SPECIAL REPORT 19

HEALTH EFFECTS INSTITUTE

November 2015

Diesel Emissions and Lung Cancer: An Evaluation of Recent Epidemiological Evidence for Quantitative Risk Assessment

HEI Diesel Epidemiology Panel

Diesel Emissions and Lung Cancer: An Evaluation of Recent Epidemiological Evidence for Quantitative Risk Assessment

HEI Diesel Epidemiology Panel

Special Report 19
Health Effects Institute
Boston, Massachusetts

Trusted Science • Cleaner Air • Better Health

Publishing history: The Web version of this document was posted at *www.healtheffects.org* in November 2015.

Citation for document:

HEI Diesel Epidemiology Panel. 2015. Diesel Emissions and Lung Cancer: An Evaluation of Recent Epidemiological Evidence for Quantitative Risk Assessment. Special Report 19. Boston, MA: Health Effects Institute.

 Cameographics, Belfast, Me., Compositor. Printed by Recycled Paper Printing, Boston, Mass. Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.

Cover paper: made with at least 55% recycled content, of which at least 30% is post-consumer waste; free of acid and elemental chlorine. Text paper: made with 100% post-consumer waste recycled content; acid free; no chlorine used in processing. The book is printed with soy-based inks and is of permanent archival quality.

CONTENTS

Special Report 19

ABOUT HEI

The Health Effects Institute is a nonprofit corporation chartered in 1980 as an independent research organization to provide high-quality, impartial, and relevant science on the effects of air pollution on health. To accomplish its mission, the institute

- Identifies the highest-priority areas for health effects research;
- Competitively funds and oversees research projects;
- Provides intensive independent review of HEI-supported studies and related research;
- Integrates HEI's research results with those of other institutions into broader evaluations; and
- Communicates the results of HEI's research and analyses to public and private decision makers.

HEI typically receives balanced funding from the U.S. Environmental Protection Agency and the worldwide motor vehicle industry. Frequently, other public and private organizations in the United States and around the world also support major projects or research programs. HEI has funded more than 330 research projects in North America, Europe, Asia, and Latin America, the results of which have informed decisions regarding carbon monoxide, air toxics, nitrogen oxides, diesel exhaust, ozone, particulate matter, and other pollutants. These results have appeared in more than 260 comprehensive reports published by HEI, as well as in more than 1000 articles in the peer-reviewed literature.

HEI's independent Board of Directors consists of leaders in science and policy who are committed to fostering the public–private partnership that is central to the organization. The Health Research Committee solicits input from HEI sponsors and other stakeholders and works with scientific staff to develop a Five-Year Strategic Plan, select research projects for funding, and oversee their conduct. The Health Review Committee, which has no role in selecting or overseeing studies, works with staff to evaluate and interpret the results of funded studies and related research. For this report, the HEI Board of Directors appointed a special Diesel Epidemiology Panel to fulfill this role.

All project results and accompanying comments by the Health Review Committee are widely disseminated through HEI's Web site (*www.healtheffects.org*), printed reports, newsletters and other publications, annual conferences, and presentations to legislative bodies and public agencies.

CONTRIBUTORS

HEI DIESEL EPIDEMIOLOGY PANEL

Daniel Krewski, Chair, *Professor and Director of the R. Samuel McLaughlin Centre for Population Health Risk Assessment, University of Ottawa, Canada*

Paul A. Demers, *Director, Occupational Cancer Research Centre, Cancer Care Ontario; Professor, Dalla Lana School of Public Health, University of Toronto, Ontario, Canada.*

David Foster, *Professor Emeritus, Department of Mechanical Engineering, University of Wisconsin–Madison*

Joel Kaufman, *Professor, Environmental and Occupational Health Sciences, Medicine, and Epidemiology; Director, Occupational and Environmental Medicine, University of Washington–Seattle*

Jonathan Levy, *Professor and Associate Chair, Department of Environmental Health, Boston University School of Public Health, Boston, Massachusetts*

Charles Poole, *Associate Professor, Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina–Chapel Hill*

Nancy Reid, *University Professor of Statistics, Canada Research Chair in Statistical Theory and Applications, University of Toronto, Canada*

Martie van Tongeren, *Director of Research, Institute of Occupational Medicine, Edinburgh, United Kingdom*

Susan R. Woskie, *Professor, Department of Work Environment, University of Massachusetts–Lowell*

PEER REVIEWERS

Eileen Abt, *Chemist, Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, College Park, Maryland*

Igor Burstyn, *Associate Professor, Department of Environmental and Occupational Health, School of Public Health, Drexel University, Philadelphia, Pennsylvania*

William Northrop, *Richard and Barbara Nelson Assistant Professor, Department of Mechanical Engineering, University of Minnesota–Minneapolis*

Noah S. Seixas, *Professor, Department of Environmental and Occupational Health Sciences, School of Public Health, University of Washington–Seattle*

Leslie T. Stayner, *Professor, Division of Epidemiology and Biostatistics, and Director, Occupational and Environmental Epidemiology Program, School of Public Health, University of Illinois–Chicago*

Jonathan Samet, *Director, Institute for Global Health, and Distinguished Professor and Flora L. Thornton Chair, Department of Preventive Medicine, University of Southern California Keck School of Medicine, Los Angeles, California*

Thomas Smith,† *Professor of Industrial Hygiene, Emeritus, Harvard T.H. Chan School of Public Health, Boston, Massachusetts*

† As a co-author with Eric Garshick on the 2012 study of trucking industry employees (Garshick et al. 2012a), Thomas Smith did not review Chapter 3.

The Panel also benefited from comments and suggestions by the original investigators of the studies reviewed in this document and by analysts who conducted additional analyses of the data from those studies:

Michael Attfield, *National Institute for Occupational Health and Safety (retired), Morgantown, West Virginia*

Kenny Crump, *Private Consultant, Ruston, Louisiana*

Eric Garshick, *Pulmonary, Allergy, Sleep, and Critical Care Medicine Section, VA Boston Healthcare System, Associate Professor of Medicine, Harvard Medical School, Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts*

Suresh Moolgavkar, *Principal Scientist, Director of the Center for Epidemiology, Biostatistics, and Computational Biology, Exponent, Inc., Bellevue, Washington*

Debra Silverman, *Chief, Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland*

Roel Vermeulen, *Associate Professor of Molecular Epidemiology and Risk Assessment, Institute for Risk Assessment Sciences (IRAS), Utrecht University, the Netherlands*

CONTRIBUTORS

HEI PROJECT STAFF

Katherine Walker, Project Leader, *Senior Scientist*
Aaron Cohen, *Principal Scientist*
Rashid Shaikh, *Director of Science*
Kate Adams, Review Coordinator, *Senior Scientist*
Jacqueline Presedo, *Research Assistant*
Adam Cervenka, *Research Assistant*

UNIVERSITY OF OTTAWA PROJECT STAFF

Nagarajkumar Yenugadhati, *Epidemiologist*
Yuanli Shi, *Research Associate*

HEI PUBLICATIONS STAFF

Carol Moyer, Editorial Coordinator, *Consulting Science Editor*
Hilary Selby Polk, *Managing Editor*
Hope Green, *Publications Associate*
Fred Howe, *Consulting Proofreader*
Ruth Shaw, *Consulting Compositor*

EXECUTIVE SUMMARY

Diesel Emissions and Lung Cancer: An Evaluation of Recent Epidemiological Evidence for Quantitative Risk Assessment

INTRODUCTION AND SCIENTIFIC BACKGROUND

Since their introduction in the early 20th century, diesel engines have become the workhorses in a wide range of industrial settings and forms of transportation. Their power and durability, better fuel efficiency, and lower emissions of some air pollutants (in particular, carbon monoxide) made them attractive in heavy-duty applications such as trucks, buses, construction, farming and mining equipment, locomotives, and shipping in marine and inland waterways. Given these attributes, dependence on diesel engines for all forms of transport, including light-duty passenger vehicles, is strong and appears likely to grow in the foreseeable future.

At the same time, exposures to emissions from diesel engines and their potential impact on human health in both environmental and occupational settings have long been a subject of concern. Over the past several decades, epidemiological and toxicological studies have reported associations between short-term and long-term exposures to diesel exhaust and its components and a range of acute and chronic adverse health effects, including lung cancer. HEI conducted the first of its comprehensive reviews of the scientific literature on diesel exhaust emissions, exposures, and health effects in 1995 (HEI Diesel Working Group 1995). In that review, HEI identified weak increases in lung cancer risk in exposed relative to unexposed workers. Diesel exhaust has also been the subject of numerous scientific reviews by national and international organizations. Most recently, in 2012, the International Agency for Research on Cancer (IARC*) reviewed the body of scientific evidence on the

What This Study Adds

- This report is a careful review by an independent scientific panel of two major epidemiological studies of historical exposures to diesel exhaust, the Diesel Exhaust in Miners Study (DEMS) and the Trucking Industry Particle Study (Truckers) to assess whether these studies could provide the basis for quantitative risk assessment.
- In the Panel's view, both the Truckers study and the DEMS were well-designed and well-conducted studies that each made considerable progress toward addressing a number of the major limitations that had been identified in previous epidemiological studies of diesel exhaust and lung cancer.
- The Panel found that the studies have many strengths, but any effort at quantitative risk assessment will need to acknowledge some key uncertainties and limitations.
- The Panel concluded that both the DEMS and the Truckers study provided results and data that provide a useful basis for quantitative risk assessments of exposures in particular to older diesel engine exhaust.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83467701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

* A list of abbreviations and other terms appears at the end of the Executive Summary.

carcinogenicity of diesel exhaust, and concluded that there was now sufficient evidence in humans and experimental animals to reclassify diesel exhaust from Group 2A (probably carcinogenic to humans) to Group 1 (carcinogenic to humans). As a result, the potential use of these studies for characterization of the exposure–response relationship and for quantitative estimation of lung cancer risk in occupational and general populations has become an issue of considerable interest in the scientific and regulatory communities.

In response to requests from its sponsors, HEI convened a panel in 2013, chaired by Dr. Daniel Krewski of the University of Ottawa (see list of contributors), to review new epidemiological studies of diesel exhaust and lung cancer that had been influential in IARC's determination. The Panel focused on two studies, the Trucking Industry Particle Study (the Truckers study) conducted by Dr. Eric Garshick of the VA Boston Healthcare System and Harvard University and his colleagues (Garshick et al. 2012a), and the Diesel Exhaust in Miners Study (DEMS) conducted by investigators led by Drs. Debra Silverman and Michael Attfield and their colleagues at the National Cancer Institute (NCI) and the National Institute for Occupational Safety and Health (NIOSH), respectively (Attfield et al. 2012; Silverman et al. 2012). The overall charge to the Panel was to make a determination whether or not their data and results could now form the basis for a quantitative characterization of the lung cancer risks associated with diesel exhaust. This report provides the Panel's detailed evaluations of the studies and its conclusions.

THE HEI PANEL APPROACH TO ITS CHARGE

OVERALL PROJECT APPROACH

Beginning in April 2013, the Panel held a series of meetings in person and through webinars and conference calls to discuss the charge to the panel, the Truckers study and DEMS, and the criteria for evaluating them. Through formal applications to NCI and NIOSH, the Panel also obtained the cohort and case–control analytical data sets for DEMS, and after replicating the main results of the study, explored additional questions raised during its evaluation of the studies.

The Panel also took into consideration several published commentaries on both studies as well as the work of two analysts who conducted extensive additional investigations of the DEMS data on behalf of a consortium of firms organized by the Engine Manufacturers Association (Crump et al. 2015; Crump et al. in press; Moolgavkar et al. 2015). The Panel held a public workshop in March 2014 to hear presentations from the original investigators on their studies, from Drs. Crump and Moolgavkar, and from other scientists with expertise in quantitative risk assessment and risk management.

The Panel prepared a draft report that was sent to external peer reviewers, to the original authors of the Truckers and DEMS studies, and to Drs. Crump and Moolgavkar. The report's major findings were presented at the HEI Annual Conference in Philadelphia in May, 2015. The report was revised in response to the many useful comments received during the review process and at the conference.

EVALUATION OF EPIDEMIOLOGICAL STUDIES FOR USE IN QUANTITATIVE RISK ASSESSMENT

Quantitative risk assessments estimate the magnitude of the health burden caused by risk factors to which human populations are exposed. The paradigm for conducting a quantitative risk assessment has long been described in terms of four components: hazard identification; exposure–response assessment; exposure assessment, and risk characterization (National Research Council 1983). The IARC decision having identified a hazard, the Panel focused on the second component and assessed the utility of the Truckers study and the DEMS for quantitative characterization of the exposure–response relationship between diesel exhaust and lung cancer. However, no one set of criteria has been agreed upon to definitively identify studies that provide data of sufficient accuracy, precision, and relevance to be useful for quantitative risk assessment. Instead, this decision remains at the intersection of basic principles of sound epidemiological study design and analysis, of the scientific issues presented by individual studies, and of the needs of risk managers who must ultimately weigh the scientific evidence with uncertainties and other factors in coming to their decisions.

The HEI Diesel Epidemiology Panel therefore evaluated the Truckers and DEMS studies according to how they: 1) addressed major limitations of earlier epidemiological studies for use in quantitative risk assessment that had been identified by a previous HEI panel in 1999 (HEI Diesel Epidemiology Expert Panel 1999); and 2) embodied the attributes of high quality epidemiological studies that make them appropriate and useful for quantitative risk assessment, systematic review, and meta-analysis.

The HEI Expert Panel convened in 1999 had the same mandate as the current panel; to review the epidemiological

literature available at that time. The 1999 Panel reviewed studies in working populations in the trucking and railroad industries and concluded that the studies had a number of limitations that precluded their use in quantitative risk assessment. These limitations related to the quality and specificity of the exposure assessments for diesel exhaust, the absence of quantitative estimates of exposure that would support the exposure–response characterization, and the lack of adequate data to account quantitatively for individual exposure to possible factors that might confound the diesel exhaust and lung cancer relationship, smoking in particular. HEI recommended that these limitations be addressed in future research.

Many publications over the past 25 years have tried to identify the attributes of well-designed, well-conducted epidemiological studies that make them most reliable and useful for quantitative risk assessments. While individual recommendations may differ in details, they share common goals, some overlapping with the research needs identified by the 1999 Panel, which helped to guide the current Panel's evaluation of the details of each of the studies. These included several factors to be considered in the strength and appropriateness of: the study design; the analytical approach to the data and reporting of results; the quality of outcome assessments and follow up; the exposure assessment including the appropriate marker for, and estimates of exposure; the exposure–response assessment; control for confounding factors in both design and analysis; and sensitivity and uncertainty analyses that test the robustness of findings to major assumptions.

EVALUATION OF THE TRUCKERS STUDY

SUMMARY OF THE STUDY

The Truckers study by Garshick and colleagues (2012a) examined the risk of lung cancer in relation to quantitative estimates of personal exposure to submicron elemental carbon (SEC) in a large cohort (31,135) of workers employed in trucking facilities geographically distributed across the United States. This study was the culmination of decades of work investigating a number of health outcomes in association with employment in the trucking industry. Several peer-reviewed publications led up to this study, laying the groundwork for the retrospective reconstruction of individual-level SEC exposure estimates (for the period 1971 to 2000) and the subsequent epidemiological analyses (Davis et al. 2006, 2007, 2009, 2011; Garshick et al. 2008; Jain et al. 2006; Laden et al. 2007, Sheesley et al. 2008, 2009; Smith et al. 2006). Individual-level data on smoking were not available and therefore were not adjusted for in this study. Garshick and colleagues (2012a) found weak associations and evidence of trends in hazard ratios for cumulative SEC, lagged 5 and 10 years, and lung cancer in the cohort excluding mechanics; those associations and trends were strengthened when adjusted for duration of employment, a proxy for a healthy worker survivor bias.

PANEL EVALUATION

The 2012 Truckers study, with its related publications, was designed to address limitations of previous epidemiological studies of diesel exhaust. Specifically, the investigators chose an appropriate metric for diesel exhaust, SEC, a form of elemental carbon (EC). EC generally has been accepted as a reasonable marker for diesel exhaust and is less subject to interference by tobacco smoke and other sources. While gasoline and propane-powered engines also emit EC, the investigators conducted source apportionment analyses in selected terminals that identified diesel engines as a primary source of the SEC measured. The Panel found the investigators' retrospective exposure assessment to be conceptually and statistically sound, relying as it did on a statistically-designed exposure monitoring survey in U.S. trucking terminals, detailed job history and work practice records, and a creative, state-of-the-art structural equation modeling approach. The Truckers study provided estimates of job-specific SEC exposures; using regional coefficient of haze measurements, a reasonable surrogate for particulate EC, they also estimated the historical trends in those exposures. The investigators were able to validate some components of their exposure model, and they tested the sensitivity of their model estimates to some key assumptions. Finally, the conduct of the exposure assessment was independent of knowledge about outcome status, which removed one potential source of differential bias.

The Truckers study embodied other attributes of well-designed and well-conducted epidemiological studies that also make them more useful for quantitative risk assessment. The study was the largest of its kind in this occupation and was geographically representative of the United States. The use of Cox proportional hazards regression to evaluate associations between exposures to SEC and lung cancer was appropriate. The investigators also fit penalized splines in regressions using the continuous SEC exposures and lung cancer to explore the potential for nonlinearities in the exposure–response relationship. They explored the sensitivity of their results to the exclusion of workers in the mechanics job category, a category where there was evidence of greater uncertainty in the exposure estimates. They made the decision to address the suggestions of healthy worker survivor bias that they had observed in their data and did so by adjusting for duration of employment.

The Panel's overall assessment is that the Truckers study can support the development of quantitative risk assessments of diesel exhaust. However, as in any epidemiological study it has some limitations, with resultant uncertainties, that warrant consideration in its interpretation and application in quantitative risk assessments for diesel exhaust.

A major challenge in the Truckers study was the reconstruction of historical exposures to SEC. Several important issues that could impact the validity or uncertainty associated with the retrospective exposure assessment include: the use of the time trends in the coefficient of haze from only one area of the country (New Jersey) was assumed to represent time trends for all the other U.S. trucking terminals in the study; there were no coefficient of haze data prior to 1971 so prior exposures were assumed to be equal to the 1971 levels; SEC was assumed to represent diesel for all workers even though for exposures on or near roads, the mixture of diesel- and gasoline-engine–related ambient EC varies according to the mixture of vehicles (diesel or gasoline) traveling. The Panel agreed that these are potentially important sources of uncertainty in the exposure estimates and therefore could impact the exposure–response relationships that might be derived from the study. To date, no alternative exposure or sensitivity analyses that examine these assumptions have been conducted on these data. Despite the quality of the retrospective exposure construction in the Truckers study, including the careful efforts to validate interim steps in the process, it is the nature of such enterprises that independent data do not exist with which to assess the accuracy and precision of the final estimates.

The investigators were unable to obtain and adjust for individual-level smoking behaviors, an important confounder for lung cancer; however the Panel did not think that smoking alone could explain the findings for the study and noted that the investigators have pointed the way toward post hoc methods for adjusting for this missing information using job-level smoking data. While the investigators have made a reasonable case for adjusting for healthy worker survivor bias in this cohort, the adjustment using duration of work creates some challenges for interpretation of the results and their comparison to the results of other studies lacking such an adjustment.

EVALUATION OF THE DEMS

SUMMARY OF THE STUDY

The DEMS is a cohort and nested case–control study designed to study associations between retrospective estimates of exposure to diesel exhaust, represented by respirable elemental carbon (REC), and health outcomes in a large (12,315) cohort of mostly white male miners engaged in work in eight underground nonmetal mines in the United States (Attfield et al 2012; Silverman et al. 2012). Five peer-reviewed publications laid out the methods and results of the retrospective exposure analysis that was designed to estimate personal-level REC exposures from 2001 back to the start of diesel equipment use in the mines (1947 to 1967, depending on the mine) (Coble et al. 2010; Stewart et al. 2010, 2012; Vermeulen et al. 2010a,b). The mines were chosen because they involved low exposure to potential lung carcinogens other than diesel exhaust (including radon, silica, asbestos, and nondiesel polycyclic aromatic hydrocarbons [PAHs]), used diesel engines over a long period of time, and had good records of both work history and surrogate measures of exposure to diesel exhaust. The nested case–control study (198 cases, 562 controls) included detailed questionnaires to collect data from subjects or next of kin on other potential risk factors for lung cancer, including smoking and employment in other occupations where exposure to lung carcinogens might have occurred. The results of the cohort and the case–control studies were each explored with multiple sensitivity analyses; their results were broadly consistent with each finding an increasing risk of lung cancer in relation to increasing cumulative exposure to REC, lagged 15 years.

PANEL EVALUATION

Like the Truckers study investigators, DEMS investigators also set out to address limitations of exposure assessments in earlier epidemiological studies. They chose nonmetal mines with records of diesel equipment use and an exposure metric, REC, that is generally accepted as a marker of diesel exhaust. The Panel thought that the DEMS retrospective exposure assessment was logically constructed, was thorough in its collection and assessment of available sources of data, and incorporated state-of-the-art methods to develop quantitative estimates of personal exposures to REC for the full period of the study. To the extent possible, the investigators confirmed or justified the decisions they made at several stages in the development of their models, using independent approaches or data where available.

The Panel thought that the process by which DEMS had been designed, conducted, independently overseen, and peer-reviewed met high standards of scientific research. The study was designed with sufficient statistical power and relevant data on covariates to test the hypothesis of an association between long-term exposure to diesel exhaust in the mines and lung cancer in the cohort of mine workers. The study design and analytical approach both included strategies for collecting data on and controlling for potential occupational exposures (i.e., low levels of occupational

carcinogens such as radon, PAHs, silica, asbestos, and respirable dust) and other confounding factors for lung cancer, in particular smoking. Ascertainment of health outcomes was of high quality and conducted independently of the exposure assessment. The statistical analyses followed a logical and standard progression beginning with the estimation of standardized mortality ratios and followed by Cox proportional hazards modeling using both categorical and continuous exposures to REC in the cohort and in the nested case–control study. The DEMS investigators also conducted numerous informative analyses of the sensitivity of their findings to alternative assumptions about exposure metrics, to alternative approaches to modeling relationships between diesel exhaust exposure and lung cancer, and to adjusting for confounding factors. The investigators also made their data and analytical information available through a public process, allowing for further analyses by other groups.

The fundamental associations between estimated exposure to REC and lung cancer were replicable by and robust to numerous investigations — by both the HEI Panel and by other analysts — of alternative statistical modeling approaches, control for confounding factors, and estimates of exposure (Crump et al. 2015; Crump et al. in press; Moolgavkar et al. 2015). The HEI Panel focused on the robustness of the case–control results to alternative adjustments for the two most important potential confounders for lung cancer — smoking and radon. The Panel's analyses affirmed the finding of negative confounding of the REC association by smoking and also found that the REC–lung cancer results were robust to measures of smoking and modeling approaches. However, the Panel noted that the investigators' use of combined work location and smoking variables made the results more challenging to apply in quantitative risk assessments. The Panel's assessment of both the radon data from the mines and the effect of different approaches to adjusting for radon in the statistical models, left Panel members with a high level of confidence that radon is not a major confounder in this study, that adjustment for it is not necessary in this study, and in fact could lead to unintended biases in the results.

As in other retrospective epidemiological studies, a major challenge in DEMS was the reconstruction of historical exposures to REC. Several important questions have been raised about the validity of the retrospective exposure assessment including: the methods for imputing missing measurements; the choice of carbon monoxide (CO) with which to model trends in airborne contaminants in the mines over time; the relationships between horsepower (HP), CO, and REC relative to emissions; and the impacts of temporal changes in diesel engine technology and fuels on the characteristics and the concentrations of diesel exhaust in the mines. The Panel agreed that these are potentially important sources of uncertainty in the exposure estimates and therefore in the exposure–response relationships that might be derived from the study.

Many of these issues have been extensively explored, both by the original investigators in their own sensitivity analyses and by Crump and van Landingham (2012) and by Crump and colleagues (2015 and in press). Crump and colleagues demonstrated sensitivity of the odds ratios and the slope of the exposure–response relationships to alternative exposure estimates and statistical models. The variability in results was considerable in some cases. However, in the Panel's view of the most relevant analyses the variability was smaller, and the results still demonstrated a clear, significant association between REC and lung cancer risk. The associations remained even with the alternative exposure models that did not rely on the HP–CO–REC relationships used in the original investigators' main exposure models.

DISCUSSION

In the Panel's view, both the Truckers and DEMS were well-designed and well-conducted studies and each made considerable progress toward addressing a number of the major limitations that had been identified in previous epidemiological studies of diesel exhaust and lung cancer. These limitations related particularly to the need for metrics more specific to diesel, better models of historical exposures, and ultimately for quantitative estimates of historical exposures to diesel exhaust. They both also demonstrated many of the attributes of high quality epidemiological studies that scientists and regulators value in evidence used to support quantitative risk assessments.

As is true of most occupational epidemiological studies, the findings of these studies are most readily generalizable to workers in other populations exposed to similar concentrations of diesel exhaust, emitted from comparable older engines, over comparable periods of time. However, as part of its charge, the Panel was also asked to consider whether data or results from these studies might also be used to quantify lung cancer risk in populations exposed to diesel exhaust at lower concentrations and with different temporal patterns, such as those experienced by the general population in urban areas worldwide. Although characterization of the exposure–response relationship at low levels of exposure is challenging, the broad and overlapping ranges of exposures to SEC and REC in these studies mitigates to a considerable extent concern about their generalizability to ambient levels. In the Truckers study, the lowest job-specific SEC level was 1.8 $\mu g/m^3$ (representing background levels experienced by clerks, for example); in DEMS, the average facility-specific REC exposure for surface-only workers was 1.7 $\mu g/m^3$. The low end of the range of exposures in each of

the studies is very close to the levels of EC that have been reported in ambient air in the United States (a range of 0.26 to 2.2 μg/m^3 of ambient EC reported from various studies).

