth workers. *Occup Environ Med* 2015;**72**(5):360-5.
6. Olsen GW, Andres KL, Johnson RA, Buehrer BD, Holen BM, Morey SZ, Logan PW, Hewett P. Cohort mortality study of roofing granule mine and mill workers. Part II. Epidemiologic analysis, 1945-2004. *J Occup Environ Hyg* 2012;**9**(4):257-68.
7. Cullen MR. Invited commentary: the search for preventable causes of cardiovascular disease--whither work? *Am J Epidemiol* 2009;**169**(12):1422-5.
8. Costello S, Brown DM, Noth EM, Cantley L, Slade MD, Tessier-Sherman B, Hammond SK, Eisen EA, Cullen MR. Incident ischemic heart disease and recent occupational exposure to particulate matter in an aluminum cohort. *J Expo Sci Environ Epidemiol* 2014;**24**(1):82-8.
9. Brown DM, Petersen M, Costello S, Noth EM, Hammond K, Cullen M, Laan M, Eisen E. Occupational Exposure to PM2.5 and Incidence of Ischemic Heart Disease: Longitudinal Targeted Minimum Loss-based Estimation. *Epidemiology* 2015;**26**(6):806-14.
10. Costello S, Picciotto S, Rehkopf DH, Eisen EA. 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):138-44.
11. 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):563-70.
12. Puett RC, Hart JE, Yanosky JD, Paciorek C, Schwartz J, Suh H, Speizer FE, Laden F. Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. *Environ Health Perspect* 2009;**117**(11):1697-701.
13. Stewart PA, Coble JB, Vermeulen R, Schleiff P, Blair A, Lubin J, Attfield M, Silverman DT. The diesel exhaust in miners study: I. Overview of the exposure assessment process. *Ann Occup Hyg* 2010;**54**(7):728-46.

14. Attfield MD, Schleiff PL, Lubin JH, Blair A, Stewart PA, Vermeulen R, Coble JB, Silverman DT. The Diesel Exhaust in Miners study: a cohort mortality study with emphasis on lung cancer. *J Natl Cancer Inst* 2012;**104**(11):869-83.
15. Robins JM. A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect. *Mathematical Modelling* 1986;**7**:1393-1512.
16. Neophytou AM, Picciotto S, Costello S, Eisen EA. Occupational Diesel Exposure, Duration of Employment, and Lung Cancer: An Application of the Parametric G-Formula. *Epidemiology* 2016;**27**(1):21-8.
17. Checkoway H, Heyer NJ, Seixas NS, Welp EA, Demers PA, Hughes JM, Weill H. Dose-response associations of silica with nonmalignant respiratory disease and lung cancer mortality in the diatomaceous earth industry. *Am J Epidemiol* 1997;**145**(8):680-8.
18. Neophytou AM, Picciotto S, Brown DM, Gallagher LE, Checkoway H, Eisen EA, Costello S. Exposure-lag-response in Longitudinal Studies: Application of Distributed Lag Non-linear Models in an Occupational Cohort. *Am J Epidemiol* 2018.
19. Silverman DT, Samanic CM, Lubin JH, Blair AE, Stewart PA, Vermeulen R, Coble JB, Rothman N, Schleiff PL, Travis WD, Ziegler RG, Wacholder S, Attfield MD. The Diesel Exhaust in Miners Study: A Nested Case–Control Study of Lung Cancer and Diesel Exhaust. *Journal of the National Cancer Institute* 2012;**104**(11):855-868.
20. Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. *Stat Med* 2014;**33**(5):881-99.
21. Govindarajulu US, Malloy EJ, Ganguli B, Spiegelman D, Eisen EA. The comparison of alternative smoothing methods for fitting non-linear exposure-response relationships with Cox models in a simulation study. *Int J Biostat* 2009;**5**(1):Article 2.
22. Gasparrini A, Scheipl F, Armstrong B, Kenward MG. A penalized framework for distributed lag non-linear models. *Biometrics* 2017;**73**(3):938-948.
23. von Klot S, Peters A, Aalto P, Bellander T, Berglind N, D'Ippoliti D, Elosua R, Hormann A, Kulmala M, Lanki T, Lowel H, Pekkanen J, Picciotto S, Sunyer J, Forastiere F. Ambient air pollution is associated with increased risk of hospital cardiac readmissions of myocardial infarction survivors in five European cities. *Circulation* 2005;**112**(20):3073-9.
24. Gasparrini A. Distributed Lag Linear and Non-Linear Models in R: The Package dlnm. *Journal of Statistical Software* 2011;**43**(8):1-20.
25. Taubman SL, Robins JM, Mittleman MA, Hernan MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol* 2009;**38**(6):1599-611.
26. Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernan MA. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. *Stat Biosci* 2011;**3**(1):119-143.
27. Neophytou AM, Costello S, Picciotto S, Brown DM, Attfield MD, Blair A, Lubin JH, Stewart PA, Vermeulen R, Silverman DT, Eisen EA. Diesel Exhaust, Respirable Dust, and Ischemic Heart Disease: An Application of the Parametric g-formula. *Epidemiology* 2019;**30**(2):177-185.