RECOMMENDATIONS ON ADDITIONAL ANALYSES OR STUDIES

As part of its charge, the Panel was asked to consider the usefulness of extending or conducting further analyses of existing data sets and for the design of new studies that would provide a stronger basis for risk assessment. The Panel had no further recommendations for major analyses that would need to be done before it could come to its conclusions. Similarly, the Panel thought it would be difficult to identify alternative research designs that would substantially improve on these two studies in the foreseeable future. The major uncertainties in the studies arise from factors largely beyond the control of these investigators — and likely any future investigators — most notably the absence of or only partial historical exposure monitoring and other records necessary to develop more accurate and precise estimates of exposure. Even if a well-designed prospective occupational cohort study were to be initiated today, with detailed personal exposure monitoring for individual workers, it would take decades for results to become available. The Panel however, saw merit in the initiation of exposure-monitoring programs to track trends in exposure to diesel emissions in the future. Data from such programs could be useful for better estimation of future exposure reductions and for evaluating concomitant reductions in human lung cancer risk while avoiding the need for the kinds of historical reconstructions of exposure that have received so much criticism in these and other occupational epidemiological studies.

CONSIDERATIONS FOR FUTURE QUANTITATIVE RISK ASSESSMENTS

The Panel's evaluation of the Truckers study and the DEMS is only one step in a more comprehensive risk assessment process for both characterization of the exposure–response relationship and its application in different risk management settings. The National Research Council risk assessment–risk management paradigm makes it clear that these steps are informed not only by a broad set of evidence, including epidemiological studies, but by the particular decision that must be made and its regulatory context.

Additional considerations in translating the results from these studies to other target populations include generalizability of risk estimates from these predominantly healthy male, Caucasian workers to subpopulations thought to be more susceptible to the effects of exposure to diesel exhaust (e.g., children, elderly people, and those with preexisting comorbidities) and differences in patterns of exposure either at work or to the general population.

Future risk assessments also need to consider major changes in diesel fuels, engines, and aftertreatment technologies that have occurred since these studies were conducted, and the implications those changes have for ambient concentrations and composition of diesel emissions and the risk associated with them. Emissions of PM mass from new technology diesel engines — that is, those equipped with a diesel particulate filter and powered by ultra-low-sulfur diesel fuel — have been reduced by about 99% compared with older engines. The composition of diesel PM from the newer technology has also changed substantially with EC dropping from about 70% by mass in emissions from older engines to as low as 13%–16% in emissions from the newer technology diesel engines. Emissions of PAHs, nitroPAHs, metals and other compounds from newer engines have dropped by about 80% to 99% relative to their levels in 2004 (Khalek et al. 2011, 2015). A study of chronic exposure of rodents to these lower emissions from 2007 technology engines found no evidence of carcinogenicity and few other biological effects (McDonald et al. 2015).

While there remains debate, or uncertainty, about what the 'right' exposure or statistical models are, or the predictions that follow from them, that in and of itself does not mean that these studies and their data are not useful. It is unrealistic to expect that individual results would be universally applicable or that all of the issues could be anticipated for extrapolating the results of the studies to other populations, time periods, and exposure conditions, including different diesel exhaust technologies. Given the basic integrity of the studies, what is important for quantitative risk assessment is that they allow exploration and communication of the nature and magnitude of those uncertainties.

CONCLUSIONS

The HEI Panel found that the epidemiological information that has accrued since the previous HEI panel reported on this issue in 1999 is both relevant and informative. The occupational studies of nonmetal miners and workers in the trucking industry represent useful contributions by investigators who have worked carefully over extended periods of time to recreate historical exposure profiles and to describe exposure–response relationships between diesel exhaust and human lung cancer. Overall, these studies made considerable progress toward addressing the deficiencies that HEI had identified in the utility of earlier epidemiological research studies of diesel exhaust for quantitative risk assessment.

The detailed evaluations of these studies by IARC, the HEI Panel, and other analysts lay the groundwork for a systematic

characterization of the exposure–response relationship and associated uncertainties in a quantitative risk assessment, should one be undertaken. In addition, the Panel has identified the challenges that should be confronted in extrapolating the results from these studies to different populations and time periods, particularly given the rapid changes in diesel technology and its deployment around the world. The Panel concluded that the DEMS and data from both the Truckers study and the DEMS can be usefully applied in quantitative risk assessments. The uncertainties within each study should be considered in any attempts to derive an exposure–response relationship.

REFERENCES

Attfield MD, Schleiff PL, Lubin JH, Blair A, Stewart PA, Vermeulen R, et al. 2012. The diesel exhaust in miners study: A cohort mortality study with emphasis on lung cancer. J Natl Cancer Inst 104:869–883.

Coble JB, Stewart PA, Vermeulen R, Yereb D, Stanevich R, Blair A, et al. 2010. The Diesel Exhaust in Miners Study: II. Exposure monitoring surveys and development of exposure groups. Ann Occup Hyg 54:747–761.

Crump K, Van Landingham C. 2012. Evaluation of an exposure assessment used in epidemiological studies of diesel exhaust and lung cancer in underground mines. Crit Rev Toxicol 42:599–612.

Crump KS, Van Landingham C, McClellan R. In press. Influence of alternative exposure estimates in DEMS miners study: diesel exhaust and lung cancer. Risk Anal.

Crump KS, Van Landingham C, Moolgavkar SH, McClellan R. 2015. Reanalysis of the DEMS nested case-control study of lung cancer and diesel exhaust: suitability for quantitative risk analysis. Risk Anal 35(4):676–700. doi: 10.1111/risa.12371.

Davis ME, Hart JE, Laden F, Garshick E, Smith TJ. 2011. A retrospective assessment of occupational exposure to elemental carbon in the U.S. trucking industry. Environ Health Perspect 119:997–1002.

Davis ME, Laden F, Hart JE, Garshick E, Blicharz A, Smith TJ. 2009. Predicting changes in PM exposure over time at U.S. Trucking terminals using structural equation modeling techniques. J Occup Environ Hyg 6:396–403.

Davis ME, Smith TJ, Laden F, Hart JE, Blicharz AP, Reaser P, et al. 2007. Driver exposure to combustion particles in the U.S. Trucking industry. J Occup Environ Hyg 4:848–854.

Davis ME, Smith TJ, Laden F, Hart JE, Ryan LM, Garshick E. 2006. Modeling particle exposure in U.S. Trucking terminals. Environ Sci Technol 40:4226–4232.

Garshick E, Laden F, Hart JE, Rosner B, Davis ME, Eisen EA, et al. 2008. Lung cancer and vehicle exhaust in trucking industry workers. Environ Health Perspect 116:1327–1332.

Garshick E, Laden F, Hart JE, Davis ME, Eisen EA, Smith TJ. 2012a. Lung cancer and elemental carbon exposure in trucking industry workers. Environ Health Perspect 120:1301–1306.

HEI Diesel Epidemiology Expert Panel. 1999. Diesel Emissions and Lung Cancer: Epidemiology and Quantitative Risk Assessment. Special Report. Cambridge, MA:Health Effects Institute.

HEI Diesel Working Group. 1995. Diesel Exhaust: Critical Analysis of Emissions, Exposure, and Health Effects. Cambridge, MA:Health Effects Institute.

International Agency for Research on Cancer. 2012. IARC: Diesel engine exhaust carcinogenic. Press Release 213. Lyon, France:IARC. Available: *www.iarc.fr/en/media-centre/pr/2012/pdfs/pr213_E.pdf*.

Jain NB, Hart JE, Smith TJ, Garshick E, Laden F. 2006. Smoking behavior in trucking industry workers. Am J Ind Med 49:1013–1020.

Khalek IA, Blanks MG, Merritt PM, Zielinska B. 2015. Regulated and unregulated emissions from modern 2010 emissions-compliant heavy-duty on-highway diesel engines. J Air Waste Manag Assoc 65(8):987–1001.

Khalek IA, Bougher TL, Merritt PM, Zielinska B. 2011. Regulated and unregulated emissions from highway heavy-duty diesel engines complying with U.S. Environmental Protection Agency 2007 emissions standards. J Air Waste Manag Assoc 61:427–442.

Laden F, Hart JE, Smith TJ, Davis ME, Garshick E. 2007. Cause-specific mortality in the unionized U.S. Trucking industry. Environ Health Perspect 115:1192–1196.

McDonald JD, Doyle-Eisele M, Seagrave J, Gigliotti AP, Chow J, Zielinska B, et al. 2015. Part 1.Assessment of carcinogenicity and biologic responses in rats after lifetime inhalation of new-technology diesel exhaust in the ACES bioassay. In: Advanced Collaborative Emissions Study (ACES) Lifetime Cancer and Non-Cancer Assessment in Rats Exposed to New-Technology Diesel Exhaust. Research Report 184. Boston, MA:Health Effects Institute.

Moolgavkar SH, Chang ET, Luebeck G, Lau EC, Watson HN, Crump KS, et al. 2015. Diesel engine exhaust and lung cancer mortality – time-related factors in exposure and risk. Risk Analysis: doi: 10.1111/risa.12315 [Online 13 February 2015]. Available: *http://dx.doi.org/10.1111/ risa.12315.*

National Research Council. 1983. Risk assessment in the federal government: Managing the process. Washington, DC:National Academies Press.

Sheesley RJ, Schauer JJ, Garshick E, Laden F, Smith TJ, Blicharz AP, et al. 2009. Tracking personal exposure to particulate diesel exhaust in a diesel freight terminal using organic tracer analysis. J Expo Sci Environ Epidemiol 19:172–186.

Sheesley RJ, Schauer JJ, Smith TJ, Garshick E, Laden F, Marr LC, et al. 2008. Assessment of diesel particulate matter exposure in the workplace: Freight terminals. J Environ Monit 10:305–314.

Silverman DT, Samanic CM, Lubin JH, Blair AE, Stewart PA, Vermeulen R, et al. 2012. The diesel exhaust in miners study: a nested case-control study of lung cancer and diesel exhaust. J Natl Cancer Inst 104:(11):855–868.

Smith TJ, Davis ME, Reaser P, Natkin J, Hart JE, Laden F, et al. 2006. Overview of particulate exposures in the U.S. trucking industry. J Environ Monit 8:711–720.

Stewart PA, Coble JB, Vermeulen R, Schleiff P, Blair A, Lubin J, et al. 2010. The Diesel Exhaust in Miners Study: I. Overview of the exposure assessment process. Ann Occup Hyg 54:728–746.

Stewart PA, Vermeulen R, Coble JB, Blair A, Schleiff P, Lubin JH, et al. 2012. The Diesel Exhaust in Miners Study: V. Evaluation of the exposure assessment methods. Ann Occup Hyg 56:389–400.

Vermeulen R, Coble JB, Lubin JH, Portengen L, Blair A, Attfield MD, et al. 2010a. The Diesel Exhaust in Miners Study: IV. Estimating historical exposures to diesel exhaust in underground non-metal mining facilities. Ann Occup Hyg 54:774–788.

Vermeulen R, Coble JB, Yereb D, Lubin JH, Blair A, Portengen L, et al. 2010b. The Diesel Exhaust in Miners Study: III. Interrelations between respirable elemental carbon and gaseous and particulate components of diesel exhaust derived from area sampling in underground non-metal mining facilities. Ann Occup Hyg 54:762–773.

ABBREVIATIONS AND OTHER TERMS

CO	carbon monoxide
DEMS	Diesel Exhaust in Miners Study
EC	elemental carbon
HP	horsepower
IARC	International Agency for Research on Cancer
NCI	National Cancer Institute
NIOSH	National Institute for Occupational Safety and Health
PAH	polycyclic aromatic hydrocarbon
REC	respirable elemental carbon
SEC	submicron elemental carbon

Chapter 1

Introduction, Scientific Background, and Overview of the HEI Project

1.0 INTRODUCTION

In 2012, the Health Effects Institute was asked by its sponsors to convene a multidisciplinary expert panel to assess the potential use of recent epidemiological studies of exposure to diesel exhaust and mortality from lung cancer for quantitative risk assessment. Quantitative risk assessment is a process by which scientists use available evidence to estimate the likelihood and severity of adverse health or other outcomes that cannot always be observed directly or with complete certainty, yet which often inform individual or societal risk management decisions. For diesel exhaust or other environmental exposures, it is a prerequisite for identifying the levels of exposure that would be protective of human health in ambient or in occupational settings.

In a 1999 review of the occupational epidemiological studies, a previous HEI expert panel had concluded that the studies available at that time were either not suitable for quantitative risk assessment or needed further analysis (HEI Diesel Epidemiology Expert Panel 1999). The California Environmental Protection Agency had already decided to conduct a quantitative risk assessment for diesel exhaust (Office of Environmental Health Hazard Assessment 1998), but the U.S. Environmental Protection Agency (U.S. EPA*) elected not to, on the basis that the evidence was not sufficient to support quantitative risk assessment (U.S. EPA 2002).

However, in June 2012, the International Agency for Research on Cancer (IARC) reassessed the body of scientific evidence from both toxicological and epidemiological studies of the effects of exposure to diesel exhaust and reclassified diesel exhaust from a Group 2A carcinogen (probably carcinogenic to humans) to a Group 1 carcinogen (carcinogenic to humans) (IARC 2012, 2014). Included in the evidence the IARC reviewed were the two recently published epidemiological studies of historical exposures to diesel exhaust in occupational settings. The first was the most recent analysis by Garshick and colleagues (2012a) of lung cancer mortality in a large cohort of male workers employed in the unionized U.S. trucking industry (hereafter, the Truckers study). The second was the National Cancer Institute–National Institute for Occupational Safety and Health (NCI–NIOSH) study of diesel exhaust exposure in a large cohort of nonmetal miners (hereafter, DEMS [Diesel Exhaust in Miners Study]) (Attfield et al. 2012 and Silverman et al. 2012).

Both studies had sought to address criticisms of earlier epidemiological evidence, including lack of quantitative historical estimates of exposure. The publication of these studies, and the subsequent IARC reclassification of diesel exhaust as a Class 1 known human carcinogen, reignited debates about the extent to which the epidemiological evidence was now sufficient and relevant for use in developing quantitative risk assessments.

Why is this debate important? And how is it informed by the HEI Diesel Epidemiology Project? This chapter begins with a brief scientific background on our reliance on diesel engines, concerns about the health impacts of exposure to their emissions, and the changes in emissions that have been brought about in response to health and other concerns. It next summarizes the series of HEI studies and panels that have addressed the study of diesel engine emissions and health that lay important groundwork for the current report. The chapter concludes with an overview of the project including: the appointment of the Diesel Epidemiology Panel, the charge to the Panel, and a summary of the Panel's overall approach to fulfilling its charge.

1.1 SCIENTIFIC BACKGROUND: A BRIEF HISTORY OF DIESEL ENGINES, EMISSIONS, AND HEALTH

Since their introduction in the early 20th century, diesel engines have over time become the workhorses in a wide range of industrial settings and forms of transportation. Their power and durability, better fuel efficiency, and lower emissions of some air pollutants (in particular, carbon monoxide [CO]) have made them attractive in heavy-duty applications such as trucks, buses, construction, farming and mining equipment, locomotives, and shipping in marine and inland waterways. Because diesel engines are more efficient than gasoline engines, they also emit less carbon dioxide (CO_2, a greenhouse gas) per unit of work, an issue of increasing importance as the total number

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83467701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

* A list of abbreviations and other terms appears at the end of the Special Report.

of vehicle miles traveled increases (HEI Diesel Epidemiology Working Group 2002). However, they have also historically emitted more black carbon (another contributor to climate change) than gasoline engines, an issue that has begun to be addressed as the newest technology engines are entering the market.

Dependence on diesel fuel for all forms of transport is strong and appears likely to grow in the foreseeable future. The 2012 report, *World Energy Outlook*, released by the International Energy Agency, forecast that the worldwide demand for oil from the transport sector, which accounts for over half of global oil production, would continue to grow substantially. Trucks used for freight transport account for 60% of all the diesel fuel consumed globally. The number of diesel-powered light-duty vehicles is also increasing steadily worldwide.

Exposures to emissions from diesel engines and their potential impact on human health in both environmental and occupational settings have long been a subject of concern. Diesel engines have historically emitted high levels of oxides of nitrogen (NO_x), a contributor to ozone formation, fine and ultrafine particulate matter, elemental carbon (EC), and a complex mixture of chemical compounds including aldehydes, aromatic compounds, 1,3-butadiene, polycyclic aromatic hydrocarbons (PAHs), nitroPAHs, and other forms of organic carbon, sulfate, and metals that are of potential concern to public health (U.S. EPA 2002; Zielinska et al. 2010).

Over the past several decades, epidemiological and toxicological studies have reported associations between short-term and long-term exposures to diesel exhaust and its components and a range of acute and chronic adverse health effects, including lung cancer (see, for example, literature cited in HEI 2007; HEI Diesel Working Group 1995; U.S. EPA 2002). Since 1981, comprehensive reviews of the scientific evidence by various state, national, and international organizations have reported mounting evidence supporting a causal relationship between exposure to diesel exhaust and lung cancer (see Table 1.1). Until the most recent review by IARC, however, most had concluded that the evidence was not conclusive about a causal relationship for humans, citing limitations in the epidemiological evidence. Other reviews also have raised questions about the strength of the evidence for an association between diesel exhaust and lung cancer risk (see for example, Hesterberg et al. 2006, 2012a).

As indicated earlier, regulatory agencies had been divided on whether or not to use the then-available epidemiological studies as a basis for developing quantitative cancer risk estimates. In 1998, the California Environmental Protection Agency had developed a quantitative risk factor based on the railroad workers studies (a cancer unit risk factor of 3×10^{-4} indicating the lifetime individual risk of developing lung cancer per $\mu g/m^3$ of exposure to diesel exhaust) (Office of Environmental Health Hazard Assessment 1998). However, in 2002 the U.S. EPA, based on its own assessment and on conclusions from the HEI Diesel Epidemiology Expert Panel's evaluation (1999) regarding the epidemiological evidence available for quantitative risk assessment, decided that the evidence was not sufficient to support quantitative risk assessment at that time (U.S. EPA 2002).

Nonetheless, concerns about other health effects have prompted regulatory agencies in the United States and in other industrialized countries to adopt regulations to control emissions from diesel engines (CONCAWE 2012; HEI 2011). These include a series of regulations by the United States in 2001 to reduce sulfur in diesel fuel, a step that both reduced particulate emissions and paved the way for newer diesel engine and emissions-control technologies that were required to meet new standards for particulate matter (PM) emissions by 2007. By 2010, the engines were required to conform to even stricter standards than in 2007 for emissions of NO_x (U.S. EPA 2001). In addition, starting in 2004 and with full implementation in 2007, all light-duty vehicles (including diesel vehicles) were required to meet stringent PM, NO_x, and hydrocarbon standards (U.S. EPA 2000). Similar efforts have been underway in Europe.

Over the last three decades, emissions from light- and heavy-duty diesel engines have declined dramatically. Compared with 1998 emissions standards, emissions of PM mass from the newer 2007 and 2010 technology diesel engines were reduced by about 99%; CO was similarly reduced by about 97% (Khalek et al. 2011, 2015). The composition of diesel PM has also changed substantially with EC dropping from about 70% by mass to 13%–16%, depending on model year (HEI 2015). Emissions of PAHs, nitroPAHs, metals and other compounds have dropped by about 80% in 2007 engines and 99% in 2010 engines relative to 2004 technology engines (Khalek et al. 2011, 2015).

The effect of these changes and of other regulations are beginning to be reflected in ambient concentrations. For example, the latest Multiple Air Toxics Exposure Study (MATES IV), conducted by the South Coast Air Quality Management District (SCAQMD) in California to evaluate the impact of emissions and other control programs on air toxics concentrations and associated cancer risk in that region, reported that average diesel PM ambient concentrations measured at their 10 monitoring sites dropped from about 3.5–3.7 $\mu g/m^3$, the levels in the 2005 MATES III study, to about 0.9 $\mu g/m^3$ in the 2014 study, an estimated 70% reduction (SCAQMD 2014). Concentrations of the EC component of $PM_{2.5}$ were estimated to drop by about 35%.

Table 1.1 Overview of Assessments of the Causal Evidence for Carcinogenicity of Diesel Exhaust

Organizational Reviews	Animal Data	Human Data	Classification	Quantitative Risk Assessment Conducted?
National Research Council (1981)	Negative	Not convincingly demonstrated	—	—
NIOSH (1988)	"Confirmatory"	"Limited"	"Potential occupational carcinogen"	None
IARC (1989)	"Sufficient" (rats)	"Limited"	"Probably carcinogenic to humans (Group 2A)"	None[a]
World Health Organization (1996)	Supportive (rats)	Suggest "probably carcinogenic"	—	Yes, based on rat data; epidemiologic data considered inadequate
Office of Environmental Health Hazard Assessment, California EPA (1998)	"Demonstrated" carcinogenicity (rats)	"Reasonable and likely"	"Toxic air contaminant"	Yes, based on epidemiologic data in railroad workers [cancer unit risk factor of 3×10^{-4} $(\mu g/m^3)^{-1}$]
U.S. EPA (2002)	"Adequate" (rats)	"Limited"	"Likely human carcinogen"	No; epidemiologic dose–response data inadequate
National Toxicology Program (2011)	"Supporting evidence"	"Limited evidence"	"Reasonably anticipated to be a human carcinogen"	None
IARC (2012, 2014)	"Sufficient evidence"	"Sufficient evidence"	"Carcinogenic to humans (Group 1)"	None[a]

[a] Note: IARC does not conduct quantitative risk assessments at this time for any chemicals.

Reductions in ambient concentrations will reduce exposures to the general population, but the changes in diesel emissions will also have an impact in occupational settings, with the nature and magnitude of exposure depending on the work location and other factors. Jobs involving engines similar to those affected by these regulations, such as those of long-haul truck drivers, are most clearly impacted. In workplaces involving "nonroad" engines — such as mining (either underground or above ground), construction, and agriculture — similar regulations have come into force but at later dates. In metal and nonmetal mines, which are governed by the Mine Safety and Health Administration (MSHA) in the United States, it is only recently that the United States has regulated workplace exposures to diesel particulate matter (DPM) specifically (although regulations for CO and respirable dust would likely have also controlled diesel exhaust exposures). In 2001, the MSHA established interim airborne limits for DPM of 400 $\mu g/m^3$ measured as total carbon in metal and nonmetal mines, a level that was to have been reduced to 160 $\mu g/m^3$ total carbon in 2008. A 2005 rule revised the interim DPM limits to 308 $\mu g/m^3$ expressed as EC, a limit revised to 350 $\mu g/m^3$ in 2006 (Department of Labor, MSHA 2005; Pomroy and Saseen 2008; *www.dieselnet.com/standards/us/ohs.php).*

The replacement of older, more polluting diesel technology has not been immediate throughout the United States and elsewhere, however. Newer engines and aftertreatment technologies are being introduced at varying rates and to varying degrees in different parts of the world depending on the regulatory climate, the business sector,

the availability and affordability of low-sulfur fuels necessary to run the newer technologies, and other factors (International Council on Clean Transportation [ICCT] 2014). In the United States where 95% of heavy-duty trucks are diesel powered, about 33% have 2007 technology or later (Diesel Technology Forum 2014: *www.dieselforum.org/diesel-at-work/delivering-for-america*). Light-duty diesel cars still make up only a very small percentage of personal passenger vehicles sold in the United States — less than 1% in 2012 (*www.eia.gov/forecasts/ aeo/tables _ref.cfm*). In contrast, diesel-powered vehicles account for more than 50% of Europe's light-duty fleet today. Because of the long lifetime of diesel vehicles, the turnover in technology to cleaner diesel engine technology is expected to take one to two decades in the United States and other industrialized countries. Fleet turnover projections by the International Institute for Applied Systems Analysis (2012) suggest that it will take another 15 years, to year 2030, for roughly 95% of light- and heavy-duty vehicles to meet Europe's more stringent emissions standards (i.e., EURO 6 for light-duty vehicles and EURO VI for heavy-duty vehicles [*http://ec.europa.eu/environment/air/transport/road.htm*]). In developing countries, given the longevity of diesel engine technology and the slow rate at which the necessary changes in diesel fuels are being implemented, older diesel technology is likely to dominate for much longer (ICCT 2014).

Given the ongoing transition from older to newer diesel technology engine systems, the question regarding what the most current scientific evidence suggests about the risks of ongoing exposure to diesel exhaust from older engines remains a relevant public health question. A major challenge in going forward is how to incorporate the complex mixture of emissions from new and old diesel technologies and fuel sources in quantitative risk assessments that reflect current and future ambient concentrations of diesel exhaust.

1.2 WHY HEI? A LONG HISTORY OF DIESEL RESEARCH

Since its inception in 1980, the Health Effects Institute has devoted a substantial portion of its research program to the study of diesel emissions, particularly the potential for diesel exhaust constituents to exacerbate or cause adverse health outcomes, including cancer. Its broad-based research program has supported more than 40 research studies to characterize emissions, model exposure and dose, and to evaluate the potential health risks of those exposures (see Appendix A for a list of related publications). Of particular relevance to the current project, HEI has a specific interest in the scientific questions surrounding the use of occupational epidemiological studies to support quantitative risk assessments and in the development of research programs to improve the quality of epidemiological studies for that purpose. Highlights of HEI's work in this area are summarized below:

- **The HEI Diesel Working Group (1995)** conducted a comprehensive review and synthesis of the scientific literature on diesel exhaust emissions, exposures, and associated health effects, with a focus on cancer. Their report concluded that the epidemiological evidence showed weak but consistent increases in the risk of lung cancer for exposed workers compared with unexposed workers (with relative risks in the range of 1.2 to 1.5). However, they cautioned that the absence of concurrent exposure measurements, and insufficient evidence on potential confounding factors, limited the utility of the studies for quantitative estimates of cancer risk.
- **HEI initiated the Diesel Epidemiology Project in 1998.** HEI assembled an expert panel to review six feasibility studies that had been commissioned to provide insight into whether a new retrospective or prospective epidemiological study could provide data that would improve our ability to estimate cancer risks from exposure to diesel exhaust (among them the feasibility study for Dr. Garshick's U.S. Truckers Cohort study). Their work was followed by a Diesel Workshop, "Building a Research Strategy to Improve Risk Assessment," held March 7–9, 1999. This workshop was designed to support a broad discussion about research strategies to improve risk assessment. Topics included: more complete characterization of vehicle emissions, changes in emissions with newer technologies, assessment of diesel exposures in varied occupational and ambient settings, and how best to characterize exposure–response relationships for both cancer and noncancer health effects (HEI 1999).
- **1999 HEI Special Report *Diesel Emissions and Lung Cancer: Epidemiology and Quantitative Risk Assessment.*** HEI appointed an expert panel to evaluate the strengths and weaknesses of the occupational epidemiological studies available at the time for use in quantitative risk assessment. They evaluated a series of studies in railroad workers (Garshick et al. 1987, 1988) and in unionized employees of the trucking industry (Steenland et al. 1990, 1992, 1998). The Panel recommended against use of the railroad worker studies for assessing the quantitative lifetime lung cancer risk from exposure to diesel exhaust. However, the Panel suggested the trucking industry cohorts might be useful if further work were done to quantitatively reconstruct past exposures and to model the exposure–response relationship.

- **HEI formed the Diesel Epidemiology Working Group in the fall of 2000** to review the final reports from the six diesel feasibility studies funded to provide information on potential study populations and on exposure assessment methods (HEI Diesel Epidemiology Working Group 2002). In its evaluation, the Diesel Epidemiology Working Group concluded that full studies of cohorts that had been characterized in the feasibility studies would not generate substantially more accurate exposure–response information, in large part due to limitations of exposure assessment methods. The Working Group's evaluations led to a second workshop, *Workshop to Improve Estimates of Diesel and Other Emissions for Epidemiological Studies*, that would define new research directions for that purpose (HEI 2003).

As researchers continued to search for better markers of current exposures to diesel exhaust, diesel fuels and the engine and emissions control technologies have continued to evolve. HEI joined forces with the Coordinating Research Council in 2005 to develop the Advanced Collaborative Emissions Study (ACES), a cooperative, multiparty effort designed to characterize the mass, composition, and potential toxicity of advanced technology compression ignition, engines, exhaust aftertreatment, and ultra-low-sulfur fuel that have been developed to meet the 2007 and 2010 U.S. EPA emissions standards. The program consisted of three phases. Phase 1 involved extensive emissions characterization of four production-ready heavy heavy-duty diesel engines (i.e., gross vehicle weight higher than 33,000 lb) equipped with control systems designed to meet the 2007 standards for reduced PM (Khalek et al. 2011). Phase 2 involved extensive emissions characterization of a group of diesel engine and control systems intended for production that met the more stringent 2010 standards (including more advanced NO_x controls) (Khalek et al. 2015). As discussed in the previous section, the results from the first two phases indicated substantial reductions in the mass of PM, EC, PAHs and other constituents that have been the hallmark of diesel composition in the past (Khalek et al. 2011, 2015). Phase 3 assessed in rodents the toxicity of exhaust from a 2007 technology engine, including a chronic inhalation bioassay of cancer and noncancer endpoints in rats and a 90-day inhalation study in mice (McDonald et al. 2015). The results of the toxicity testing indicate that lifetime exposure of rats to "new-technology" diesel exhaust from a 2007-compliant engine does not induce tumors in the lungs and has few biological effects (McDonald et al. 2015). These findings differ markedly from those of earlier studies of lifetime exposure to "traditional" diesel exhaust from older engines (see for example, Hesterberg et al. 2005, 2006).

1.3 THE HEI DIESEL EPIDEMIOLOGY PROJECT II

The current HEI Diesel Epidemiology Project builds on HEI's extensive experience in this area. The specific focus of the current work is to conduct a thorough new assessment of the current diesel epidemiological literature on the associations of diesel exhaust with lung cancer and its potential use in the development of quantitative risk assessments.

1.3.1 APPOINTMENT AND CHARGE TO THE DIESEL EPIDEMIOLOGY PANEL

The first step of this project was for the HEI Board of Directors to appoint an Expert Panel to evaluate the studies. The Board appointed Dr. Daniel Krewski to chair the panel because of his high-level committee leadership experience, scientific expertise, and understanding of risk assessment issues. Dr. Krewski is currently a Professor and Director of the R. Samuel McLaughlin Centre for Population Health Risk Assessment at the University of Ottawa, Canada, and a fellow of the Society for Risk Analysis and of the American Statistical Association. The Board also appointed eight additional distinguished scientists to the Panel with substantial expertise in epidemiology, biostatistics, internal combustion engines, industrial hygiene, exposure reconstruction, and risk assessment. Panel members and their expertise are listed under Contributors at the beginning of this report.

The Panel was charged with:

1. Reviewing the findings of the previous Panel's 1999 HEI Special Report, *Diesel Emissions and Lung Cancer*.
2. Reviewing the design, data, and exposure estimates for epidemiological studies that have recently become available and that may form the basis of quantitative risk assessment for diesel exhaust, and analyzing such data as needed.
3. Exploring the question as to whether the data from these new studies enables analyses to extend exposure–response relationships to lower concentration levels, similar to those encountered in everyday, nonoccupational environments.
4. Identifying data gaps and sources of uncertainty.
5. Making recommendations about the usefulness of extending or conducting further analyses of existing data sets.
6. If necessary, making recommendations for the design of new studies that would provide a stronger basis for risk assessment.

These charge questions provided the focus for careful evaluations of the published studies. The Panel's charge did not include a complete reanalysis of the studies such as those HEI conducted on the American Cancer Society Cohort and the Harvard Six Cities Study (HEI 2000).

1.3.2 OVERALL PROJECT APPROACH

1.3.2.1 Internal Panel Deliberations

Beginning in April 2013, the Panel held a series of meetings in person and through webinars and conference calls to discuss the charge, the epidemiological studies that were to be the focus of their review, and the criteria for evaluating them.

The primary focus of the review was on the two recently published studies that had been influential in the IARC deliberations, the Truckers study (Garshick et al. 2012a) and the DEMS (Attfield et al. 2012; Silverman et al. 2012). The investigators for these two studies had undertaken efforts to address a number of shortcomings in earlier epidemiological studies, in particular the development of quantitative estimates of exposure to diesel exhaust. Both studies measured a form of EC by mass, a well-accepted marker for diesel exhaust; the Truckers study measured submicron elemental carbon (SEC, the concentration in $\mu g/m^3$ of EC less than 1 micron in aerodynamic diameter), whereas the DEMS focused on respirable elemental carbon (REC). REC is the fraction of particulate EC that is estimated to reach the alveolar region and is defined by a 50% cut-off diameter of approximately 3.5 μm in aerodynamic diameter. The DEMS also collected some measurements of SEC. Together the two studies span a broad range of exposures which makes them potentially useful for evaluation of exposure–response relationships (see Chapter 2, Figure 2.2).

The Panel also took into consideration several published commentaries on both studies (Boffetta 2012; Borak et al. 2011; Crump and Van Landingham 2012; Gamble et al. 2012; McClellan 2012; Morfeld 2012a,b; Morfeld and Erren 2012; Spallek and Morfeld 2012; Tse and Yu 2012) and the original investigators' responses to them (Garshick et al. 2012b; Silverman and Attfield 2012, 2013; Stewart et al. 2011).

The HEI Panel corresponded with the original investigators of both studies to explore questions about their work. Specifically, the Panel contacted Dr. Silverman and colleagues and Dr. Garshick with follow-up questions about each of their studies and to obtain the "data not shown" that had been referenced in the reports in support of particular analytical decisions but that had not been published. Given the terms of a court order on the release of unpublished data, Dr. Silverman was able to respond to most, but not all, of the Panel's requests (Silverman D, personal communication, 2013). Dr. Garshick was able to respond fully to the Panel's requests (Garshick E, personal communication, 2013).

Over the course of its deliberations, the Panel had the opportunity to consider the work of two analysts who conducted extensive additional analyses of the DEMS data. Working with the DEMS cohort data, Dr. Suresh Moolgavkar† and his colleagues explored the use of an alternative approach to modeling lung cancer risk (i.e., using the three-stage clonal expansion [TSCE] model of carcinogenesis) to take into account time-dependent exposure patterns. They also evaluated the contribution of different mine types to the overall cancer risk (Moolgavkar et al. 2015). Dr. Kenny Crump† and his colleagues examined the impact on cancer risk estimates of alternative estimates of historical exposure to diesel exhaust, to alternative statistical approaches to modeling lung cancer risk and control for radon exposures, and to alternative groupings of the mine workers (Crump et al. 2015). In September 2014, the Panel heard updated presentations on the completed work conducted by Drs. Moolgavkar and Crump and their colleagues. While not the primary focus of the Panel's charge, these analyses provided important insights for the Panel's evaluation of the DEMS. The Panel focused greater attention on the DEMS than on the Truckers study, in part because of these additional analyses but also because the DEMS data were available to the Panel. The Panel was then able to conduct multiple types of sensitivity analyses considering major risk factors for lung cancer (such as smoking and radon).

1.3.2.2 Public Workshop

The HEI Panel held a public workshop Boston, Massachusetts on March 6, 2014 (See Appendix B for the agenda and list of attendees).‡ The purpose of the workshop was to provide the Panel and other interested parties with an opportunity to hear presentations from and ask questions of the original investigators for the DEMS and the Truckers study and of Drs. Moolgavkar, Crump, and Boffetta, who presented their progress on their additional analyses of the DEMS data. Representatives of the U.S. EPA and of NIOSH with responsibilities for quantitative risk assessment and an

† The work of these investigators has been coordinated by the Truck and Engine Manufacturers Association (EMA) on behalf of the American Petroleum Institute (API), European Automobile Manufacturers Association (ACEA), American Trucking Association (ATA), International Organization of Motor Vehicle Manufacturers (OICA), Alliance of Automobile Manufacturers (Alliance), European Research Group on Environment and Health in the Transport Sector (EUGT), Association of Equipment Manufacturers (AEM), Association of American Railroads (AAR), and European Association of Internal Combustion Engine Manufacturers (EUROMOT).

‡ This workshop was originally scheduled for October 26, 2013, but the closure of the U.S. Government resulted in postponing the workshop, which resulted in delays of several months for the project.

author of the reports *Science and Decisions* (National Research Council [NRC]) 2009 and *Environmental Decisions in the Face of Uncertainty* (Institute of Medicine 2013) capped the day with their perspectives on the needs of quantitative risk assessment, for characterization of uncertainty, and the roles of science and uncertainty in decision making. The workshop was attended by over 100 people representing academia, regulatory agencies, nongovernmental organizations, industries, consultants, and law firms.

1.3.2.3 Selected Analyses of the DEMS Data

In the course of its deliberations, the HEI Panel identified several questions about the DEMS case–control and cohort studies that it wanted to explore in greater depth. NCI and NIOSH had each established a process by which investigators could obtain the analytical data sets used to create the published results for the DEMS nested case–control and cohort studies. Each data set required a separate research application from the Panel that included a research proposal and a request to be a signatory to a strict Data Use Agreement that protects the confidentiality of the study subjects. These applications were reviewed both by NCI and NIOSH and by the Institutional Review Boards of the University of Ottawa, where HEI had contracted to have the data analyzed, and of the U.S. EPA (an HEI sponsor). The Data Use Agreement required that the data be received and held in a secure, restricted-access facility; which was available at the University of Ottawa. The Data Use Agreement also required that no attempt be made to link the cohort and case–control data sets either to each other or to any other data sets. Linkage of the cohort and case–control studies, which would be necessary for certain types of analysis (for example, assignment of alternative exposure estimates), could only be done at the National Center for Health Statistics Research Data Center and required a separate research application. The HEI Panel did not undertake analyses that would require such a linkage.

The summary DEMS exposure data used in the DEMS publications are available for download directly from links on the NCI website (see downloadable files on *http://dceg.cancer.gov/research/what-we-study/environment/diesel-exhaust-miners-study-dems*). In response to a Freedom of Information Act request from the EMA, the NCI also provided detailed raw background exposure information used in developing the inputs to the DEMS exposure models. These were subsequently released directly to HEI (Milliard S, personal communication, 2013).

1.3.2.4 External Reviews

A draft version of this report was reviewed by seven independent peer-reviewers who had not been involved in the studies or their original review. The external peer-reviewers were selected based upon their experience with one or more of six relevant areas of expertise: occupational epidemiology, exposure reconstruction, biostatistics, engine and combustion science, mine health, and risk analysis. They were asked to evaluate the Panel's response to the initial charge questions, whether there were other analyses or evaluations of the studies that would give clearer insights, and whether the Panel's conclusions were appropriate. All reviewer comments and concerns were conveyed without attribution to the Panel, which considered them carefully and addressed them as appropriate. The complete list of reviewers can be found in the Contributors page at the beginning of the report. The principal investigators of the original studies — DEMS (Drs. Silverman, Attfield, and Vermeulen) and Truckers (Dr. Garshick) — were also invited to review the report, as were Drs. Crump and Moolgavkar.

1.3.3 ORGANIZATION OF THE REPORT

This report provides the methods and results of the HEI Diesel Epidemiology Panel's critical evaluation of these studies — the design, data collection, exposure assessment and statistical methods, and findings and conclusions. In the context of its charge questions, the Panel puts the work of these studies into a broader perspective and concludes with observations on the strengths and limitations of the use of these specific studies for quantitative assessment of the lung cancer risks associated with exposure to diesel exhaust.

The remainder of the report is organized as follows:

- **Chapter 2** defines more specifically the elements of quantitative risk assessment that are the focus of this report and outlines the Panel's approach to addressing its charge,
- **Chapter 3** provides the Panel's evaluation of the recent Truckers study,
- **Chapter 4** provides the Panel's evaluation of the DEMS,
- **Chapter 5** presents the Panel's conclusions about the DEMS and the Truckers study relative to the charge questions and recommendations for consideration in the application of these studies to quantitative risk assessment.

Chapter 2

Quantitative Risk Assessment and the Role of Epidemiology: The Panel's Approach

2.0 INTRODUCTION

The central charge to the Panel was to evaluate the recent epidemiological studies of diesel exhaust and lung cancer and to explore questions about their potential use in quantitative risk assessment. Quantitative risk assessment, in its broadest definition, involves a comprehensive assessment of the available data on exposure and outcome; human exposure to the agent of interest; modeling of exposure–response relationships; and characterization of risks and uncertainties associated with the exposure of interest (National Research Council [NRC*] 1983). The purpose of this chapter is to define the Panel's focus within the broader set of considerations involved in risk assessment and management, and to define the criteria by which it evaluated the studies for their role within that framework.

Quantitative risk assessment has been developed to estimate the likelihood and severity of outcomes that we cannot always observe directly, but that are factors in decisions we face as individuals and as a society (NRC 2009). Whether or how to build nuclear power plants, to send a space shuttle into orbit, to undertake medical treatment, to set public health priorities to reduce the burden of disease, and where to set limits on exposure to occupational or environmental pollutants are all examples of decisions that depend on scientists to evaluate the data at hand, often imperfect or incomplete, and to provide an assessment of what the data indicate.

In occupational and environmental settings relevant to the consideration of diesel exhaust, the paradigm for quantitative risk assessment has long been described in four basic steps, essentially codified in a report of the NRC (1983) that with some modifications, have continued to define it in the decades since: hazard identification; dose–response assessment; exposure assessment; and risk characterization. Figure 2.1 illustrates how in this paradigm different streams of data from human, animal, and mechanistic toxicity

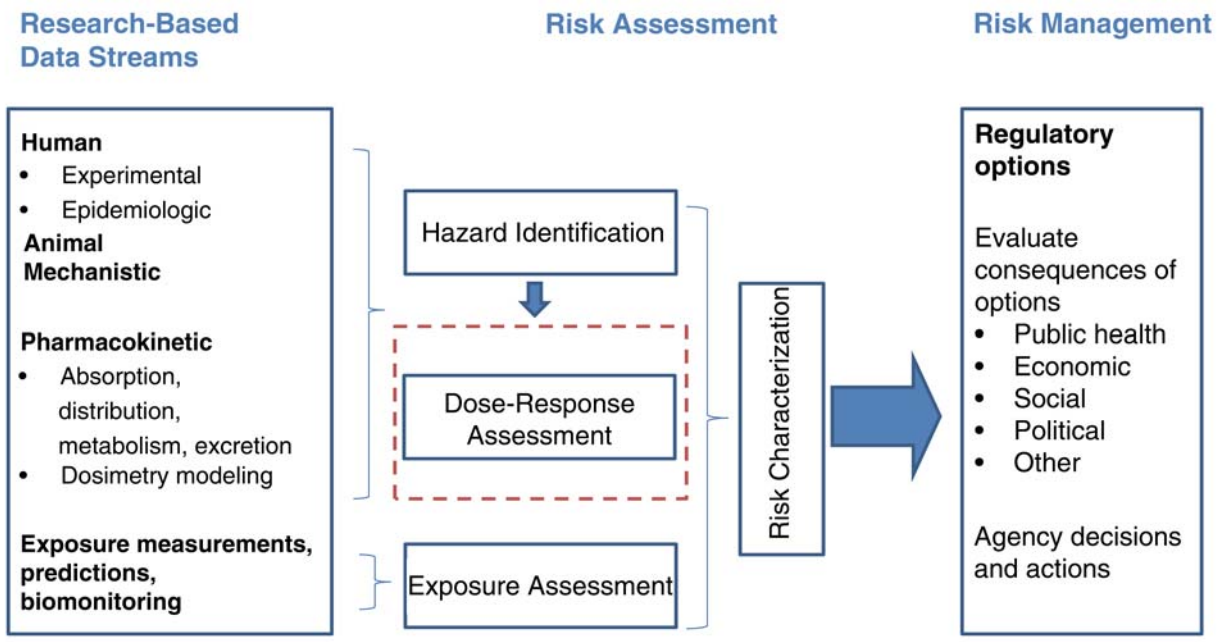

Figure 2.1. The National Research Council risk assessment-risk management paradigm. The HEI Diesel Epidemiology Panel's charge focused on the utility of the DEMS and Truckers studies for Dose–Response Assessment, outlined in dashed red lines. (Adapted from National Research Council 2014 with permission from the National Academies Press.)

* A list of abbreviations and other terms appears at the end of the Special Report.

studies inform evaluations of which adverse health effects may be associated with exposures to a particular pollutant (hazard identification), and of the relationship between the doses in those studies and the probability of adverse effects in the study population or the target population(s) of interest (dose–response assessment). The measurement or estimation of exposure to the pollutant in the target population(s) of interest (exposure assessment) is then coupled with information about dose–response to characterize the nature, likelihood, and magnitude of adverse effects in the exposed population (risk characterization).

Subsequent reviews of this basic paradigm by the National Academy of Sciences and by others have elaborated on a number of key elements in an effort to improve the scientific integrity and utility of risk assessment for risk management decision making. These include recommendations on the appropriate role of judgment in scientific evaluations of risk; how to assess and integrate risk data from different study designs in a transparent way; improvements in the characterization of variability and uncertainty throughout the process of risk assessment; better anticipation of the decisions that risk assessments are intended to inform; and how uncertainties in risk assessment inform environmental decisions (Institute of Medicine 2013; NRC 1993, 1994, 1996, 2009). They provide important additional context for this evaluation.

For diesel exhaust, the hazard identification step has been conducted by others. As discussed in Chapter 1, and summarized in Table 1.1, a number of broad-based scientific literature reviews have been undertaken, culminating in the decision to categorize diesel exhaust as a risk factor for lung cancer in the most recent International Agency for Research on Cancer (IARC) review and reclassification of diesel exhaust as a known human carcinogen (IARC 2012, 2014). The Panel did not reexamine this evidence and operated under the premise that diesel exhaust was a known human carcinogen warranting consideration for future risk assessment applications.

The Panel's charge focused primarily on the evaluation of the DEMS and the Truckers study for use in quantitative characterization of the relationship between exposures to diesel exhaust and the risk of lung cancer (i.e., the exposure–response relationship). The Panel's role was to examine the technical quality and integrity of the exposure and health data generated by the studies, the potential for bias in the results, and the ability of the data to support sensitivity and uncertainty analyses that might be useful to decision makers applying the results in different risk-decision contexts. However, the Panel did not develop or recommend the use of specific quantitative exposure–response functions from either of the studies.

Risk characterization, the ultimate application of exposure–response functions that might be developed from these studies to estimate risks of exposure to diesel exhaust in other target populations, also lies beyond the charge of this HEI Panel. As indicated in Table 2.1, there are a number of potential risk-management activities that quantitative risk assessments of diesel exhaust and lung cancer might eventually inform. The Panel recognized the value in anticipating the additional demands that risk management–risk assessments place on the utilization of epidemiological studies for risk assessment (Fann et al. 2011; NRC 2009; U.S. EPA 2013). Figure 2.2 illustrates that a number of additional modeling steps, assumptions, and other data may be required to adapt or extrapolate the findings from one population and setting to another — in this example from studies of male, largely white workers to settings involving populations for whom the composition, levels and patterns of exposure to diesel exhaust likely differ from those under which the original studies were conducted. Adaptation of any finding based on historical diesel exhaust exposures will need to account for the complex mixture of emissions from new and old diesel technologies and fuel sources reflected in current and future ambient concentrations of diesel exhaust.

2.1 EVALUATING THE ROLE OF EPIDEMIOLOGICAL STUDIES IN QUANTITATIVE RISK ASSESSMENT

Epidemiology, the study of patterns and determinants of health in human populations, has long played an important role in decisions about how to improve public health, ranging from the safety and efficacy of medical interventions to the potential risks associated with exposures to hazards in environmental or occupational settings. The U.S. EPA Integrated Risk Information System (IRIS), the science-based program responsible for developing quantitative exposure–response values for either cancer or noncancer endpoints used by many state, federal, and international agencies for quantitative risk assessment, has recently reaffirmed its preference for reliance on data from well-designed human studies (U.S. EPA 2013).

The reasons for this preference are easy to understand. Epidemiological studies involve the species of interest, if not the actual population group of interest, for many public health decisions, which eliminates many of the challenges of extrapolating results. They often involve large populations and "real-world" occupational or environmental exposure levels and conditions, so extrapolation from the effects of high to low exposures is less of a challenge than in experiments with animals or other systems. Further,

Table 2.1. Potential Risk Management Activities for Diesel Exhaust

Risk Management Activities	Specific Applications
Ambient air quality standards, guidelines or regulations	• U.S. EPA Integrated Science Assessments • Regulatory Impact Assessments to evaluate the risks and benefits of alternate standards or regulations • National Ambient Air Quality Standards • WHO Air Quality Guidelines
Engine fuel and emission standards	• Regulatory Impact Assessments for controls on: ◦ Heavy and light duty onroad vehicles ◦ Nonroad vehicles ◦ Locomotives and marine engines
Occupational guidelines or regulations	• NIOSH Recommended Exposure Limits • OSHA Permissible Exposure Limits • Mine Safety and Health Administration Permissible Exposure Limits • Rules regarding work practices
Burden of disease studies	• Quantitative estimate of the national or global health burden associated with exposures to: ◦ General population ◦ Working populations

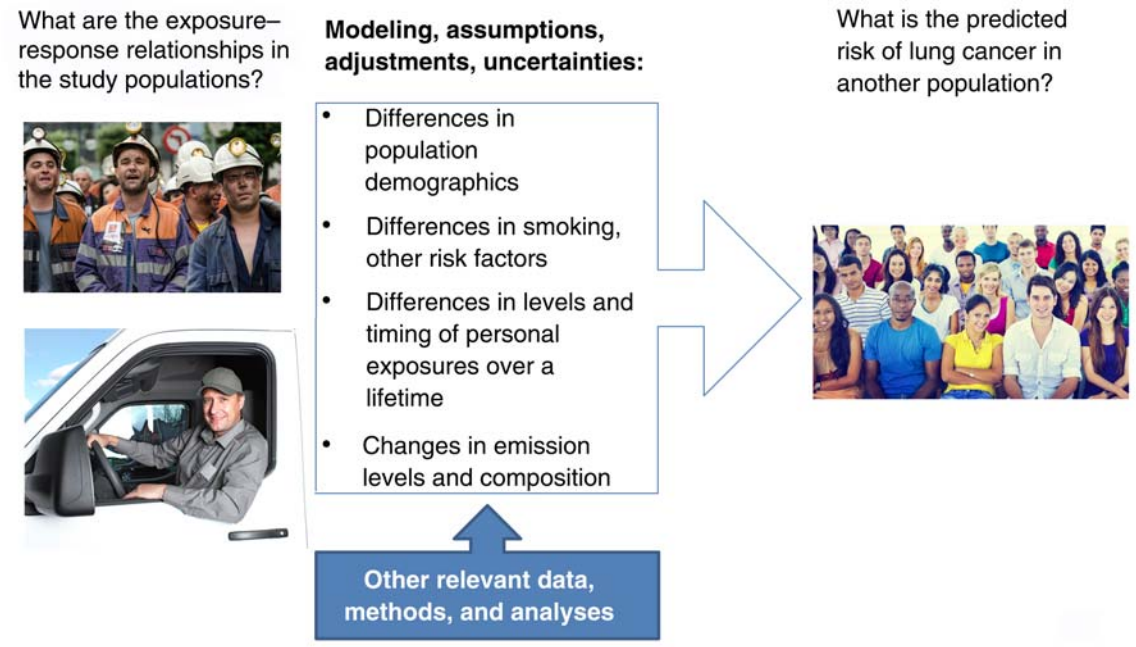

Figure 2.2. Risk Characterization: translating from study results to estimates of risks in different populations. Within the NAS risk assessment paradigm, risk characterization is the final step. It builds on the determination that a hazard exists, on the exposure–response function developed from individual, or multiple studies, and on any necessary adjustments to account for differences in demographics, exposures, and other factors in the study and target populations. Top left: © *matthi/dreamstime.com*. Bottom left: © *kurhan/dreamstime.com*. Right: © *rawpixelimages/dreamstime.com*.

experimental designs that involve assigning human subjects to long-term exposures are commonly either infeasible or unethical.

Despite the general preference for epidemiological studies, it is relatively uncommon that they actually form the basis for quantitative exposure–response assessments. Of the 557 substances for which quantitative risk assessments have been conducted under U.S. EPA's IRIS program, about 80% rely on animal data. Among substances characterized as inhalation carcinogens, 63 quantitative estimates of cancer potency have been developed. Of these, only 14 are based on human evidence of any kind. The DEMS and the Truckers study offer an unusual opportunity to examine the effect of historical exposures to diesel exhaust in large human populations.

2.2 THE HEI DIESEL EPIDEMIOLOGY PANEL'S APPROACH

Given the important role that epidemiological studies have played in our understanding of the potential risks to human health from exposure to toxic substances in the workplace and environment, it is understandable that the studies and the evidence that they provide receive intense scrutiny. To date, no one set of criteria has been agreed upon to definitively identify studies that provide data of sufficient accuracy, precision, and relevance for quantitative risk assessment. Instead, this decision remains at the intersection of basic principles of sound epidemiological study design and analysis, of the scientific issues faced in individual studies, and of the needs of risk managers who must ultimately weigh the scientific evidence with other factors in coming to their decisions.

The HEI Diesel Epidemiology Panel therefore conducted its evaluation of the Truckers study and the DEMS in the context of two sets of criteria: 1) the research needs identified during the 1999 HEI Diesel Epidemiology Panel and subsequent evaluations for addressing the deficiencies in the epidemiological studies available at that time for quantitative risk assessment, and 2) the broad guidance that has emerged in the scientific literature on the design and selection of epidemiological studies for quantitative risk assessment, systematic review, and meta-analysis.

2.2.1 RESEARCH NEEDS IDENTIFIED FOR EPIDEMIOLOGICAL STUDIES ON DIESEL EXHAUST AND LUNG CANCER

Previous HEI panels identified a number of specific limitations in the epidemiological studies available at the time that decreased their utility for quantitative risk assessment (HEI 1999, 2003; HEI Diesel Epidemiology Expert Panel 1999; HEI Diesel Epidemiology Working Group 2002). Table 2.2 summarizes the research needs identified by those panels to address those limitations in future epidemiological studies. The majority of the recommendations were to improve the quality and specificity of the exposure assessment for diesel exhaust, to provide quantitative estimates of exposure that would support the exposure–response characterization, and to quantitatively account for exposure to possible factors that might confound the diesel exhaust and lung cancer relationship, smoking in particular. These recommendations helped to focus this Panel's evaluation of the DEMS and the Truckers study, which were just underway when HEI's 1999 Panel made its recommendations.

2.2.2 PRINCIPLES FOR THE EVALUATION OF EPIDEMIOLOGICAL STUDIES

Observational epidemiological studies, particularly those that are retrospective in design, face many challenges in identifying what might be small increases in risk against an often noisy background of other environmental or personal risk factors. They also need to guard against assigning risk to one agent when other factors may also be responsible to some degree. These challenges of avoiding false negatives on the one hand, and false positives on the other, have led to thoughtful efforts by individual researchers and institutions not only to improve the design and conduct of studies, but also to establish frameworks for deciding when studies are of sufficient quality to be included in comprehensive reviews of the weight of scientific evidence or in meta-analyses for purposes of hazard assessment, exposure–response assessment, or both.

Much has been written over the past 25 years about the characteristics of epidemiological studies that make them best suited for use in quantitative risk assessment (e.g., Fann et al. 2011; Federal Focus 1995; HEI Diesel Epidemiology Expert Panel 1999; Krewski et al. 1990; Loomis et al. 2014; NRC 2014; Schwartz 2002; Stayner et al. 1995; Turner et al. 2010; U.S. EPA 2005; Vlaanderen et al. 2008; World Health Organization [WHO] 2005). In some organizations, these characteristics have been codified in the form of more formal frameworks or checklists, for example, by IARC as part of their systematic reviews of the scientific evidence in support of a causal association between a particular agent and cancer in humans (*http://monographs.iarc.fr/ENG/Preamble/index.php*) and by the Cochrane Collaboration for review of the safety and efficacy of medical interventions (*www.cochrane.org/*). An international collaboration of epidemiologists, methodologists, statisticians, researchers and journal editors, has proposed the STROBE initiative (STrengthening the Reporting of OBservational studies in Epidemiology)

Table 2.2 Research Needs for Quantitative Risk Assessment of Diesel Exhaust[a]

Research Needs for QRA	Specifically
Better measures of exposure	• Measures of diesel constituents. • Of particular importance are the selection and validation of a chemical marker of exposure to the complex mix of diesel exhaust emissions. • Specific biomarkers of diesel exposures, health outcomes, and susceptibility are needed.
Better models of exposure	• Exposure models may include data from personal monitors, area monitors placed where diesel exposure is likely to occur, and current and historical data regarding emission sources. • In any such modeling effort, the effects of environmental tobacco smoke should be removed as completely as possible. • Reliable estimates of past emissions and of factors affecting historical exposures in a range of settings are needed to improve the characterization of uncertainties, both quantitative and qualitative, in historical models of exposures.
Design needs for new studies of exposure–response	• Exposures should be adequately and accurately characterized with respect to magnitude, frequency, and duration, rather than solely by duration of employment. • The exposures considered should be close to levels of regulatory concern, including a range of exposures to provide a base for understanding the relation between exposure and health effects. • Errors and uncertainties in exposure measurements should be quantified where possible. • These should be fully reported to users, and taken into account in both power calculations and exposure–response analyses. • Cigarette smoking is a potent risk factor for lung cancer, and it must be controlled for in any study of risk factors for this disease. • Smoking histories obtained for a cohort study subset that uses a case–control or case–cohort design will strengthen the interpretation of results.

[a] Sources: HEI 1999, 2003; HEI Diesel Epidemiology Expert Panel 1999; HEI Diesel Epidemiology Working Group 2002.

(Vandenbroucke et al. 2007). In the United States, similar approaches are being used by the National Toxicology Program's Office of Health Assessment and Translation — the group within the National Institute of Environmental Health Sciences responsible for conducting evaluations of substances that may be of concern for public health (National Toxicology Program 2013). The National Academy of Sciences, in its recent review of the U.S. EPA's IRIS program recommended that the U.S. EPA consider developing a set of criteria for evaluation of epidemiological and other studies; this review also emphasized the role of formal systematic reviews (NRC 2014).

Despite differences in the particular objectives for these various frameworks, the principles they are based upon have substantial areas of overlap. All have as their common goal the development of systematic and transparent approaches that can help identify well-designed, well-conducted epidemiological studies that provide the most reliable basis for risk assessments or other analyses. The Panel drew on these common principles for its evaluation of the DEMS and the Truckers study including:

- a study design that is clearly documented and scientifically justified to test the study hypotheses, including adequate power and precision, the appropriate study

population, and plans for the evaluation of effect modification and control for confounding variables;

- an analytical approach that is appropriate to the data and hypotheses, including complete reporting of results, both positive and null;
- an approach to health outcome assessment that is complete, reliable, and verifiable and that is blind to assignment of exposure;
- an exposure assessment that includes an appropriate measure of exposure, includes a range of exposures relevant to exposure–response assessment in the populations of interest, provides some insight to the magnitude and potential influence of key uncertainties in exposure assignment, and is blind to identification of health outcomes;
- an exposure–response assessment based on models that fit the data well, reflect a range of plausible alternatives, including where possible consideration of biological relevance; and
- sensitivity and uncertainty analyses that test the robustness of findings to major assumptions in the design and analysis of the study and that characterize the impact of potential sources of bias or uncertainty on the outcomes.

All assessments of epidemiological studies are made substantially easier by efforts to share the data and methods used in the study, both to confirm reproducibility of results and to explore alternative analyses for use in risk assessment and other applications. For example, studies that provide adequate information to guide adjustment of the models or results for differences in population demographics, susceptibility or personal risk factors (e.g., smoking habits, disease state, or socioeconomic status) make them more useful for extrapolation of results to other populations. These characteristics of epidemiological studies, while not necessarily essential to the internal validity of the study, can improve the utility of the study for quantitative risk assessment (Fann et al. 2011).

2.3 SUMMARY

The charge to the HEI Panel was to evaluate the recent epidemiological studies of diesel exhaust and lung cancer, their strengths and weaknesses, and their sensitivities and uncertainties for use in quantitative risk assessment. Specifically the Panel focused on attributes of the studies that were necessary to support development of quantitative exposure–response relationships between exposure to diesel exhaust and lung cancer. In the Panel's view, the existence of sensitivities or uncertainties does not necessarily disqualify studies for use in quantitative risk assessment; what is of value to any quantitative risk assessment derived from these studies is that they allow for a careful accounting of the potential uncertainties in the study data and derived estimates of risk. Uncertainty in various forms is an inherent part of science and a necessary issue for decision makers to confront (Institute of Medicine 2013; Morgan and Henrion 1990; NRC 1983, 2009). The Panel agreed with its earlier counterpart, the 2002 HEI Diesel Epidemiology Working Group, which wrote that the "… judgments as the level of uncertainty to be tolerated are not scientific but rather reflective of the policy-making process. Here, there should be substantial, continuing dialogue between scientists and policy makers."

Chapter 3

Evaluation of Lung Cancer and Elemental Carbon Exposure in the Trucking Industry

3.0 OVERVIEW OF STUDY, METHODS, AND MAIN FINDINGS

The Trucking Industry Particle Study (hereafter, the Truckers study) was designed as a joint effort in exposure assessment and epidemiology in cooperation with the International Brotherhood of Teamsters and four large U.S. trucking companies. The Truckers study comprises an extensive body of work represented by several papers that document the progressive development of the cohort and of the exposure assessment (Davis et al. 2006, 2007, 2009, 2011; Sheesley et al. 2008, 2009; Smith et al. 2006) as well as of the epidemiological and related analyses (Garshick et al. 2008, 2012a; Jain et al. 2006; Laden et al. 2007). The Panel's focus in this report is largely on the latest study by Garshick and colleagues (2012a) because of the greater relevance of its exposure assessment and epidemiological findings for quantitative risk assessment. However, because the earlier studies laid the groundwork for design and analytical choices in the final study, they are also briefly discussed.

Table 3.1 provides an overview of the Truckers study, with an emphasis on the paper by Garshick and colleagues (2012a). The investigators identified 54,319 men and 4,007 women who were employed for at least one day in 1985 by four unionized trucking companies. Women were excluded from subsequent analyses, so are not discussed further here. Laden and colleagues (2007) calculated standardized mortality ratios (SMR*) in the remaining cohort for major causes of death and found a greater than expected number of deaths from ischemic heart disease and lung cancer. The cohort was followed for mortality through the year 2000.

A subsequent paper (Garshick et al. 2008) presented results for a subcohort of 31,135 men 40 years of age or older, who were employed in 1985 by at least one of these four trucking companies, had worked for at least one year in a trucking industry job, and who had follow-up data through 2000. This study used proportional hazards regression to analyze relationships of lung cancer mortality with duration of employment, expressed as hazard ratios (HRs), in different job categories. It found elevated HRs in job categories associated with "regular exposure to freshly emitted vehicle exhaust" — in particular long-haul truck drivers, pick-up and delivery drivers, dockworkers, or those who had some combination of these jobs, but not among clerks whose exposure was assumed to be low.

The most recent of these analyses (Garshick et al. 2012a), the primary focus of the Panel's evaluation, used estimated personal exposures to submicron elemental carbon (SEC) for each member in the 2008 cohort. The exposure assessment built on SEC measurements taken in and around 36 trucking terminals randomly selected to be regionally representative of a total of 139 large terminals operating in the United States in 2000 (Davis et al. 2007; Smith et al. 2006). Using these data, structural equation models were then developed to predict baseline personal SEC exposures (year 2000) for workers employed at the trucking terminals or as drivers (Davis et al. 2006, 2009). Finally, the investigators developed historical estimates of exposure dating back to 1971 by adjusting the baseline SEC exposures for changes in background air pollution levels (represented by coefficient of haze), for some changes in fuel use, and for job-related changes in exposure over time (Davis et al. 2011).

The primary epidemiological analyses involved proportional hazards regressions to assess the relationship between SEC and lung cancer mortality (Garshick et al. 2012a). Multiple exposure metrics were evaluated, including average and cumulative SEC with multiple (0-, 5-, and 10-year) lag times, and with both categorical (exposures divided into quartiles) and continuous measures of exposures. Analyses were conducted on both the full cohort (n = 31,135) and on a cohort excluding 1811 mechanics from the analysis. As in the earlier 2008 study, analyses with and without adjustment for duration of work were performed to address the potential for a healthy worker survivor bias — where individuals who are unhealthy or more susceptible are underrepresented in the workplace, leading to underestimation of risk. Given the absence of individual-level smoking data, smoking was not explicitly controlled for in the analysis.

The investigators found weak associations and evidence of trends in the HRs for cumulative SEC and lung cancer in the full cohort. The findings were stronger when the mechanics were excluded from the analysis (Table 3.1). For both the full cohort and the cohort excluding mechanics, the associations and trends were somewhat stronger after adjustment for duration of work. The exposure–response function using continuous SEC appeared linear and showed a borderline significant association that strengthened with increasing lag. No statistically significant association was seen with average SEC as a continuous covariate.

* A list of abbreviations and other terms appears at the end of the Special Report.

Table 3.1. Overview of the Truckers Studies with a Focus on Garshick et al. 2012a

Truckers Cohort	Exposure Assessment	Outcome Assessment	Analysis	Selected Results
Original Cohort (Laden et al 2007): 58,326 unionized trucking industry workers (93% male) who worked 1 day or more in 1985	**None**	**All cause, and cause-specific mortality**	**Standardized mortality ratios** Expected numbers of all cause and cause-specific deaths calculated from person-years in each race, 10-year age- and calendar period-specific stratum, and national reference rates. SMRs = ratio observed/expected deaths.	**Standardized mortality ratios:** Ischemic heart disease (1133 cases) SMR = 1.41, 95% CI = 1.33–1.49) Lung cancer (769 cases) SMR = 1.04, 95% CI = 0.97–1.12
Garshick 2008, 2012a Cohort: 31,135 male workers ≥ 40 years of age in 1985 with at least one year of work **Cohort description:** Ethnicity: 85% Caucasian, 9% Black, 6% Other/unknown Mean age in 1985: 49.1 years Mean total work: 21.6 years End of follow up: year 2000 Person-years of follow up: ≈106,000 for each quartile of exposure. **8 Job Groups:** *Drivers:* Long-haul, Pick-up & delivery / dockworker (combination), Hostler *Nondrivers:* dockworker, mechanic, clerks, other	**Metric:** Submicron ($PM_{1.0}$) Environmental carbon (SEC) in μg/m^3 **Exposure model:** *Current Exposures:* From 2001–2006, over 4000 personal and area measurements were taken for cross-shift (8–12 hr) SEC at 36 large terminals randomly chosen to be representative of all terminals. Personal SEC exposures were calculated using Structural Equation Modeling as a function of job category, terminal characteristics, and background EC *Historical Exposures:* Ambient SEC levels were modeled based on the ratio of SEC: Coefficient of haze data available from 1971–2000. Comparison of 1988–1989 data to 2001–2006 data was used to calculate job-specific multipliers for 1971–1989 and extrapolated linearly for 1990–2000 *Personal Exposure:* Modeled current exposures were combined with the historical models, including job-specific multipliers, to extrapolate SEC exposures.	**Primary outcome:** Lung cancer, as indicated anywhere on death certificate **Total male deaths**: 4306 (779 lung cancer) **Ascertainment:** National Death Index (NDI-Plus) matched with Social Security Administration files, date of birth, and first, last, and middle names	**Internal cohort analysis:** Proportional hazard regression models, separate baseline hazards based on decade of hire, age in 1-year increments, exposure measured as cumulative SEC and average EC, lags of 0, 5, and 10 years as continuous, and in quartiles. **Adjustments:** age, lung cancer secular trends, calendar year, race, census region of residence, total years of employment (as a time-dependent covariate). There was no control for smoking. **Mechanic Workers:** Separate analyses also performed excluding workers present for ≥ 1 year as a mechanic (n = 1811) due to inconsistency in exposure modeling for mechanics.	**Cohort excluding mechanics, adjusted for duration of work (Garshick et al. 2012a)** **Cumulative SEC (μg/m^3-months)** / **HR (95% CI)** No lag < 530 / 1 530 to < 1061 / 1.25 (0.99–1.60) 1061 to < 2076 / 1.30 (0.99–1.72) ≥ 2076 / 1.24 (0.89–1.71) 5-yr lag < 371 / 1 371 to < 860 / 1.31 (1.01–1.71) 860 to < 1803 / 1.38 (1.02–1.87) ≥ 1803 / 1.48 (1.05–2.10) 10-yr lag < 167 / 1 167 to < 596 / 1.17 (0.88–1.57) 596 to < 1436 / 1.26 (0.90–1.78) ≥ 1436 / 1.41 (0.95–2.11)

3.1 PANEL EVALUATION

Understanding the utility of epidemiological data for quantitative risk analysis requires a full and transparent examination of attributes of the study that affect confidence in both the underlying data and the results of analyses of that data. The Panel conducted a critical evaluation of the Truckers study considering the study design, retrospective exposure assessment, and statistical/analytical methods (including control for potentially confounding exposures) that were outlined in Chapter 2 as desirable attributes of a well-conducted epidemiological study intended for use in quantitative risk assessments. The Panel decided that it would not attempt to obtain and further analyze data from the Truckers study in light of a modest concern about confounders other than smoking in this setting. While smoking is always a candidate confounder of concern for an epidemiological study of lung cancer, no individual smoking data were available in Garshick and colleagues (2012a). The strategy used by Garshick and colleagues to evaluate the potential for confounding due to smoking is discussed in Section 3.2.2 below.

3.2 STUDY DESIGN

The Panel thought the Truckers retrospective occupational cohort study was well designed to assess the mortality risks associated with exposure to diesel and other vehicle exhaust. The study and its methods have been formally peer reviewed at various stages in their development and have been published in leading medical journals. The Truckers study has also been reviewed by the International Agency for Research on Cancer (IARC) as part of that agency's evaluation of the evidence on the relationship between historical diesel exhaust exposure and lung cancer (IARC 2012, 2014).

3.2.1 COHORT SELECTION

The first analyses appropriately focused on a fixed cohort of 54,319 men who made up 93% of the unionized employees who qualified for the cohort (Laden et al. 2007). The rationale for using a fixed 1985 cohort was not explicitly stated, but was presumably based on the availability of computerized records. However, one company had no computerized records before 1993, and only the employees that had been continuously employed at this company between 1985 and 1993 were included in the cohort. This company represented 18.5% of the whole study population. The Panel requested clarification from Dr. Garshick regarding how the person-times for these workers were handled, citing concern that counting those workers' person-times as at risk prior to 1993 could have introduced a negative bias in the study results for analyses compared to the general population. Dr. Garshick, however, confirmed that the "trucking company whose records were only available starting in 1993 did not contribute to the assessment of mortality risk until 1993," which eliminated that concern (Garshick E, personal communication, 2013).

Detailed work history information was obtained for all employees, including date of hire, last date of work, layoff dates, and job title and terminal locations for the employee work histories of the four companies. One company lacked computerized records prior to 1972 (representing 1.5% of all work history time); work before that date was assumed to be the same as the job held in 1972. (As most workers were reported to have remained in the same job category during their career at the same company, this is probably of minor consequence). Job titles and duties were the same across the four companies. Jobs were categorized into eight groups (Laden et al. 2007). Further groupings of drivers and nondrivers were used in some analyses.

In later papers, (Garshick et al. 2008, 2012a), only male workers 40 years of age or older in 1985 and employed for more than one year (n = 31,135) were included in analyses. The rationale provided for this decision was that 96% of lung cancer deaths occurred among workers over 40 years old. The Panel agreed that this was a reasonable approach.

National mortality follow up was performed using the National Death Index or NDI (Laden et al. 2007). Matching criteria included social security number, month and year (± 1) of birth, first name, middle initial, and last name. Lung cancers mentioned anywhere on the death certificate were classified as cases; 734 such cases were identified as the underlying cause of death and 45 appeared elsewhere on the death certificate. Mortality follow up was from 1985 through 2000. No assumptions regarding loss to follow up were stated in the publications, perhaps indicating that cohort members were assumed alive if death was not ascertained. This assumption could result in a small downward bias in death rates in the cohort in comparison with rates in the general population, depending on the quality of records for linkage with the NDI. Workers averaged 42 years of age at entry into the original cohort with about 19 years of employment; however, in the subcohort used in the 2008 and 2012 analyses, workers averaged 49 years of age at entry with almost 22 years of employment.

3.2.2 CONTROL FOR CONFOUNDING FACTORS

In the earlier studies, data on smoking and other potential confounding factors were collected. A smoking survey was mailed to a stratified random sample of 11,986 current

or recently retired (as of 2002) employees of three of the companies (Jain et al. 2006). The questionnaire was modeled on the American Thoracic Society (ATS) questionnaire. The ATS questionnaire is a standardized survey designed to assess chronic respiratory disease in epidemiological studies. The general format includes questions about chronic lung conditions, as well as occupational history, tobacco smoking, and family history of disease. This particular questionnaire also included questions regarding education level and work history prior to working in the trucking industry (Jain et al. 2006).

Survey results were merged with company records to obtain demographic and work history information. After exclusion of bad addresses, the response rate was 40.5%. The distribution of age, sex, job title, region, terminal size, and terminal location were similar among respondents and nonrespondents. Analyses were subsequently restricted to 3362 white males. Long-haul drivers had the highest rates (67%) of respondants who reported ever having smoked, while clerks had the lowest (44%) (Jain et al. 2006). Regional differences were also identified. When stratified by current and ever smokers and by birth cohort, both drivers and nondrivers in the cohort had smoking rates similar to the general population (Laden et al. 2007).

Smoking, the most important potential nonoccupational confounder for lung cancer, was not included in the core analyses of SEC and lung cancer in the latest Truckers study (Garshick et al. 2012a). As is the case with many retrospective occupational cohort studies, there was a lack of information on individual smoking status in this cohort. Using the smoking data described in the previous paragraph (Jain et al. 2006), Garshick and colleagues (2008) found that indirect adjustment for smoking in the job group analysis led to modest reductions in the HRs for long-haul drivers and modest increases for others (ranging from −15% to + 8%).

Garshick and colleagues (2012a) argue that individual-level control for smoking, had it been possible, would not likely have had an appreciable influence on their findings. They point out that not only were the above adjustments in the earlier study small, but that the similarity among the cohort members with respect to socioeconomic status and adjustment in their analyses for age and birth year, also correlates of smoking, were likely to limit confounding by smoking. The Panel agreed with the investigators that smoking was unlikely to account for the observed associations between diesel exhaust and lung cancer, although the absence of formal adjustment for smoking status does contribute some uncertainties in the quantitative exposure–response functions. We discuss the implications of the lack of individual-level smoking data on the applicability of the Truckers study for risk assessment in Section 3.6, Conclusions.

3.3 RETROSPECTIVE EXPOSURE ASSESSMENT

One of the most difficult challenges in retrospective observational epidemiological studies is the characterization of the exposures that members of the cohort were likely to have experienced over the course of their work experience. Historical measurements are often lacking entirely or are incomplete in various ways, measurement technologies have changed, and exposures themselves change over time. For these reasons, most of the epidemiological studies of worker exposure to diesel exhaust have used job exposure matrices to assign workers to exposure categories (e.g., low, medium, and high) or use duration of time worked rather than quantitative measures of exposure to diesel exhaust itself. Such studies still have value for estimating whether qualitatively greater exposure to diesel exhaust is associated with higher risk of disease, but are of more limited use in estimating the risk associated with specific exposure scenarios in quantitative terms.

Consequently, investigators have developed approaches to reconstructing historical levels of exposures over several decades using a broad array of information. One of the best known dose-reconstruction efforts was that conducted to support retrospective analyses of cancer risks among the atomic bomb survivors in Hiroshima and Nagasaki (Preston et al. 2003, 2007; Sawada et al. 1986). In the absence of actual radiation dose measurements for the members of the Life Span cohort established among survivors in those cities, a dose-reconstruction effort was undertaken that required painstaking efforts to document the precise location of individual survivors at the time of detonation, taking into account the orientation of the individual with respect to the epicenter, and the effects of any physical objects such as buildings and clothing that may have had a shielding effect.

Retrospective exposure ascertainment is also widely employed in occupational epidemiology. Job–exposure matrices have been constructed by a number of groups (FINJEM ['Finnish Information System on Occupational Exposure']) to infer typical workplace exposures that may be experienced in a wide variety of occupations (Kauppinen et al. 1998; Lavoue et al. 2012). Alternatively, exposure assessors can utilize exposure measurements grouped into similar exposure groups or job categories combined with individual work histories from employer records or questionnaires to assign personal exposure levels that can be cumulated over a working life. Examples include studies of occupational exposure to silica (Dosemeci et al. 1993), formaldehyde (Stewart and Blair 1994; Stewart et al. 1990), materials in the semiconductor industry (Hammond et al. 1995), ethylene oxide (Hornung et al. 1994), sawdust (Friesen et al. 2006), benzene (Friesen et al. 2012; Glass et al.

2000), and asbestos (Williams et al. 2007). Collectively, this large body of literature demonstrates the ability of retrospective exposure ascertainment methods in epidemiology to identify and quantify cancer risks associated with exposure to agents in occupational settings.

3.3.1 CHOICE OF SEC FOR HISTORICAL EXPOSURES TO DIESEL EXHAUST

The choice of the marker for diesel exhaust exposure was an important early step in this series of studies. The Truckers study investigators chose SEC which corresponds to elemental carbon (EC) measured in PM $\leq$ 1.0 µm in aerodynamic diameter ($PM_{1.0}$). Particles less than 1 µm aerodynamic diameter were selected using a cyclone separator, collected for laboratory analysis on a quartz tissue filter, and analyzed using a thermal optical method (National Institute for Occupational Safety and Health [NIOSH] method 5040; NIOSH 1998).

Garshick and colleagues (2008) chose EC as it is known to be a major component of diesel engine emissions. The intensity of EC in engine emissions is generally substantially greater from heavy-duty diesel trucks than from gasoline-powered cars and propane-powered vehicles. Their measurement studies had indicated diesel exhaust as the primary source of SEC in specific terminal locations (Davis et al. 2006). A source apportionment using chemical mass balance analysis that was conducted at a single site in St. Louis, Missouri, found that the majority of the personal EC measured ($\geq$ 80%) was from diesel exhaust, with spark ignition exhaust and lube-oil impacted exhaust contributing less (Sheesley et al. 2009). However, the source–apportionment analyses of measurement data from their representative sample of 36 U.S. terminals and 1 Mexican terminal also indicate exposures to other mobile sources, represented primarily by organic carbon (Sheesley et al. 2008). Another advantage of EC is that it represents only a very small component of cigarette smoke (0.49%).

3.3.2 EXPOSURE ASSESSMENT SURVEY

A comprehensive exposure assessment survey of the U.S. trucking industry was carried out during which over 4000 SEC samples were collected between 2001 and 2006. Thirty-six different trucking terminals were visited, randomly selected to be regionally representative of the full set of 139 large terminals in operation in 2000 (Davis et al. 2006; Smith et al. 2006). Full-shift (8–12 hr) SEC personal samples were collected from dock workers (who load and unload cargo) and mechanics; SEC area measurements were made in loading docks, offices (to represent clerks and other jobs), and truck cabs to represent hostlers (on-site drivers that move trailers using small specialized tractor units), long-haul drivers, and pick-up and delivery drivers. Ambient background conditions were measured at the periphery upwind of the terminal. SEC area measurements, including 214 loading dock samples, were also collected at 44 smaller terminals (1–2 per trip) within 75 miles of the sampled terminals.

3.3.3 BASELINE EXPOSURE MODEL

Structural equation modeling techniques were used to predict shift-specific personal SEC levels for the on-site terminal workers, including dockworkers, mechanics, clerks, and hostlers. Briefly, structural equation modeling entails simultaneous fitting of multiple nested equations, in this case considering predictors of background exposures, which along with other factors predicts work-area exposures, which in turn are used to predict personal exposures. In the original structural equation modeling conducted by Davis and colleagues (2006), personal job-specific exposures were predicted by smoking status and work area exposures ($R^2 = 0.64$); work area exposures were predicted by terminal-specific characteristics, ventilation, job location in the terminal, and matching background exposures from the area surrounding the terminals ($R^2 = 0.64$). Background exposures were predicted by local weather characteristics, proximity to major roads, industrial land-use characteristics around the terminal, and regional location within the United States ($R^2 = 0.51$). The modeling approach was validated using additional exposure data collected during a series of six repeat site visits conducted after the initial 36 terminal sampling trips (Davis et al. 2009).

Separate exposure models were constructed for drivers who worked off-site delivering and picking up freight (Davis et al. 2007), including local pick-up and delivery drivers and long-haul drivers whose exposures could not be modeled explicitly within the structural equation models because of the dynamic nature of their exposures. Driver SEC exposures were moderately correlated with background EC levels measured at their home terminals, with stronger correlations for local pick-up and delivery drivers ($r = 0.4–0.5$; $P < 0.01$) than for long-haul drivers ($r = 0.2–0.4$; $P < 0.01$). Measured SEC levels inside the driver cabs were also significantly higher when the windows were predicted to be open versus shut ($P < 0.05$). Differences in SEC measurements across driver groups (pick-up and delivery vs. long-haul) and by driver smoking status were not statistically significant.

3.3.4 SPATIAL EXTRAPOLATION

The structural equation modeling and characterization of driver exposures were used to extrapolate SEC exposures to workers at the 103 additional large terminals in the epidemiological cohort that were not part of the original exposure assessment. Using input data from each of the terminals and the estimated coefficients from the modified structural equation modeling, exposures were extrapolated spatially across the entire cohort for the year 2000, and a monthly exposure estimate was calculated for each job and terminal combination. The year 2000 was chosen as the base year for the exposure extrapolation efforts because it represents the final year of follow up in the epidemiological cohort. Smoking was excluded in models used for prediction and extrapolation to terminals not part of the original exposure assessment, due to lack of individual-level smoking in the other terminals and because "its impact on personal exposure to [S]EC is small by comparison with the impact of work area EC levels" (Davis et al. 2006).

3.3.5 JOB GROUP SCALING

Scaling factors were constructed that related measured driver SEC to model-based background predictions. Specifically, ratios of median driver to terminal background SEC were obtained for each driver type. This ratio resulted in a multiplier of 2.1 for long-haul drivers and hostlers, indicating that exposures for these drivers were typically 2.1 times higher than terminal background conditions. For pick-up and delivery drivers, separate multipliers were calculated for warm- and cold-weather conditions to account for the impact of open cab windows in the truck cabs that were not air conditioned. The multiplier for pick-up and delivery driver exposures to background levels was 2.3 (window open) in warm-weather conditions (> 10°C) and 2.0 for colder temperatures (window shut). Office workers were assigned background conditions.

3.3.6 TEMPORAL EXTRAPOLATION OF EXPOSURES

To account for changes in job-related exposure characteristics over time, a comprehensive historical review of work practices in the trucking industry was carried out, including the introduction of diesel-fueled vehicles across job groups and companies. Three historical multipliers were developed to extrapolate baseline exposure model estimates derived for the year 2000 to earlier periods, based on the structure of the exposure model along with historical input data on the model covariates, to predict job-specific SEC exposures (Davis et al. 2011). Each of the three multipliers dealt with different factors that could have influenced historical exposures levels.

The first historical multiplier was developed using SEC exposure measurement data in an earlier trucking industry study by Zaebst and colleagues (1991). The ratio of the model-based predictions to the measured values was used to adjust baseline model-based predictions for changes in work-related conditions over time. The second multiplier was focused on fuel use since forklifts used during the exposure assessment survey (discussed in Section 3.3.2) were powered by propane fuel only, and in some locations diesel forklifts were used in the 1980s and 1990s. Data from Zaebst and colleagues (1991) were used to develop fuel use multipliers by comparing EC concentrations from propane with concentrations related to diesel- and gasoline-powered forklifts.

The third multiplier was developed from the trend in monthly average coefficient of haze levels at 26 locations in New Jersey between 1971 and 2000 to adjust for the effect of changes in background ambient SEC levels on work-related exposures. Coefficient of haze was widely used in the 1960s and 1970s to monitor air pollution and has been shown to be a strong predictor of EC ($R^2 = 0.94$) (Cass et al. 1984; Wolff et al. 1983). It has more recently been used to characterize changes in diesel-related PM exposure conditions over time (Davis et al. 2010; Kirchstetter et al. 2008). Ratios comparing the median annual coefficient of haze value in each year with the estimate for base year 2000 were used to adjust annual background SEC predictions for the period 1971–1999. Job-specific SEC values before 1971 (8% of total person-years) were assigned 1971 exposures because coefficient of haze data were not available to estimate background levels prior to that time.

Table 3.2, taken from Table 2 in Davis and colleagues (2011), shows the SEC estimates by job and time period. Results indicated that estimated SEC exposures were higher for 1981–1990 than for 1991–2000 for all job groups (range of about 23%–57% difference between decades). The biggest change was observed for the mechanics in cold climates (-17.14 μg/m^3 [51% decrease]). The temporal trends in exposure profiles across job categories reflect both changes in work practices and elevated background conditions over time. Past use of diesel-powered forklifts on terminal docks was by far the largest historical multiplier, resulting in the largest impact on job-related estimates of exposure.

Regional differences both within and between the various job groups were observed. Mechanic exposures were significantly higher in the Midwest and Northeast ($P < 0.01$), whereas exposures in the other job groups were comparatively higher in the South and West ($P < 0.01$). The investigators hypothesized that this is likely due to reduced ventilation in colder climates.

In contrast, exposures for pick-up and delivery drivers, which come from outside of the truck cab, are increased

Table 3.2. Summary Statistics of Shift-Level SEC Predictions by Job per Decade (μg/m³).

	1971–1980			1981–1990			1991–2000			Change in Means: 1971–1980 & 1981–1990		Change in Means: 1981–1990 & 1991–2000	
Job Group	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	μg/m³	%	μg/m³	%
Background/ clerks	1.79	1.65	0.74	1.25	1.20	0.40	0.80	0.75	0.31	−0.54	−30.17	−0.45	−36.00
Dockworkers (diesel)	40.80	37.25	17.28	32.06	29.86	12.19	24.73	22.83	9.96	−8.74	−21.42	−7.33	−22.86
Dockworkers (gasoline)[a]	8.20	7.49	3.47	6.44	6.00	2.45	4.97	2.00	4.59	−1.76	−21.46	−1.47	−22.83
Dockworkers (propane)[a]	1.95	1.78	0.83	1.53	1.43	0.58	1.18	1.09	0.48	−0.42	−21.54	−0.35	−22.88
Mechanics (all)	19.66	9.72	22.48	15.23	7.66	16.98	7.64	3.86	9.91	−4.43	−22.53	−7.59	−49.84
Mechanics (warm climate)	7.75	6.33	5.05	6.08	5.07	3.75	3.16	2.56	2.33	−1.67	−21.55	−2.92	−48.03
Mechanics (cold climate)	43.23	37.72	24.86	33.57	29.77	18.27	16.43	13.19	12.79	−9.66	−22.35	−17.14	−51.06
LH drivers/ hostlers	6.40	5.88	2.64	4.46	4.26	1.45	2.21	2.01	1.04	−1.94	−30.31	−2.25	−50.45
P&D drivers (warm)	10.41	9.59	4.16	7.23	6.97	2.25	3.09	2.77	1.64	−3.18	−30.55	−4.14	−57.26
P&D drivers (cold)	4.56	4.15	1.99	3.18	2.95	1.12	1.79	1.64	0.80	−1.38	−30.26	−1.39	−43.71

[a] Dockworker exposure predictions not relevant to all time periods; based on company reported fuel-use profiles.

P&D = pick-up and delivery; LH = long haul.

Adapted from Davis et al. 2011, Table 2.

during warmer temperatures and in warmer climates because these trucks are often not equipped with air conditioning and cab windows are more likely to be open when it is warm outside. This effect was not as evident for long-haul drivers, whose truck cabs were equipped with air conditioning throughout the study period. Because of the greater contribution of background conditions to exposures in pick-up and delivery drivers, pick-up and delivery exposure levels were higher than long-haul exposures in the past, consistent with higher background exposures.

3.3.7 STRENGTHS OF THE EXPOSURE ASSESSMENT APPROACH

The Panel identified several strengths of the exposure assessment. The HEI Panel agreed that SEC was an appropriate marker for exposure to diesel exhaust. Although diesel exhaust is a complex mixture that creates challenging questions about the choice of marker, a general consensus has emerged that EC is the most reasonable option available (Birch and Cary 1996; Birch and Noll 2004; Bunn et al. 2002; HEI Diesel Epidemiology Working Group 2002).

The use of structural equation modeling is a creative and statistically sound approach for taking advantage of the known contributors to personal exposure and the availability of measurements in a variety of microenvironments. The emphasis on personal exposure assessment takes account of some of the unique exposure profiles of different trucking industry employees, reducing exposure misclassification relative to approaches that rely on ambient or selected microenvironmental monitoring measurements. Finally, the exposure assessment was conducted without knowledge of outcome status, removing one potential source of differential bias in outcome ascertainment.

While any retrospective exposure assignment must rely on certain assumptions, the investigators were able to validate multiple elements of their exposure ascertainment

algorithm. For example, the coefficient of haze data showed that levels in 1988–1989 were 2.2 times higher than in 2000, an identical ratio as seen when comparing the geometric mean background EC concentrations from the 1988–1989 trucker study (Zaebst et al. 1991) with those of the current study (Davis et al. 2011). Similarly, Garshick and colleagues (2012a) report that the predicted geometric mean EC in 1988–1989 for dockworkers who drove propane forklifts in the study terminals was 1.36 μg/m^3 (Davis et al. 2011), which compares very well with the value of 1.30 μg/m^3 measured by Zaebst and colleagues (1991) after background adjustment. The investigators also constructed the structural equation modeling with the initial set of 36 sampling trips and validated the approach through application in repeat trips to 6 terminals.

3.3.8 LIMITATIONS OF THE EXPOSURE ASSESSMENT APPROACH

The investigators constructed an exposure metric that was reasonably specific to diesel and well-quantified. As has been discussed, EC has been generally favored as an appropriate marker for diesel exhaust. Source apportionment analyses conducted in support of this study provide some support for the argument that most of the SEC measured is attributable to diesel exhaust in the terminal yard and in the urban background sites studied (Sheesley et al. 2008, 2009). However, it is also clear from other measurements in the source apportionment studies that the workers are also exposed to exhaust from other mobile sources. Other studies of traffic-related exposures, not specifically of truckers, suggest similar findings. On or near roads, the mixture of diesel- and gasoline-engine–related ambient EC varies according to the mixture of vehicles traveling (Riddle et al. 2008), and individual-level exposures to vehicle exhaust are influenced not just by the vehicle in which one is traveling but by emissions from other vehicles on those roads (see, for example, Zuurbier et al. 2010).

As in all historical exposure reconstructions, an important limitation of this exposure assessment is the retrospective extrapolation of current exposures. This back-extrapolation relies on coefficient of haze taken from only one area of the country (New Jersey), which is then assumed to represent trends for all the other U.S. trucking terminals in the study. This trend line, shown in Figure 3.1, was chosen because it was consistent with the results in the study by Zaebst and colleagues (1991); that is, the ratio of the current SEC to SEC levels in 1988–1989 Zaebst study matched that trend line. The temporal trends do differ somewhat in other areas of the country, which could contribute some uncertainty to the extrapolations to historical background ambient concentrations, an uncertainty that could vary by location. In addition, coefficient of haze data

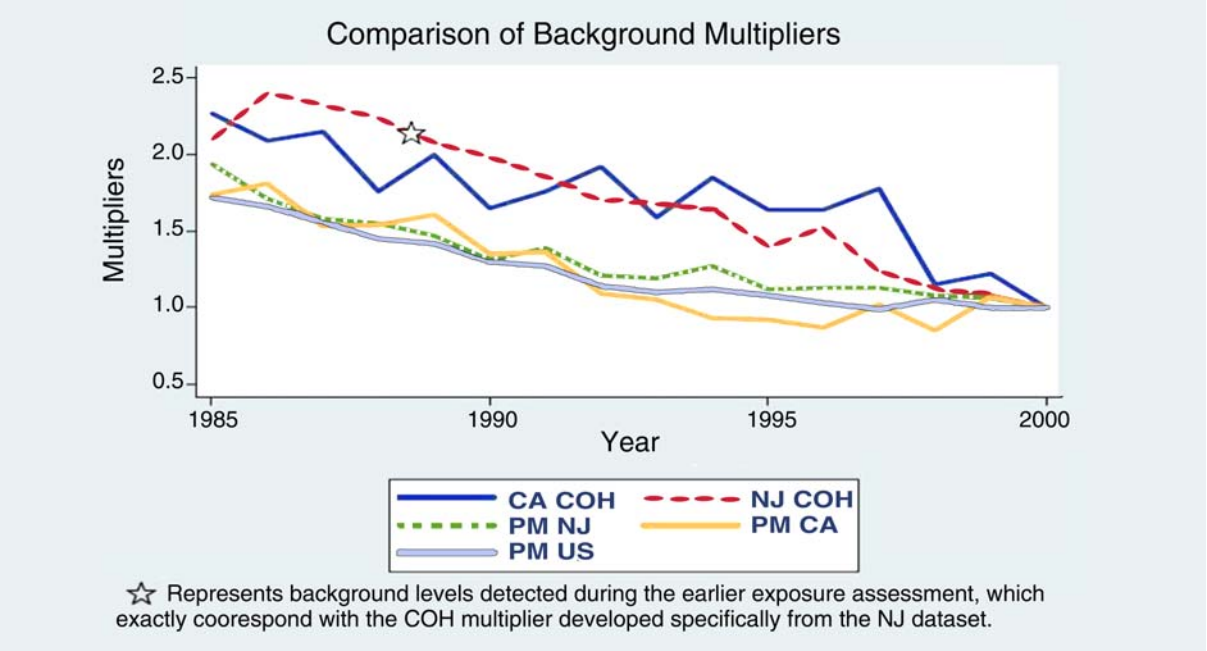

Figure 3.1. Comparison of trends in background multipliers from 1985–2000. COH = coefficient of haze. (Reproduced from Davis et al. 2011.)

did not exist prior to 1971, so the values for 1971 were assumed to be the same for all previous years. Although the absence of data before 1971 accounted for only about 8% of the total person-years in the cohort as a whole, it has some potential to influence estimates of exposure to the long-haul drivers, for whom dieselization began in the 1950s, as opposed to the pick-up and delivery drivers and dock-workers for whom dieselization began in the 1970s. The Panel thought this assumption would most likely contribute to an underestimate of cumulative exposures for a subset of long-haul drivers, which could affect the estimated slope of the exposure–response function; however, the potential magnitude of the effect of this assumption has not been evaluated. In general, there are multiple uncertainties related to the need to characterize the timing of fuel use transitions, which may not be accurately reported by the trucking companies and may be more uncertain for dates distal in time.

It is difficult to fully validate this retrospective exposure model since the main source of comparison data, the 1988–1989 study by Zaebst and colleagues (1991), also influenced the model by informing the choice of coefficient of haze data to characterize the historical trend in SEC levels and provided data for development of the fuel use multiplier for forklift use. Nevertheless, some support for the general validity of these estimates is provided by: 1) the match between the predicted SEC and estimated SEC for dock-workers in 1988–1989, 2) the similar ratios of 1988–1989 versus 2000 coefficient of haze data and Zaebst and colleagues (1991) background SEC data versus 2000 ambient measurement data, and 3) the validation of structural equation models to predict SEC by using data from repeat trips to 6 of the 36 sampled trucking terminals.

Other minor concerns include lack of discussion of how limits of detection for SEC measurements were handled, and that the ratio of SEC from Zaebst and colleagues (1991) and SEC measured in 2000 may have been influenced by the different sampling techniques used. (Note: Zaebst and colleagues [1991] used a modified dichotomous sampling cassette, described as essentially a single stage impactor, to collect submicrometer-sized particles on a 37 mm quartz filter, whereas in the Truckers study, a precision machined cyclone separator [SCC1.062 Triplex, BGI, Inc., Waltham, MA] was used to remove particles greater than 1.0 μm in aerodynamic diameter before collecting the smaller particles on a 22 mm quartz filter [Smith et al. 2006]). Differences in sampling efficiencies between the methods could lead to corresponding biases in the measurements. Hence, it is possible that the historical modifiers over- or under-adjust for historical changes in the working conditions.

Mean SEC exposure levels for the different job groups (Table 3.2) were generally lower than those in the Diesel Exhaust in Miners Study (DEMS) and other occupational settings but spanned those observed among more highly exposed members of the general population. Cumulative SEC exposures varied substantially across the cohort, with more than an order of magnitude difference between the 5th and 50th percentile and more than an order of magnitude difference between the 50th percentile and the maximum exposure.

In summary, the Truckers study investigators undertook a creative approach in the retrospective estimation and assignment of exposures in this study, with back-extrapolation approaches based on reasonably well-calibrated models. For analyses of historical trends, the retrospective exposure assignment relied on a pollutant measurement (coefficient of haze) that has long been considered a reasonable surrogate for particulate EC, reducing concerns regarding the choice of pollutant. However, the limited spatial coverage of coefficient of haze data for the locations included in the Truckers study, as well as the lack of haze data prior to 1970, leaves the potential for some uncertainty in the SEC exposure estimates. In the absence of alternative exposure assignments, quantifying the implications of key assumptions is challenging; however, the Panel found no obvious elements of the analysis that would invalidate the use of the Truckers study for risk assessment applications.

3.4 STATISTICAL METHODS AND ANALYSIS

The initial analyses of the Truckers study data involved calculation of SMRs in the original cohort in which mortality rates for several diseases, including lung cancer, were compared to those in the general U.S. population, adjusted for race, calendar year (i.e., 1-year intervals), and 10-year age groups (Laden et al. 2007). Population rates were obtained from the Center for Disease Control's WONDER database (Centers for Disease Control 2005). SMR analyses were not repeated for the older, longer duration of employment subcohort used in subsequent publications (Garshick et al. 2008, 2012a). These latter studies focused on survival analyses within the subcohort of workers $\geq$ 40 years old using proportional hazards survival models.

In this study, outcome is a failure time, attained age at death from lung cancer; there are many censored observations, since most workers did not die of lung cancer, and there are a number of additional variables potentially associated with lung cancer failure times. In this context, the Cox proportional hazards model, which is very widely used in survival data analysis, seems appropriate. It can readily

accommodate censoring, and allows modeling of the effect of explanatory variables in a familiar regression form. Age in 1-year increments was used as the timeline for the proportional hazards models, as well as calendar year of follow up (1985–2000), and penalized splines were used to examine potential nonlinearities in the exposure–response relationship.

Analyses by Garshick and colleagues in 2008 focused on the risk of lung cancer based on duration of employment within eight different job groups. Models were fit with all eight job groups included to adjust for the effects of other jobs held by the study participants. As well as implicitly adjusting for attained age based on the proportional hazards model, baseline hazards were stratified by decade of age at entry, calendar year, and decade of hire. Analyses were adjusted for race, census region (Northeast, Midwest, South, and West, based on last address), years employed, and years off work (the latter two were intended to adjust for the healthy worker survivor effect). Thus, the model includes many time-related variables that are likely correlated with each other. Assuming standard lung cancer relative risks from the literature for current/former/never smokers, the investigators constructed smoking adjustment factors by job title, which ranged from 0.92 (for pick-up and delivery drivers) to 1.17 (for long-haul drivers).

The methods used by Garshick and colleagues in 2012 largely followed the structure used in their 2008 paper but focused on the lung cancer risk associated with estimated personal exposures to SEC rather than with job group. The investigators state that in order to meet the assumptions of proportional hazards, baseline hazards were stratified by decade of hire and age in 1985 (10-year groups). Results were further adjusted for race and region of residence as in the 2008 study. Analyses were also conducted with and without duration of employment, to assess the impact of a potential healthy worker survivor bias. They were also conducted with the full cohort or with the cohort excluding mechanics (n = 1811 with 38 cancer deaths). Garshick and colleagues (2012a) justified exclusion of mechanics by positing that exposure characterization was weaker (i.e., likely to be subject to greater exposure measurement error) given substantial historical changes in job duties over time and due to differences in the nature of exposure experienced by mechanics compared to workers exposed on roadways and loading docks. Smoking was not adjusted for in these analyses given a lack of information on individual-level smoking status.

3.5 PRESENTATION AND INTERPRETATION OF MAIN FINDINGS

SMR analyses were conducted for the major causes of death for the full male cohort. Only ischemic heart disease (1133 cases, SMR = 1.41, 95% confidence interval [CI] = 1.33–1.49) and lung cancer (769 cases, SMR = 1.04, 95% CI = 0.97–1.12) had a greater number of observed than expected deaths (Laden et al. 2007). For other major chronic disease categories (i.e., diabetes, nervous system diseases, circulatory diseases [excluding ischemic heart disease, which has also been linked to fine particulate air pollution], respiratory diseases, and digestive diseases) a strong healthy worker effect was observed; the rates of these diseases were lower than in the general population, reflecting the tendency that healthier workers are more likely to remain in the workforce.

In the 2008 study, employment in four job groups (long-haul, pick-up and delivery, dockworker, and combination) was associated with an increased risk of lung cancer (expressed as HRs greater than 1) (Garshick et al. 2008). Employment as a mechanic, hostler, clerk, or other job was not associated with an increased risk. Lung cancer risk was inversely associated with overall duration of employment, which the investigators interpreted as an indication of a healthy worker survivor bias. Adjusting for smoking at a job-group level led to modest reductions in the HRs for long-haul drivers and modest increases for others.

In the most recent paper, Garshick and colleagues (2012a) estimated the HRs associated with their quantitative estimates of exposure to SEC, adjusted for race, calendar year, and census region. They presented HRs for both cumulative and average SEC exposures, unlagged or lagged by either 5 or 10 years. They included exposures to SEC in their models either in quartiles or as continuous variables.

Figure 3.2 displays boxplots showing the HRs and 95% CIs for the cumulative SEC exposures lagged 5 and 10 years, with and without adjustment for duration of work, and with and without exclusion of the mechanics from the cohort. The P values for the tests for trend are also provided for each analysis. No associations were observed between lung cancer and average SEC exposure (results not shown here, but available in Garshick et al. 2012a)

Using categorical exposure metrics, the investigators found limited evidence of association between cumulative SEC and lung cancer for the full cohort or for the cohort excluding mechanics prior to adjustment for duration of work; all HRs for individual quartiles were greater than 1 but non-significant, and there was little evidence of a positive trend in HR with increasing exposure. After adjusting for duration of work, however, the HRs generally increased, consistent

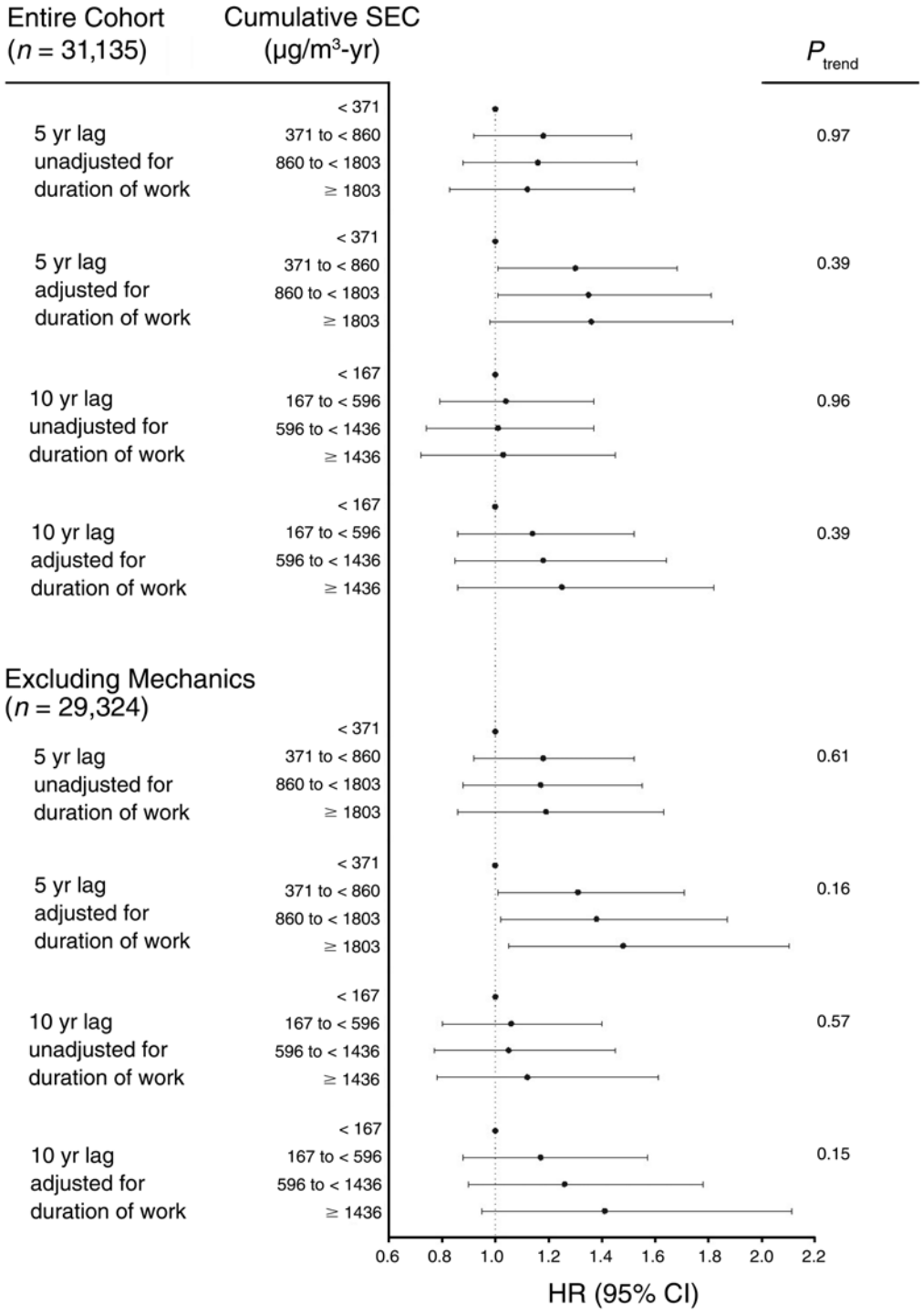

Figure 3.2. Lung cancer hazard ratios (HRs) associated with each quartile of cumulative SEC, lagged 5 and 10 years, with and without adjustment for duration of work, for both the full cohort and the same cohort excluding mechanics. Each box plot represents the maximum likelihood estimate (•) and 95% CI (whiskers) for each hazard ratio. The linear *P* values for trend were derived using an ordinal value that was based on the median of each quartile. Models were adjusted for race, calendar year of follow-up, and census region. (Data obtained from Table 4, Garshick et al. 2012a.)

with the negative confounding found previously in workers with longer tenure in their jobs (Garshick et al. 2008). In particular, in the analyses with the cohort excluding mechanics, there was some evidence of increasing associations with increasing quartiles of exposure (5-year and 10-year lag cumulative EC). The exposure–response function using continuous EC showed a borderline significant association that strengthened with increasing lag. No statistically significant association was seen with average SEC as a continuous covariate.

There was little evidence of a trend in the exposure–response relationship without the adjustment for duration of work. The investigators argue for the importance of this adjustment step, stating that duration acts as a surrogate for time-varying health status and acts as a negative confounder. Specifically, Garshick and colleagues (2012a) argued that the influence of work duration is attributable to a combination of downward bias resulting from "left truncation in a cohort composed of prevalent hires" (see Applebaum et al. 2011 for more discussion) and a healthy worker survivor effect.

Excluding the mechanics job group also strengthened the exposure–response results. The rationale given by the investigators for excluding mechanics was that "[m]echanics experienced significant historical changes in job duties that weaken the validity of extrapolation of current exposure to historical estimates" and other differences in the nature of diesel exhaust exposure in this job group (i.e., shorter duration exposures to more aged exhaust [Garshick et al. 2012a]). In other words, they hypothesized that there would be increased exposure measurement error in this subpopulation which, if included in the analyses, would lead to a weakening in the observed exposure–response relationship.

Continuous exposure metrics are generally preferable for risk assessment applications as they permit quantification of risk for precisely defined exposure scenarios. Continuous measures of cumulative SEC were linearly associated with increasing lung cancer risk in the cohort excluding mechanics, with borderline statistical significance; no results were provided for the cohort including mechanics. As discussed earlier, graphical presentations of the penalized spline model results provided to the Panel by the study investigators confirmed that the relationship did not significantly depart from linearity (Garshick E, personal communication, 2013. See Additional Materials 1, available on the HEI Web site). In the cohort excluding mechanics, the risk per 1000 $\mu g/m^3$-months of cumulative SEC exposure increased with longer lags, with the slope increasing from 0.0345 (standard error [SE] = 0.0349, $P = 0.32$) for no lag, to 0.0665 (SE = 0.0379, $P = 0.08$) for a 5-year lag, to 0.0849 (SE = 0.0501, $P = 0.09$) for a 10-year lag. Expressed as relative lung cancer hazards, these values were 1.04, 1.07, and 1.09 respectively (see Table 5 in Garshick et al. 2012a and related discussion).

As discussed earlier, the most important potential confounder in any epidemiological study of lung cancer is smoking, which was not included in the core analyses of the Truckers study due to the lack of information on individual smoking status. In the absence of ancillary analyses, this could pose some challenges for interpretation of the study findings and their application in risk assessment. Lacking individual smoking characterization and given the use of SEC rather than job title for exposure assignment, the investigators argued that adjustments similar to those conducted by job group in the earlier analyses (Garshick et al. 2008) would not have been possible, and in any event, would not have substantially altered their findings of an association between cumulative SEC exposure and lung cancer. However, the Panel noted that there are examples in comparable epidemiological contexts in which sensitivity analyses were conducted using Monte Carlo analyses to test alternative assumptions about individual smoking assignments (see for example, Steenland and Greenland 2004). Such analyses were beyond the scope of the Panel's charge but could be considered for future quantitative risk assessments.

3.6 CONCLUSIONS

The HEI Panel evaluated the Truckers study using a broad set of study attributes introduced in Chapter 2. The focal areas included the potential for confounding by smoking, the possibility for exposure misclassification error related to the historical exposure reconstruction strategy, and the overall degree of uncertainty given alternative model formulations.

The Panel's overall assessment is that the Truckers study is informative for the development of quantitative risk assessments of diesel exhaust. The study was the largest of its kind in this industry and demonstrates several strengths that make it useful for quantitative analyses of lung cancer risk, particularly of exposures to diesel and other vehicle exhaust (as represented by SEC) found in the ambient onroad environment. The retrospective exposure assessment attempted to assess historical personal exposures to diesel exhaust quantitatively, a crucial component for quantitative risk assessment. The choice of SEC was based on its specificity for diesel exhaust in this work environment and was closely connected with the coefficient of haze measurement used for historical extrapolation. The use of structural equation modeling represents a creative and sound approach to integrating information on known contributors to personal exposure and measurement data that had been obtained in several microenvironments. Given that the occupational exposure levels

were far lower than seen in the DEMS (in particular, those in the mines), accounting for exposures away from the worksite was an important step in reducing error in estimates of overall personal exposure. The investigators were able to validate multiple elements of the retrospective exposure assessment, although they had limited independent data with which to validate their predictions of SEC exposure.

The investigators' decision to adjust for duration of work creates some potential challenges in interpretation, given that cumulative SEC metrics depend on duration of employment and are therefore correlated (r = 0.55–0.74, depending on lag) in this cohort. Critics have argued that adjustment for duration of work can therefore lead to over adjustment (Morfeld 2012a). Garshick and colleagues (2012b) have responded that it was a necessary step to address negative confounding from the factors listed above. The Panel recognizes that adjustment for a healthy worker effect is often an important consideration in occupational epidemiology, although it is typically thought to be more of a problem for health outcomes other than cancer that can involve longer periods of disability and that cause people to drop out early from the workforce. Garshick and colleagues (2012a) cite several recent publications indicating that the healthy worker survivor bias may be operating for cancer as well.

With respect to use of duration of work to adjust for the healthy worker survivor bias, the Panel notes that the science on this issue is in an unsettled state. Unusual associations have been noted for many years between exposure or employment duration and outcomes, even for established causal associations such as the one between asbestos and lung cancer (Doll 1985). Theoretically, there are not only the issues of susceptible depletion and healthy worker survivor bias, but the recent recognition that an effect of 30 years of exposure requires an intervention that would immortalize exposed persons and keep them free of other competing risks (i.e., other than death) for at least that long (Flanders et al. 2014). In the Panel's view, adjustment for duration of exposure is a source of some uncertainty but not a definitive basis for precluding use of the Truckers study results in quantitative risk assessment.

The lack of individual-level data on, or control for smoking is a limitation in the Truckers study. While this omission clearly contributes some uncertainty to the risk estimates, the Panel concluded that it did not preclude their use in quantitative risk assessments. The Panel agrees that the investigators' earlier analytic approaches to estimating the potential impact of smoking by job group (Garshick et al. 2008; Jain et al. 2006) provide some reassurance that smoking is not a major explanation for the associations between SEC and lung cancer in this study and recognized the challenges in applying the same method to this study of personal exposures. However, future quantitative risk assessments could include alternative approaches to adjusting or modifying the diesel-related lung cancer results for smoking.

Beyond the challenges in controlling for key confounders, as in most studies that rely on retrospective exposure assessments, the principal uncertainties in the study results are likely related to the development of historical estimates of exposure. For the Truckers study, these have been enumerated in detail in previous sections and include important questions about: the specificity of the SEC exposure metric to diesel exhaust, particularly regarding in-cab exposures experienced in traffic which are known to reflect the full mixture of vehicles on the road; the specificity of coefficient of haze measurements in New Jersey as a historical marker for background trends in diesel/SEC levels not only for New Jersey but for all U.S. locations in the study; and the implications of transitions in fuel and engine characteristics over time. The Truckers study's investigators did not conduct extensive analysis of the sensitivity of their models to alternative exposure assumptions; however, such sensitivity analyses could be useful to pursue further in the context of a comprehensive quantitative risk assessment. The sensitivity of the results to the choice of exposure metric was evaluated to a greater extent, with consideration of average and cumulative measures with multiple lag times, and with presentation of estimates with and without adjustment for duration of employment and with and without the mechanics group. These analyses reinforced the importance of incorporating a substantial (10-year) lag and of using a cumulative exposure metric, while providing a sense of the degree of uncertainty related to selecting one specific model formulation.

Ultimately, there are many additional decisions to be made regarding how to utilize the results from the Truckers study in quantitative risk assessments in other settings or population groups, whether alone or as part of a meta-analysis. Given the availability of other studies of diesel exhaust and lung cancer, it seems likely that the Truckers study would not be the sole basis for quantitative risk assessment. It will be interpreted and applied within the broader scientific literature both on diesel exhaust and lung cancer, exposure–response modeling, and the emerging literature on emissions from newer technology engines. These issues will be discussed further in the concluding chapter of this report.

Chapter 4

Evaluation of the Diesel Exhaust in Miners Study

4.0 OVERVIEW OF STUDY DESIGN, ANALYTICAL METHODS, AND MAIN FINDINGS

The Diesel Exhaust in Miners Study (DEMS*) was designed to study associations between exposure to diesel exhaust and health outcomes in a cohort of workers in underground mines. The overall cohort consisted of 12,315 mostly white male miners engaged in work in eight nonmetal mines in various locations around the United States. These mines were chosen because of their low concentrations of other potential lung carcinogens (including radon, silica, asbestos, and nondiesel polycyclic aromatic hydrocarbons [PAHs]), use of diesel engines over a long period of time, and having good records of both work history and surrogate measures of exposure to diesel exhaust. Table 4.1 provides an overview of the DEMS cohort and nested case–control studies, the exposure and outcome assessment, and selected results that are discussed in greater detail below.

In a series of five papers (Coble et al. 2010; Stewart et al. 2010, 2012; Vermeulen et al. 2010a,b), the investigators describe the extensive exercise undertaken to estimate and validate historical exposures to diesel exhaust, for which exposure to respirable elemental carbon (REC) was used as a surrogate. A distinguishing feature of this study was the effort to develop and assign quantitative estimates of average annual and cumulative REC exposure to every individual, providing a broad range of diesel exhaust exposures with which to explore exposure-related mortality. By doing so, the investigators sought to rectify a major shortcoming of previous epidemiological studies of historical exposure to diesel exhaust that relied on qualitative or semiquantitative indicators of exposure such as job titles, duration of employment, or job–exposure matrices.

The mortality experience of the workers was followed to December 31, 1997, and ascertained by matching between the National Death Index and the Social Security Administration death files. Less than 1% of the cohort (111 individuals) could not be matched.

The association between exposure to REC and mortality was explored in both the full cohort (Attfield et al. 2012) and in a nested case–control study, which included additional information on participants (Silverman et al. 2012, 2014). The first step in the investigators' exploration of the DEMS cohort was calculation of standardized mortality ratios (SMRs) to compare observed mortality in the study population to expected mortality based on age/gender/race and state-specific mortality rates. The investigators reported slightly lower overall mortality than in the general population, which is not uncommon for active, working populations who are typically healthier than the general population. However, they found that SMRs for lung cancer, esophageal cancer, and pneumoconiosis were elevated (see lung cancer results in Table 4.1; all SMR results can be found in Table 3 of Attfield et al. 2012). In all subsequent analyses the investigators separated out two cohorts of workers: those who had ever worked underground (*ever-underground*), and those who had worked only at the surface (*surface-only*). This decision by the investigators to stratify by work location was made after a priori specified analyses of the complete cohort did not find a "clear relationship of lung cancer mortality with DE exposure" (Attfield et al. 2012). The investigators reported different patterns of lung cancer mortality by work location, related to striking differences in exposure levels between surface-only and ever-underground workers, and as discussed later, in smoking histories and other factors, that lent further support to this decision.

The primary analyses of the cohort study examined time to death from lung cancer (malignant neoplasms of the bronchus and lung, excluding tracheal) using Cox proportional hazards (CPH) models, with attained age as the time axis and including time-independent variables for race/ethnicity, sex, and birth year; the baseline hazard was also stratified by state. The final cohort included 200 deaths from lung cancer. These analyses showed increasing lung cancer hazard ratios (HRs) with increasing exposure to REC (see selected results in Table 4.1, with full details available in Attfield et al. 2012). The HRs were larger among ever-underground workers and statistically greater than 1 (where a HR of 1 indicates no effect relative to the lowest exposure group) at the highest two quartiles of exposure. Their analysis focused on quartiles of average and cumulative personal exposures to REC, either unlagged or lagged by 15 years, but also included several sensitivity analyses with respect to number of years worked in the mines (excluding workers with < 5 years of tenure), extended numbers of quantiles, exclusion of the highest exposures, and other factors. They evaluated trends in exposure–response by fitting a number of continuous models to cumulative and average REC exposures; their primary models used untransformed exposure values (log-linear models). Secondary analyses fit models to exposures restricted to less than 1280 μg/m^3-year and to log-transformed exposures (referred to as "power" models). The models were fit by

* A list of abbreviations and other terms appears at the end of the Special Report.

Table 4.1 Overview of the Diesel Exhaust in Miners Study

DEMS Cohort	Exposure Assessment	Outcome Assessment	Analysis	Selected Results
8 Nonmetal Mines 1 – limestone (Missouri) 3 – potash (New Mexico) 1 – salt (Ohio) 3 – trona (Wyoming) Year of dieselization in the mines: 1947–1967 Mean year of first exposure to DE: 1971 Mean underground tenure: 8.0 years **Cohort Population** 12,315 workers Sex: 96% male Ethnicity: 88% white, 2% black, 10% Hispanic Mean age at start of exposure: 29 years End of follow up: year 1997 278,041 person-years of follow up 2200 deaths at end of follow up 200 lung cancer deaths	**Metric**: Respirable elemental carbon (REC) in μg/m^3 **Historical REC exposure reconstruction:** Models based on REC measurements, horsepower (HP), CO, ventilation (see details in report text). Individual-level annual and cumulative REC exposure assigned based on estimated historical REC levels and job history in the mines. **Confounding exposures measured:** Silica, asbestos, nonDE PAHs, radon, respirable dust. Mines selected for low levels of these contaminants.	**Primary:** Lung cancer defined as malignant neoplasm of the bronchus and lung, excluding tracheal cancer, as underlying cause of death. **Other**: all-cause mortality, other malignant neoplasms, and chronic diseases. **Ascertainment:** National Death Index (NDI-Plus) matched with Social Security Administration files.	**Cohort 'external' analysis:** Standardized mortality ratio analysis, externally standardized to state-, age-, gender-, and ethnic group-specific death rates for each mine. **Cohort 'internal' analysis:** Cox proportional hazards (CPH) models, quantiles of average and cumulative REC, unlagged and lagged 15 years. By work location (ever-underground, surface-only). Adjustments: race, birth year, sex, state Sensitivity analyses: exposure metric, tenure exclusion, continuous models, work location.	SMRs (95% CI): Lung Cancer (LC) Complete cohort: 1.26 (1.09–1.44) Ever-underground: 1.21 (1.01–1.45) Surface-only: 1.33 (1.06–1.66) Cohort CPH analysis: Hazard Ratios (95% CI) Cumulative REC (μg/m^3-years), 15-yr lag Complete cohort (200 LC deaths) 0 to < 2.5: 1.00 (referent) 2.5 to < 56: 0.55 (0.35–0.85) 56 to < 583: 1.03 (0.60–1.77) ≥ 583: 1.39 (0.78–2.48) Ever-underground workers (122 LC deaths) 0 to < 108: 1.00 (referent) 108 to < 445: 1.50 (0.86–2.62) 445 to < 946: 2.17 (1.21–3.88) ≥ 946: 2.21 (1.19–4.09) Surface-only workers (78 LC deaths) 0 to < 0.70: 1.00 (referent) 0.70 to < 4.6: 1.28 (0.64–2.58) 4.6 to < 14: 0.73 (0.35–1.53) ≥ 14: 1.00 (0.44–2.28)
Nested Case–control: 198 lung cancer cases, 562 incidence density matched controls; matched on mine, sex, race/ethnicity, and birth year (Note that they had 666 controls for analysis purposes as some subjects served as controls for more than one case subject.)	**Nested Case–control: In addition to REC and confounding exposures above, individual questionnaire data collected on:** Smoking, medical history, occupational history, other personal risk factors		**Nested Case–control:** Conditional logistic regression: with quantiles of average and cumulative REC, unlagged and lagged 15 years, duration of exposure. Two-sided Wald test for linear trend. Various continuous models fit. Control for confounding: Smoking status and intensity, location (as a joint variable), history of high risk jobs. Sensitivity analyses: exposure metric, tenure exclusion, work location, continuous models, smoking metrics.	**Nested case–control CPH analysis: Odds Ratios (95% CI)** Cumulative REC (μg/m^3-years), 15-yr lag 0 to < 3: 1.00 (referent) 3 to < 72: 0.74 (0.40–1.38) 72 to < 536: 1.54 (0.74–3.20) ≥ 536: 2.38 (1.28–6.26) P_{trend} **= 0.001**

location (surface-only and ever-underground) and in the complete cohort, adjusted for location. Exposure–response results in the full cohort, adjusted for location, were suggestive of an increasing trend in lung cancer mortality in relation to higher levels of REC exposure. Full results for all of these models can be found in Attfield et al. 2012, Tables 4 through 6. The continuous models were also evaluated for the impact of cumulative exposures to silica, asbestos, non-diesel PAHs, radon, and respirable dust. The investigators were unable to control for smoking and other potential confounding variables in the cohort study; these were, however, explored in the nested case–control study.

The case–control study (Silverman et al. 2012), nested within the full cohort study, enabled a more detailed exploration of the potential association of REC with lung cancer because of the detailed information on smoking and other potential confounding variables that was obtained for the workers. This study included 198 subjects who died from lung cancer, for whom next-of-kin could be interviewed and who could be individually matched with up to four controls (n = 562). All members of the study cohort who were alive before the day the case subject died were eligible to serve as controls (i.e., incidence-density controls), and these were matched to cases using mining facility, sex, race/ethnicity, and birth year. For each case and control, comprehensive questionnaires were completed either by the individual (self) or by a next of kin (proxy) to obtain data on important factors that might confound or modify lung cancer risk, including smoking habits, lifetime occupational history, location of work in the mines, individual and family medical history, and diet.

The case–control study used conditional logistic regression to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for lung cancer in relation to REC exposure. As in the cohort study, the case–control study explored associations of lung cancer with quartiles of average and cumulative REC exposure, unlagged and lagged by 15 years, as well as with an expanded number REC exposure categories (eight) and with duration of exposure. In these analyses, individuals were assigned the median exposure in each quartile. Models were adjusted for smoking status (never, former, current, unknown) and intensity (packs per day), history of respiratory disease 5 or more years before date of death, and history of having worked 10 years or more in a job with a high risk of lung cancer. Results were presented by location of employment (surface-only or ever-underground) or, in the case of analyses involving all subjects (surface and ever-underground), adjusted for location of employment. Analyses by mine type (potash, trona) were also presented. The investigators also fit several continuous exposure–response models using continuous versions of the same 15-year lagged average and cumulative REC exposures as in the categorical models. See Figure 1, Silverman et al. 2012 for plots of power, linear, and linear exponential models fit to the data. The caption indicates that results for a log-linear model were excluded due to poor fit to the data.

The results from the case–control study indicated that the risk of lung cancer mortality increased with increasing exposure to REC (see Table 4.1 for cumulative REC, 15-year lag). The assessment of the potential confounding effects of smoking suggested a complicated interaction between smoking and location of work, with a stronger exposure–response relationship with smoking (as measured in packs per day) for surface-only workers than for ever-underground workers. The investigators reported attenuation of the REC effect at higher levels of smoking.

4.1 REPLICATION OF THE MAIN STUDY RESULTS

Given the potential importance of the DEMS in risk assessment, the questions raised about potentially important confounders as well as alternative exposure metrics, and the breadth of the data available from the study to evaluate these and other issues, the Panel chose to conduct additional analyses with the DEMS data. As would be normal prior to any such evaluation, the Panel's first step was to replicate the main results of the cohort and case–control studies using the analytical data sets provided under the respective Data Use Agreements with the National Cancer Institute (NCI) and the National Institutes of Occupational Health, described in Chapter 1.

The data sets and the accompanying data libraries were clear and well organized. Although the cohort analyses required more steps and assumptions to create the final input variables, with clarification from the original investigators, the Panel was generally able to reproduce the main cohort results from Attfield and colleagues (2012). The Panel replicated exactly the main case–control study results in Tables 1–7 from Silverman and colleagues (2012). The lists of analytical variables available for the cohort study (both external and internal data sets) and for the case–control study are provided on the HEI Web site in Additional Materials 2; tables demonstrating replication of the main results from the case–control study, whose data the Panel explored in further sensitivity analyses, are also provided. This replication of the original results is reassuring, not only with respect to the basic reproducibility of the results, but also for the Panel's subsequent analyses.

4.2 PANEL EVALUATION

Using the attributes outlined in Chapter 2 as a general guide, the Panel assessed the key elements of the DEMS

study design and analytical approach. The evaluation then focused on three key issues that the Panel thought merited more detailed discussion, and in some cases further analyses, because of the potential impact on bias or uncertainty in risk estimates derived from the study data. These issues included: the control for smoking; the control for exposure to radon; and uncertainties in the retrospective exposure assessment.

4.2.1 STUDY DESIGN

The Panel thought that the DEMS, in its conception, design, conduct, oversight and review, demonstrated a number of strengths that are considered desirable in high quality epidemiological studies.

The nested case–control study within DEMS was a major strength of the study, addressing one of the principles of study designs desirable in epidemiological studies for quantitative risk assessment and one of the key research needs identified in 1999 (Chapter 2, Table 2.2) — that is, the need to analyze carefully for potentially important variables that may contribute to, confound, or modify the main effect of interest. As is often the case with large cohorts, it was not feasible to obtain the necessary detailed data on all members of the cohort, and so this step was left to the nested case–control study. The Panel viewed this approach to be an appropriate use of a nested case–control design.

The study was well planned in advance of conducting the work. The study objectives, hypotheses, and proposed approaches were clearly developed and defined in an extensive analytical protocol published jointly by the NCI and the National Institute for Occupational Safety and Health (NIOSH) (NCI–NIOSH 1997). The investigators first undertook a detailed feasibility study from 1992–1994 to evaluate whether or not the number of nonmetal miners exposed to diesel exhaust, and records of their work histories, were adequate for detecting a potential association between diesel exhaust and cancer (NCI–NIOSH 1997). They studied: 1) the number of nonmetal miners exposed to diesel exhaust in underground mines between 1960 and 1979; 2) the completeness of the work history data available since diesels were introduced into the mines; 3) whether work history data contain possible exposure surrogates (such as job title, department, equipment assignment, work area) that could be used to develop individual exposure estimates; and 4) availability of historical industrial hygiene data on diesel exposure and the feasibility of using these data, as well as current exposure data, to estimate past exposure (NCI–NIOSH 1997). The study protocol was reviewed and approved by an independent, external review panel (the NIOSH Board of Scientific Counselors), whose membership included epidemiologists, statisticians, industrial hygienists, engineers, and individuals with mining expertise.

The study was adequately powered to evaluate the association between diesel exhaust and lung cancer. The investigators estimated that the cohort study had 90% statistical power to detect a doubling in lung cancer risk (Attfield et al. 2012; NCI–NIOSH 1997). For the case–control study, the investigators estimated that their cohort would yield a minimum of 140 lung cancer cases and 560 controls and, under various assumptions on the overall OR for lung cancer, that the power was between 80% and 95%, depending on the assumed true magnitude of the underlying association (NCI–NIOSH 1997). The final study included 198 cases of lung cancer and 562 matched controls, and estimated ORs greater than those assumed in the power calculations.

The ascertainment of vital status and of cause-of-death is unlikely to be an important source of error or bias in the study. The investigators ascertained the vital status of 99% of cohort members by linking to the National Death Index Plus and Social Security mortality files, and identified lung cancer cases via the National Death Index Plus or death certificates. Lung cancer as a cause of death was confirmed by an independent pathologist's review of the pathology report and/or slides for about 35% of the cases (Silverman et al. 2012). While misclassification of outcome is often a potential problem when using death certificates it is generally less of a problem with lung cancer. The Panel viewed the approaches taken in this study to identify cases to follow current epidemiological standards.

In the nested case–control study, Silverman and colleagues (2012) defined the risk sets by calendar time of follow up. This is a legitimate design choice for occupational studies, although others could also be considered (Langholz and Goldstein 1996). Attfield et al. (2012) chose attained age as the main time scale for most of their proportional hazards analyses. They also conducted sensitivity analyses changing the main time scale to time since the start of follow up, which they referred to as "time since cohort entry" (Attfield et al. 2012, Supplementary Table 17) and reported "similar findings."

The overall approach to matching controls to cases was also appropriate. With respect to age, an important risk factor for cancer, Silverman et al. (2012) matched the controls to the cases "by birth year (within 5 years)." This matching, in combination with the use of calendar time of follow up to define the risk sets, amounted to matching by attained age to within 5 years. Calendar time of birth and therefore attained age could have been matched more finely than the 5-year categories used here (Breslow et al. [1978]; Greenland [1986, 1997, 2008]; Greenland and Lash [2008]). However, the Panel did not consider this matching necessary for its

evaluation of Silverman et al. (2012) for use in quantitative risk assessment.

Although the exposure assessment will be discussed in greater detail later in the chapter, there are several features of the design related to exposure assessment that should be noted here. Based on data from the feasibility study, the mines selected for study were ascertained to have used diesel equipment during the period of study, to have had the most complete personnel and other records available with which to develop job categories and to assign exposures, and to have had low levels of exposure to other pollutants that also have associations with lung cancer (e.g., silica, asbestos, radon, respirable dust, and nondiesel PAHs). These prespecified design elements reduce the concern about bias from missing information and for potential confounding by these exposures, although objective confirmation through additional analyses is advisable (these were undertaken by the original investigators and are discussed later). Another important feature of the study design was that the exposure assessment and the assignment of exposures to individuals in the cohort were done without knowledge of vital status or cause of death. This makes it less likely that the investigators' choices and assumptions for the exposure assessment could be influenced by prior knowledge of health status, and thus, in principle, are preferable to those conducted post hoc.

The publications from the study have undergone extensive review both prior to and subsequent to publication. They underwent formal peer review before publication in leading medical journals. In addition, they were subsequently reviewed by the International Agency for Research on Cancer (IARC) as part of that agency's evaluation of the evidence on the relationship between historical diesel exhaust exposure and lung cancer (IARC 2012, 2014).

These characteristics serve individually and collectively to provide greater confidence in the study results by limiting the potential for the kinds of selection, investigator, and other biases that can distort the results of the study. As in any large retrospective epidemiological study, however, questions may remain about the other analytic decisions made by the original investigators with respect to the utility of the studies for quantitative risk assessment. The Panel evaluates several of these in the sections that follow.

4.2.2 RETROSPECTIVE EXPOSURE ASSESSMENT

The important role that retrospective exposure assessment has necessarily played in observational epidemiological studies was introduced in the context of the Truckers study in the previous chapter (see Section 3.3). A measurement-based retrospective exposure assessment was also employed in the DEMS; their process and the Panel's evaluation of it are presented in the sections that follow.

4.2.2.1 Choice of REC as a Marker of Exposure to Historical Diesel Exhaust

The first major decision that had to be made was the choice of an appropriate marker of exposure to diesel exhaust. The choice of elemental carbon (EC) for this purpose was reasonable and appropriate, even though virtually no historical data on EC were available for the mines prior to the DEMS survey in 1998–2000 (see Table 2 from Stewart et al. 2010, provided in Appendix Table C1). The question of what markers should be used to indicate exposure to diesel exhaust had been a major topic for debate in the years leading up to this study (e.g., see discussions from the HEI workshop on research directions, HEI Diesel Epidemiology Working Group 2002). Those discussions and others concluded at the time that EC was the most reasonable option available for characterization of diesel exhaust in mines (Birch and Cary 1996; Birch and Noll 2004; Bunn et al. 2002; HEI Diesel Epidemiology Working Group 2002). Although REC is not specific to diesel, it strongly indicates the presence of diesel exhaust in these mines since other sources of REC such as gasoline- or natural-gas fueled engine exhaust, wood smoke, and cigarette smoke were not generally present. The HEI Diesel Epidemiology Working Group had concluded then that specific reliable biomarkers of DE did not exist, and in any case would not have been possible to assign retrospectively to individuals in the cohort.

A number of methods for the collection and analysis of EC particulates have been developed. In the DEMS the choice was made to collect REC using a 10-mm Dorr-Oliver nylon cyclone (50% cut point of 3.5 microns) at a flow rate of 1.7 L/min on a single quartz filter. The samples were analyzed using a thermal–optical method (NIOSH method 5040, NIOSH 1998). The investigators conducted personal and area monitoring for REC and several gaseous pollutants (carbon monoxide [CO], carbon dioxide [CO_2], nitrogen dioxide [NO_2], nitric oxide [NO]) during 1998–2001 in seven of the eight mines that were still open. The surveys supported their hypotheses that there was sufficient variability in exposures among miners in the nonmetal mines to warrant an investigation of the associations between long-term exposures to diesel exhaust and lung cancer. Figure 4.1, taken from Coble and colleagues (2010), shows for each mine the distributions of personal REC measurements taken during the DEMS 1998–2001 survey for underground and surface jobs. Coble and colleagues (2010) reported that average REC exposure for underground jobs with five or more measurements ranged from 31 to 58 $\mu g/m^3$ at the facility with the lowest average exposure levels and from 313 to 488 $\mu g/m^3$ at the facility with the highest average exposure levels. The average REC exposure levels for surface jobs ranged from 2 $\mu g/m^3$ in two of the mines

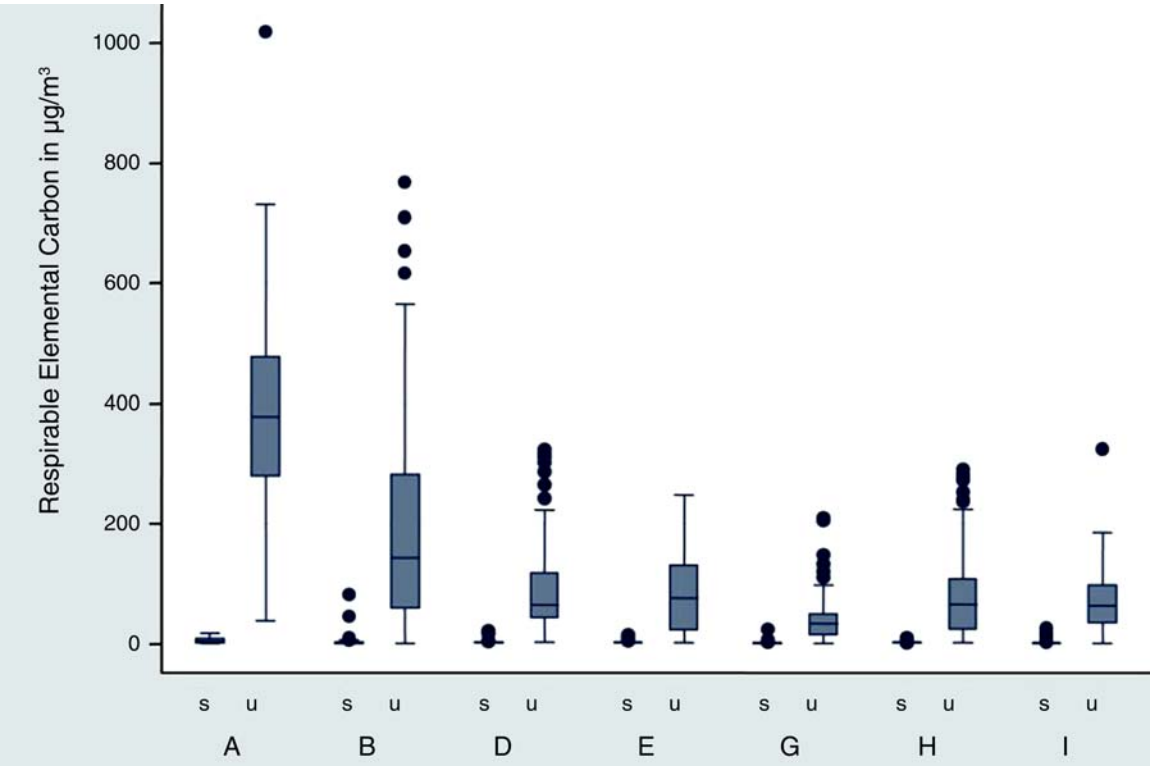

Figure 4.1. Personal respirable elemental carbon measurements (µg/m³) for surface and underground jobs by mining facility. Full-shift time-weighted concentrations; s = surface, u = underground. The boxes display the 25th and 75th percentiles, and the horizontal line within each box displays the median. The vertical whiskers extend 1.5 times the interquartile range above and below the boxes. Low or high values located outside the vertical whisker lines are displayed as points. (Source: Coble et al. 2010, Figure 1.)

(G and H) to 6 µg/m³ in Mine A. A subset of 101 measurements was taken for workers with both underground and surface responsibilities during the shift; their mean personal exposures ranged from 3 to 160 µg/m³.

4.2.2.2 Development and Assignment of Retrospective Exposures to REC

Despite agreement on REC as an appropriate marker for exposure to diesel exhaust in the mines, as in most retrospective epidemiological studies, measurements of REC were not available for most of the history of the mines. The investigators therefore needed to develop methods for historical reconstruction of exposure, combining contemporary measurements of the exposure of interest with a variety of historical measurements and data on predictors of exposure (exposure determinants).

The Panel began its evaluation of the DEMS retrospective exposure assessment with a careful review of the process the investigators undertook to estimate historical exposures to REC. It sought to understand the data and assumptions underpinning the original investigators' development of the REC estimates, to evaluate the extent to which they had tested the sensitivity of their estimates to alternative data and assumptions, to identify any major sources of uncertainty, and to place their work in the broader context of current methods for retrospective exposure assessment.

The DEMS investigators systematically described their historical exposure reconstruction process in a set of five papers (Coble et al. 2010; Stewart et al. 2010, 2012; Vermeulen et al. 2010a,b). In their first paper, Stewart and colleagues (2010) provide an overview of the process, describing the DEMS exposure surveys, their extensive collection of data from historical exposure surveys, mining methods, engine type and horsepower (HP), ventilation records, job descriptions, details of specific job activities, and employee work histories from both historical mine records and from interviews with mine employees. The individual steps taken to collect and process data for development of a historical record of personal exposure for each member of the cohort are described in detail in the remaining papers and summarized briefly below.

As part of its evaluation, the Panel compiled a summary of the several analyses undertaken by the investigators to assess the reliability of their data, the robustness of their analytical choices, and the historical modeling results at various points in the process. That summary, the results of their analyses, and the Panel's interpretation of the findings can be found in Table 4.2.

Table 4.2. Summary, Evaluation and Validation of Steps in the DEMS Retrospective Exposure Assessment[a]

Step in Exposure Assessment (References)	Assessment Used and Analysis / Comparison	Result	Interpretation
Step 1: Collection of Baseline Data Measurements			
Collection of REC measurements (Stewart et al. 2012, p 397, Supplement) (Coble et al. 2010, p 758)	Personal REC measurements were compared to area REC measurements collected concurrently on the same days at five of the facilities (Cohen et al. 2002). The five facilities were grouped into three mine types (potash, limestone, and salt) and data were compared for 3 job types (production, maintenance, and surface).	In 8 of 9 comparisons, mean differences in average REC ranged from −18% to 40%. For one surface comparison, mean difference of −550% (39 vs. 6 µg/m^3) was statistically significant.	Suggests that the personal DEMS data used to estimate 1998–2001 exposure levels were comparable to area data collected concurrently.
	The average of personal underground samples (n = 124) taken at one of the potash facilities in the DEMS survey was compared to the average of personal underground samples (n = 46) in the same facility taken during a feasibility study conducted in 1994 (Stanevich et al. 1997).	DEMS survey mean: 191 µg/m^3 Feasibility study mean: 190 µg/m^3	Suggests the DEMS survey results are reliable.
Step 2: Processing of Work Histories			
Assignment of location (underground vs. surface) of jobs of unknown location held for > 2 years (5% of job entries) (Stewart et al. 2010, p 731) (Stewart et al. 2010, Table 7)	Work histories of subjects who held a job with an unknown location for at least 2 years were reviewed with long-term facility employees, during site visit interviews. Where there was consensus among interviewees, jobs were assigned that location. Where there was no consensus, the main location of the subject, or of the job in that facility, was assigned. For the jobs with known location, the location reported by interviewees was compared to the location in the personnel file.	There was 93% overall agreement between the location reported by long-term workers and the location stated in personnel records. The percent agreement ranged from 86% to 100%, by facility.	Suggests interviews with long-term facility employees are a reliable source of information regarding assignment of job location (underground vs. surface) for subjects with missing job location information.

Table continues next page

[a] Sources: Adapted from Table 7 of Stewart et al. 2010 and Table 1 of Stewart et al. 2012.

[b] r_p is the Pearson correlation coefficient for REC 8-hr TWA personal levels compared to DEMS REC full-shift personal measurements for underground jobs.

AdjHP/CFM = adjusted horsepower divided by cubic feet per minute; AIC = Akaike information criteria; AM = arithmetic means; MIDAS = the MSHA mine information data system; ND = nondetectable; TWA = time weighted averages.

Table 4.2 (*continued*). Summary, Evaluation and Validation of Steps in the DEMS Retrospective Exposure Assessment[a]

Step in Exposure Assessment (References)	Assessment Used and Analysis / Comparison	Result	Interpretation
Step 3: Defining Exposure Determinants to Estimate Historical CO Levels			
Selection of CO to use for historical back-extrapolation (Vermeulen 2010b p769, Table 1)	Correlation between area measurements of REC and gaseous components (NO, NO_2, CO, and CO_2) from the DEMS survey, using natural-log transformed values.	Correlation with REC (r_p)[b]: 0.72 (NO), 0.66 (CO_2), 0.52 (NO_2), 0.41 (CO)	Among gases, correlation with CO was weakest.
	Factor analysis of diesel exhaust components (measures of EC, OC, gases, and particulates) loading on three factors: diesel exhaust, mine dust, and organic carbon.	Factor analysis: all gaseous components loaded most strongly on the same factor as EC.	Suggests that gases are a good surrogate of REC.
	Linear regression was conducted between REC and CO, using a mixed-effects model allowing facility-specific intercepts (fixed effects), and facility-specific slopes (random effects); natural-log transformed values. Analysis was also conducted between REC and NO_2, because NO_2 is used frequently as a surrogate of diesel exhaust in other studies.	Mean of facility-specific slopes: (CO) 0.58 (95% CI; 0.22–0.94), range 0.13–1.17 by facility Model fit: AIC = 516.8 (NO_2) 0.44 (95% CI; 0.13–0.75), range 0.16–1.04 by facility Model fit: AIC = 562.6	CO and NO_2 performed similarly in these models, though NO_2 had poorer model fit.
	Nonparametric regression analyses allowing facility-specific intercepts using generalized additive models (GAMs) were used to explore possible nonlinear relationships between REC and CO.	The association of REC with CO was essentially linear in log-log space (data not shown)	Nonlinear relationships were explored and ruled out.
Step 4: Development of REC Exposure Estimates			
Selection of the AM as the exposure metric vs. median (Stewart et al. 2012, p 392)	The arithmetic means (AM) of full-shift personal REC measurements for each underground job were calculated from the DEMS survey as the reference for the period 1998–2001. To evaluate the robustness of the AM, medians of full-shift person REC measurements were calculated as the reference estimates, and using the same primary prediction models, cumulative exposure estimates were calculated for all underground subjects based on median exposure levels.	Cumulative exposures based on means and medians were highly correlated: r_P = 0.98 (0.98 to > 0.99 by facility)	AMs are considered the best statistic for calculating cumulative exposure in evaluations of chronic disease (Seixas et al. 1991).

Table continues next page

[a] Sources: Adapted from Table 7 of Stewart et al. 2010 and Table 1 of Stewart et al. 2012.

[b] r_P is the Pearson correlation coefficient for REC 8-hr TWA personal levels compared to DEMS REC full-shift personal measurements for underground jobs.

AdjHP/CFM = adjusted horsepower divided by cubic feet per minute; AIC = Akaike information criteria; AM = arithmetic means; MIDAS = the MSHA mine information data system; ND = nondetectable; TWA = time weighted averages.

Table 4.2 (*continued*). Summary, Evaluation and Validation of Steps in the DEMS Retrospective Exposure Assessment[a]

Step in Exposure Assessment (References)	Assessment Used and Analysis / Comparison	Result	Interpretation
Step 4: Development of REC Exposure Estimates (*continued*)			
Development of underground exposure groups (Stewart et al. 2012, p 394, Table 3)	Hierarchical underground exposure groups U1, U2, and U3 were developed, independent of REC measurements. U1 is comprised of each standardized job title, grouped into U2 groups where jobs required similar proportions of time in four major underground areas (production face, haulage and travel ways, shop and office area, and crusher area). U3 groups are combinations of U2 groups with similar historical CO air concentrations. Assignment of REC estimates for the 1998–2001 period for each job depended on the number of personal measurements available and its exposure group. If ≥ 5 personal samples were available, the mean was assigned. If < 5 personal samples were available, the mean of all jobs in its U2 group was assigned. If all jobs in a U2 group had < 5 personal samples, the mean of all U2 groups in its U3 group was assigned. The estimates of time spent in each of the four underground areas were validated by calculating TWA from stationary samples and estimated time spent in each underground area, compared to full-shift personal measurements of underground workers.	TWA based on assignment of time spent in each area overestimated the full-shift measurements by a median relative difference of −19% (−48% to 20% by facility), but were positively correlated for all facilities. r_P = 0.83 for all facilities. Range across facilities: 0.15–0.72	TWAs based on estimates of time spent in areas were comparable to full-shift personal measurements in most facilities.
	Investigated between-group and within-group variance in REC measurements explained by U1–U3 job groups. Ideal grouping strategy would maximize between-group variance and minimize within-group variance.	Within-groups variance was greater than between-groups variance.	Lack of contrast between jobs reflected homogenous levels for most underground jobs.

Table continues next page

[a] Sources: Adapted from Table 7 of Stewart et al. 2010 and Table 1 of Stewart et al. 2012.

[b] r_P is the Pearson correlation coefficient for REC 8-hr TWA personal levels compared to DEMS REC full-shift personal measurements for underground jobs.

AdjHP/CFM = adjusted horsepower divided by cubic feet per minute; AIC = Akaike information criteria; AM = arithmetic means; MIDAS = the MSHA mine information data system; ND = nondetectable; TWA = time weighted averages.

Table 4.2 (*continued*). Summary, Evaluation and Validation of Steps in the DEMS Retrospective Exposure Assessment[a]

Step in Exposure Assessment (References)	Assessment Used and Analysis / Comparison	Result	Interpretation
Step 4: Development of REC Exposure Estimates (*continued*)			
Development of surface exposure groups (Stewart et al. 2012, pp 393–394)	Surface jobs were categorized into three groups, based on frequency of use, and proximity to diesel equipment of different sizes (based on interviews with long-term workers). AM and % ND of each groups' REC measurements from the DEMS surveys were examined to determine if REC levels increased across exposure groups and %ND decreased with increasing frequency and proximity.	In the three groups defined by increasing contact with diesel equipment, REC AMs increased overall (1, 3, 5 µg/m^3) and %NDs decreased overall (75%, 57%, and 47%).	The available measurements supported the grouping strategy.
Building predictive models based upon historical data of DE surrogates (Vermeulen et al. 2010a, pp 777–778)	Prediction models based on historical CO measurements, which were available to varying extent for all facilities. (Median % by facility of the CO measurements > LOD was 61%; the range was 40%–80% by facility; weighted average of the CO measurements > LOD 51% [calculated from Vermeulen et al. 2010a]). Prediction models using NO_2 developed for three facilities (most NO_2 measurements in the other four facilities were < LOD (up to 90%). Model using CO_2 measurements was considered, but > 70% historical CO_2 area measurements below typical background levels.	NO_2 model had poor fit (data not shown), and CO_2 data did not appear to be valid.	Modelling using other gaseous components was considered, but CO-based models were the most valid.
Step 5: Evaluations			
Evaluation of predictive model for historical CO levels (Vermeulen et al. 2010a, pp 779–780, 782)	Compared model-predicted CO concentrations to validation dataset: means of CO stationary measurements from six of the study facilities (1976–1977 MESA/BoM, Sutton et al. 1979), which were not used in the development of the models.	Primary CO model underestimates levels with a median relative difference of 29% (5 of 6 ranged from 24%–49%, one over-estimated: 25%)	CO levels, and thus EC, may have been higher than predicted in the majority of facilities.

Table continues next page

[a] Sources: Adapted from Table 7 of Stewart et al. 2010 and Table 1 of Stewart et al. 2012.

[b] r_P is the Pearson correlation coefficient for REC 8-hr TWA personal levels compared to DEMS REC full-shift personal measurements for underground jobs.

AdjHP/CFM = adjusted horsepower divided by cubic feet per minute; AIC = Akaike information criteria; AM = arithmetic means; MIDAS = the MSHA mine information data system; ND = nondetectable; TWA = time weighted averages.

Table 4.2 (*continued*). Summary, Evaluation and Validation of Steps in the DEMS Retrospective Exposure Assessment[a]

Step in Exposure Assessment (References)	Assessment Used and Analysis / Comparison	Result	Interpretation
Step 5: Evaluations (*continued*)			
Evaluation of predictive models to estimate REC levels (Vermeulen et al. 2010a, Table 4)	Compared model-predicted REC underground concentrations for two of four underground jobs that could be matched to data from a 1994 feasibility study in the same facility (Stanevich et al. 1997).	Model overestimated mean exposures by 10% for continuous miner (272.7 µg/m^3 vs. 248.4 µg/m^3), 6% for foreman (175.9 µg/m^3 vs. 166.3 µg/m^3)	Suggests that prediction model provides accurate estimates of REC exposure during this time period.
Alternative models of DE exposure estimates over time (Vermeulen et al. 2010a, pp 780–782) (Stewart et al. 2012, pp 395, Suppl)	For the estimation of historical REC based upon historical CO data and other exposure determinants, a 1:1 relationship between REC and CO was used in the primary models. An alternate model was explored based upon the relationship found between REC and CO in the DEMS survey data: REC = CO$^{0.58}$ (0.58 is mean of facility-specific slopes). Without modelling based on the use of determinants, actual 5-year averages of CO concentration levels from the 1976–2001 MIDAS data were used. For the period prior to 1976, extrapolation based on facility-specific annual relative changes in AdjHP/CFM was used.	The cumulative exposures of underground workers were calculated using the alternate models and were correlated with the primary model estimates. **REC = CO$^{0.58}$ Model** r_P = 0.88 (0.96–0.99 by facility) **5-year Average CO Model** r_P = 0.87 (0.95–0.99 by facility)	The models produced highly correlated estimates.
Relationship of REC to CO over time (Stewart et al. 2012)	Calculated the average year of study diesel engines by year and adjusted underground. REC exposure levels using Yanowitz et al. 2000 data. Compared these REC estimates with the primary REC exposure estimates.	Yanowitz et al. 2000 studied 1976–1997 diesel engines and found DPM, which is almost entirely REC, increased slightly less than CO (−0.003) per year back to 1976	76 REC estimates would have been 10% lower.

[a] Sources: Adapted from Table 7 of Stewart et al. 2010 and Table 1 of Stewart et al. 2012.

[b] r_P is the Pearson correlation coefficient for REC 8-hr TWA personal levels compared to DEMS REC full-shift personal measurements for underground jobs.

AdjHP/CFM = adjusted horsepower divided by cubic feet per minute; AIC = Akaike information criteria; AM = arithmetic means; MIDAS = the MSHA mine information data system; ND = nondetectable; TWA = time weighted averages.

4.2.2.2.1 Collection of Baseline DEMS 1998–2001 Data

Coble and colleagues (2010) describe the DEMS exposure survey that was carried out in seven of the eight mines in the study (one mine was no longer in operation). The surveys were conducted between 1998 and 2001 and involved personal and area measurements of REC, respirable organic carbon, and nitrogen oxides (NO and NO_2). They collected area measurements of CO, CO_2, total and submicron elemental carbon (TEC and SEC, respectively), and other agents (see Appendix Table C.1, reprinted from Stewart et al. 2010). The exposure monitoring was carried out during periods of four to five consecutive days at each underground operation and three to five days above ground. A total of 1156 personal REC measurements were taken during the DEMS surveys at the seven mining facilities. These included 779 full-shift personal measurements for underground jobs and 265 full-shift measurements for surface jobs. Finally, 101 personal measurements were collected on workers who worked both underground as well as on the surface. The investigators also collected air pollutant measurement data for potentially confounding covariates (silica, asbestos, radon, respirable dust, nondiesel PAHs) from any agencies and time periods for which they were available (i.e., from the Mine Safety and Health Administration [MSHA], Bureau of Mines [BoM], state agencies, and the mining facilities) (Stewart et al. 2010).

4.2.2.2.2 Processing of Work Histories (Coble et al. 2010)

Since the DEMS survey did not include measurements for all jobs in the facilities, the investigators devised a grouping strategy to try to ensure that every job at each of the mining facilities could be assigned a mean exposure estimate. They assigned each underground job to one of five exposure groups based on information ranging from the most to least specific: 1) standardized job titles; 2) groups of standardized job titles combined based on the percentage of time in the major underground areas; 3) larger job groups based on similar historical CO concentrations; 4) jobs that took place in the mine underground; and 5) jobs assigned based on expert judgment ("overrides"). Surface jobs were categorized based on the size of the diesel engine, the amount of time used, and the proximity of the job to the equipment. This resulted in three job categories for surface workers: group A (no or limited contact with diesel equipment) 69% of the surface exposure-years; group B (bystander or incidental contact) 23% of the surface exposure-years; group C (operation of large pieces of diesel equipment) 4% of the surface exposure-years.

4.2.2.2.3 Historical Extrapolation of REC Exposures to Surface Workers

The investigators chose a simple approach to assigning historical exposures to REC for workers on the surface of the mines given limitations in the exposure measurement data (fewer samples in the DEMS 1998–2001 survey and a high percentage of values [63%] below the limit of detection [LOD]) (Coble et al. 2010; Stewart et al. 2010). Detected REC values ranged only from 2–6 μg/m^3. To characterize historical exposures, values were first imputed for missing data and arithmetic mean exposures were estimated and assigned to the different surface job groups (defined above) using a specified decision framework (Stewart et al. 2010). Using these estimates from the DEMS survey, these exposures to diesel exhaust were assumed to be constant over the period of the study, extending back to a year "either the first year the particular type of diesel equipment was used by the job (which was reported on facility records or was estimated from information collected during the interviews) or to the year when diesel equipment was first introduced in the area where the job was located" (Stewart et al. 2010).

4.2.2.2.4 Selection of CO for Back Extrapolation and of Exposure Determinants with which to Estimate Historical CO in the Mines (Vermeulen et al. 2010b)

Given that REC measurements were not available for most of the study period, the investigators had to develop predictive models based on other markers of exposure, associated with REC, for which more historical data were available for the periods prior to the DEMS survey. Vermeulen and colleagues (2010b) investigated the interrelations between various particulate and gaseous markers of diesel exhaust concentrations, using side-by-side area measurements taken during the 1998–2001 surveys. Although the Pearson correlations between the natural log values of REC and CO were lower than for other gaseous markers of diesel exhaust (NO, NO_2, and CO_2), the investigators chose CO for historical modeling purposes based on the relative completeness of the data available including the number of samples, coverage of mines and percentage of results above the LOD. In addition, CO had a better fit in other exploratory analyses of the relationships between REC, gaseous pollutants, and particulates in the mines (see Table 4.2, Step 3).

The investigators reported that nonparametric regression analyses of the DEMS measurement data showed that the association of the natural log of REC [Ln(REC)] with Ln(CO) was essentially linear across a wide range of concentrations, although the data were not provided in the paper (Vermeulen et al. 2010b). The results of a linear mixed effects regression model that allowed for fixed facility-specific intercepts and random facility-specific slopes indicated that the mean slope was less than 1 (specifically 0.58; 95% CI: 0.22 to 0.94). Nonetheless, the investigators argued for a 1:1 relationship between Ln(REC) and Ln(CO) concentrations

(β = 1) in their main models, citing a number of reasons in their paper and in presentations at the HEI Diesel Epidemiology Workshop in March 2014: 1) data from Yanowitz and colleagues (2000, Figure 3) showed similar trends by model year in particulate matter (PM) and CO from heavy-duty diesel engines; 2) the argument that large-scale increases in the HP and ventilation affect CO and REC concentrations similarly; and 3) a concern about over-reliance on regression coefficients based on the cross-sectional data collected in 1998–2001 that might not represent historical relationships over time. As will be discussed in Section 4.2.2.3, the sensitivity of both the exposure estimates and the risks of lung cancer to this assumption were tested both by the investigators and by other investigators.

4.2.2.2.5 Development, Testing, and Application of the Retrospective Exposure Model to Predict Historical REC Concentrations in the Mines (Vermeulen et al. 2010a)

Having selected CO for the development of their retrospective models, the investigators then focused on the other determinants of diesel exhaust concentrations in the mines. Databases had been constructed for each facility by year, dating back to 1947, with air sampling data and with information on numerous potential determinants of diesel exhaust concentrations.

The investigators developed facility-specific regression models based on data for the period 1976–2001 when CO levels and other facility-specific determinants were available for each mine. The basic form of the regression model is given by:

$$\begin{aligned} Ln(CO) = {} & \propto + \beta_1 \times Ln\left(\frac{AdjHP}{CFM}\right) \\ & + \beta_2 \times Ln\left(AdjHP_{1990+}\right) \\ & + \beta_3 \times Season \\ & + \beta_4 \times Survey \\ & + \beta_{5\ldots i} \\ & \times \left(Additional\ facility\text{-}specific\ determinants\right)_{5\ldots i} + \varepsilon. \end{aligned} \quad (1)$$

These models included two fundamental factors contributing to airborne levels of pollutants in occupational environments: 1) an indicator of emission rate — the HP of the diesel equipment, adjusted for the percentage of time the equipment was in use (adjusted HP [AdjHP]); and 2) a measure of dilution rate — the total rate of airflow exhausted from the underground operations in cubic feet per minute (CFM). The AdjHP for vehicles purchased after 1990 ($AdjHP_{1990+}$) was included as a separate variable in the model to account for lower emissions from newer, cleaner engines. They also included variables to account for the time of year (Season) in which the measurements were taken and the source of the measurement data (Survey) (details for the other determinants were not given). The investigators indicated that they considered a number of other possible facility- and year-specific determinants in the development of the models, but did not include any that did not achieve statistical significance (i.e., fuel use, ore production rates; mining methods, ore haulage methods, choice of explosives, and various engineering controls, work practices, and work place characteristics).

By inputting determinants from earlier time periods into each facility-specific model, they estimated annual CO concentrations [predicted in the models as Ln(CO)] in each facility for the full study period (1947–2001).

The investigators next needed to develop a method for converting the changes in annual CO estimated from their model to changes in annual REC for each facility. They did so by developing a factor they called *RELtrend*, which scales the ratio of the estimated CO concentrations in a particular year *i* to the estimated CO concentration for the reference year in each facility (i.e., the year of each facility's DEMS survey), using a coefficient of proportionality, β, assumed in the main models to be 1:

$$RELtrend_i = \left(\frac{Estimated\,CO\ for\ year\ i}{Estimated\,CO\ \ for\ reference\ year}\right)^{\beta}. \quad (2)$$

For each mining facility, they estimated annual REC exposure for each year, *i*, and exposure group(job), *k*, by adjusting the REC level for each job in the reference year, *R*, of the survey (1998–2001, depending on when the measurements were made in the mine), by the year $RELtrend_i$ and adjustments for the percentages of time workers in that job group spent underground in mine air versus underground in intake air:

$$\begin{aligned} REC_{ik} = {} & REC_{kR} \\ & \times \left(\left(RELtrend_i\left(\frac{\%T_{Mine\,air}}{\%T_{Underground}}\right)\right) + \left(\frac{\%T_{Intake\,air}}{\%T_{Underground}}\right)\right). \end{aligned} \quad (3)$$

These exposure group-specific REC exposures were then combined with individual employee work histories to estimate average annual REC (μg/m^3) and cumulative REC (μg/m^3-yrs) exposures for each individual in the cohort. Figure 4.2, taken from Vermeulen and colleagues (2010a), shows the estimated REC exposures in μg/m^3 over time for one of the more highly exposed categories of workers, the mine operator, for each of the eight mines (except Mine A for which the loader operator was used).

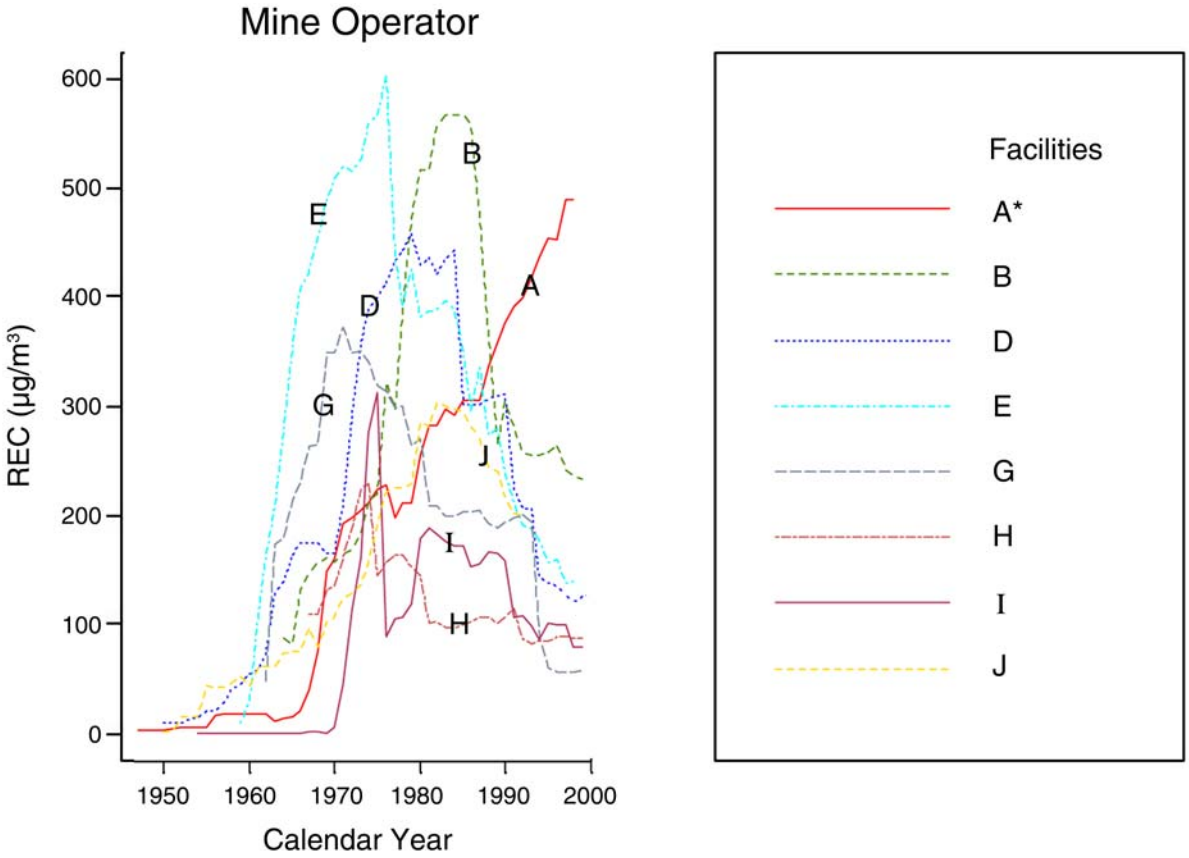

Figure 4.2. REC historical predictions (µg/m³) for the mine operator, based on the primary facility-specific CO models, by mining facility. * Facility A had no mine operator and therefore the loader operator is depicted. (Source: Vermeulen et al. 2010a, Figure 3.)

4.2.2.2.6 Evaluation of the Retrospective Exposure Assessments Using Independent Data and Methods At several points in this process the investigators made efforts to evaluate the reliability of the data they collected for the exposure assessment (Table 4.2, Step 5). They examined the robustness of the predictions of REC when varying some of the underlying assumptions (e.g., using different regression models and a different relationship between Ln(REC) and Ln(CO) other than 1:1 proportionality [i.e., $\beta = 0.58$]; see Table 4.2). Stewart and colleagues (2012) also compared the predicted CO levels with measured CO levels from the Mine Safety Enforcement Administration/Bureau of Mines (MESA/BoM surveys) conducted in 1976–1977, a data set not used in the development of the investigators' model. No CO data were available prior to 1976 that could be used in model validation. These evaluations of their analyses were completed prior to the publication of the epidemiological exposure–response analyses.

4.2.2.3 Panel Assessment

The Panel's overall assessment of the DEMS retrospective exposure analysis was that it was logical, thorough and meticulous, and used state-of-the-art methods for quantitatively estimating personal exposures for different job categories. The data collection for this study was extensive, and to the extent that it could be evaluated from materials provided to the Panel, comprehensive. Nonetheless, as is common in retrospective epidemiological studies, they had to contend with important gaps in data, and had to make a number of choices at every step to which the final estimates of exposure and risk could be sensitive. The fundamental concern about any exposure assessment relates to the nature and magnitude of error in the exposures assigned to individuals over time that can lead to under- or over-estimation of the association between exposure and health outcome of interest — on average or across the range of observed exposures —

by altering the shape of the exposure–response relationship. In addition, while systematic biases in exposure assignment (e.g., if all participants had been assigned exposures a factor of 2 greater than their true exposures) would have a limited influence on statistical significance in epidemiological studies, such biases can be extremely important in the development of exposure–response relationships for risk assessment applications. In this study, as in most studies, there may be multiple sources of error and both kinds of influences may be operating. For these reasons, the kinds of sensitivity analyses conducted by the DEMS investigators are particularly important and reflect the kind of sensitivity analyses that should be conducted with any alternative exposure reconstruction analysis.

In its evaluation of the retrospective exposure assessment conducted by the original investigators, the Panel focused on the models used to estimate historical exposures to CO and REC in the mines; the Panel did not undertake a detailed evaluation of the job assignments histories and other assumptions on which individual subject exposures were based (i.e., for determination of job group, *k*). The REC measurements collected in the DEMS survey appeared consistent with measurements taken in earlier surveys for the feasibility study (see Table 4.2, Step 1). They were replicable by others; in their detailed evaluation of the DEMS exposure assessment, Crump and Van Landingham (2012) reproduced several summaries of REC and other exposures by mining facility and job title that were originally reported by Coble and colleagues (2010).

The extrapolations involved in the DEMS historical exposure assessment relied on a number of important assumptions, specifically:

1. that CO is the most practical marker available for this study with which to model historical trends in exposure levels;
2. that the trends in CO area concentrations can be estimated using a model, based predominantly on the annual number and HP of engines used in the mines, the annual ventilation rates, and other mine-specific determinants of diesel exhaust;
3. that there is a relationship between CO and REC concentrations and that relationship (β) over time is most reasonably represented by a 1:1 proportionality factor; and
4. that the rate of diesel exhaust emissions represented by a unit of HP is relatively constant, with the exception for engines introduced in the period post-1990, for which an additional term was included in the model to account for cleaner engines and reduced emissions.

These assumptions have been the subject of considerable commentary and debate. Appendix F provides background material on diesel fuel combustion and emissions that underpin concerns about the basis for the assumptions. The discussion that follows reflects the Panel's assessment in light of both those concerns and of the broader set of factors involved in empirical modeling of the retrospective exposures to REC.

1. CO is the most practical marker available for this study with which to model historical trends in exposure levels.

Critics of the DEMS work have questioned whether CO was the best marker with which to model historical trends in REC exposures, arguing that the CO measurements were imprecise (Borak et al. 2011); and that the numbers of CO measurements were limited (Crump and Van Landingham 2012). The Panel reviewed measurement data collected for all the gases made available to both the Engine Manufacturers Association (EMA) consortium and to HEI, and concluded that the CO data were the most complete (more samples, fewer samples below detection limits, better distribution among the mines) among the options considered by the original investigators (see Sidebar on next page). Crump and colleagues (2015) also came to the conclusion that "the shortcomings in the data available for the other gaseous contaminants were even greater" than the limitations in CO, with the implication that use of markers other than CO for modeling would likely lead to more uncertainty in the REC exposure estimates.

2. The trends in CO area concentrations can be estimated using a model, based predominantly on the annual number and HP of engines, the annual ventilation rates, and other mine-specific determinants of diesel exhaust.

The Panel concluded that the logical basis for construction of the model was sound. The investigators assumed that as more diesel engines are used in a mine, more diesel exhaust is emitted and that as ventilation is added to the mine, it dilutes the diesel exhaust, thus reducing its concentration. This assumption is consistent with industrial hygiene theory and practice for meeting workplace limits on exposure (as well as with basic exposure modeling principles). Haney and Saseen (2000) demonstrated that diesel particulate exposures in mines could be predicted using a simple deterministic model and the necessary model inputs ("diesel particulate emission rates, engine horsepower, number of engines, engine operating time, length of the work shift, quantity of ventilating air, fuel properties, and efficiency of applied control technology"). Unfortunately,

the data needed to use such a deterministic model retrospectively with this cohort were not available, in particular the diesel particulate emissions rates; however, the basic relationships described in Haney and Saseen (2000) represent the conceptual framework for the development of the DEMS exposure model. Thus, within the DEMS model, the term AdjHP/CFM is a primary determinant of the historical REC concentrations in the mines (see Appendix Figure F.2). Other factors in the retrospective exposure assessment process, including the reference REC levels for particular job groups, and the fraction of time workers in the job spent in areas with fresh intake air or in areas with general mine air affect the magnitude of actual concentrations at given points in time.

The DEMS investigators conducted a number of evaluations to assess the fidelity of their predictive models to actual historical data, where available, and the sensitivity of predictions to alternative models and model assumptions (details in Table 4.2, Step 5). They were able to conduct a limited validation of their predicted CO concentrations by comparing them with CO data from a MESA/BoM survey conducted in 1976–1977 in six of their facilities that were not used in their models. For that time period, they showed that their model estimates of CO differed from measured concentrations by 24%–49% (overall median difference: 29%). Choice of alternative data from the mine information data system (MIDAS) survey in their models led to poorer agreement of their predictions but did not ultimately affect *RELtrend* used to predict historical REC concentrations (Crump and van Landingham 2012). They were able to show for two underground jobs that their model-predicted REC underground levels were within 10% of measured values for those same jobs from the 1994 feasibility study. That they were unable to validate model estimates from earlier dates is a limitation of the available data, not of their approach.

3. That there is a relationship between CO and REC concentrations, and that relationship (β) over time is most reasonably represented by a 1:1 proportionality factor.

The choice of a 1:1 proportionality ($\beta = 1$ in equation 2 for *RELtrend*) between time trends of concentrations of CO and REC has been particularly criticized for both theoretical and empirical reasons (Borak et al. 2011; Crump and Van Landingham 2012). The theoretical argument focuses on CO and REC emissions (rather than concentrations) and emphasizes that a constant relationship would not be anticipated. In this line of argument, outlined in Appendix F, the CO and REC emission relationship is unique to specific engine types, years, fuel compositions and duty cycles and there is "no universal relation between CO and particulate matter" emissions across an engine fleet (Clark et al. 1999). The empirical argument relates to whether an alternative value for β would be more appropriate to select, based on analyses of the measurements made in the mines.

First considering the theoretical argument, the sequence of Figures F.4 to F.6 presented in Appendix F illustrates how variability in the relationship between CO emissions and particulate matter (PM) emissions increases as one moves from a single engine operated on a single driving cycle, to a group of different engines on the same driving cycle, to a group of different engines operated over different driving cycles (the role of HP was not accounted for). For another set of engines and operating conditions, this sequence of relationships might look quite different.

Rationale for Selection of CO to Model Trends in REC

CO was the most frequently measured gas in the mines over time (11,124 area measurements, 46 personal measurements), followed by NO_2 (5042 area, 1798 personal measurements), then CO_2 (501 area measurements) (Stewart et al. 2010, Table 3-2). For NO_2, three facilities had up to 90% nondetectable measurements (Coble et al. 2010); only three facilities had useable data for modeling NO_2 changes over time relative to mine characteristics, and the resulting models demonstrated somewhat poor fit (Vermeulen et al. 2010b). For CO_2, the quality of historical area measurements was uncertain given that so many values (i.e., > 70%) were below the typical background level of 375 ppm (Vermeulen et al. 2010b). CO had the largest number of samples available at each location and although a large percentage of CO measurements were also below the limit of detection (ranging from 20%–60% < LOD and 39% of all samples overall), the percentages were lower than for the other gases. Thus, CO appeared the best choice for retrospective modeling (Vermeulen et al. 2010a). Also, the investigators reported that "CO correlated moderately to REC and it loaded most strongly on the factor that included EC" (Vermeulen et al. 2010a).

Given this background, the scatter observed in the empirical measurements of CO and REC concentrations in the individual mines, or in all mines combined, during the DEMS survey period (1998–2001) is understandable (from Vermeulen et al. 2010b; shown in Appendix Figure F.7). Nevertheless, the correlation between all CO and REC concentration measurements was reported as 0.41 (Vermeulen et al. 2010b), reflecting the important additional influence of the amount of diesel activity and of ventilation on concentrations (even in the presence of heterogeneous emissions characteristics).

Turning to the empirical argument, Vermeulen and colleagues (2010b) provided a statistical description of the empirical relationships between the natural log transformed REC and CO concentration measurements obtained in the mines in the DEMS survey (1998–2001) for both individual mines and for all mines combined. They found variation in that relationship among the mines, expressed by mine-specific mean regression coefficients (β values) ranging from 0.13 to 1.17. This variation could reflect statistical uncertainty in these regression coefficients but could also be explained by mines operating different numbers and types of equipment along with differences in other factors that affect concentrations of individual pollutants (see Table 4.2, Step 3). The overall regression coefficient, $\beta = 0.58$, represents the mean relationship observed in the data from all mines derived from fitting a linear mixed-effects model that allowed for fixed facility-specific intercepts and random facility-specific slopes.

As discussed earlier in the presentation of the model (see Section 4.2.2.2.4), Vermeulen and colleagues (2010b) ultimately assume a value of $\beta = 1$ in the estimation of *RELtrend* (Equation 2) for their main model, a value they note is equivalent to assuming that a given estimated change in CO concentrations relative to the reference year will be associated with a directly proportional change (a 1:1 proportionality) in REC concentrations over time. They have assumed this proportionality factor to remain constant over the period of study (as do any of the other models using alternative coefficients).

The Panel thought that their rationale for this broad assumption was reasonable given that historical changes in the ventilation rates and in the amount of diesel activity within the mines could be expected to have a similar impact on CO and REC over time. As they note, the empirically estimated coefficient of 0.58 is derived from the cross-sectional data obtained in the 1998–2001 DEMS survey and might not be representative of the CO–REC relationship over time. The empirical relationship reflects the underlying conditions and operations in the mines at that time, including the presence of particular combinations of engines and their associated relative emissions of CO, REC, and other elements of diesel exhaust. The earlier discussion based on material in Appendix F illustrates how differences in diesel engines and operating conditions can lead to heterogeneity in CO–REC relationships and our lack of knowledge about influence, if any, of these factors in the mines on historical trends in these relationships leads to uncertainty in the actual relationship between CO and REC in any year and in the exposure estimates.

With respect to the empirical argument, the key question is how this uncertainty can be explored and to what extent it affects the estimated exposure–response relationship for REC and lung cancer. This and other questions have been explored by both the DEMS investigators and by Crump and van Landingham (2012) and Crump and colleagues (2015) by varying the choice of β (whether a value of 1, 0.58, or an alternative value is utilized in the exposure reconstruction) and through modifications to other assumptions.

The DEMS investigators created two alternative REC models:

- Using $\beta = 0.58$ rather than $\beta = 1$ in the *RELtrend* equation, which when the REC estimates were compared produced highly correlated results Table 4.2, Step 5).
- Using an alternative to *RELtrend* to predict historical averages in CO using 5-year moving averages of the CO measurement data. A comparison of the REC estimates again found the two sets of estimates to be highly correlated (Table 4.2, Step 5).

The high correlations between alternative estimates indicate that the relative rankings of exposures between the subjects would not change; the relative risk for different exposure groups would also be unlikely to change. The results of supplemental analyses conducted by Silverman and colleagues (2014) using these alternative metrics have demonstrated this to be the case. However, depending on the actual magnitude of the differences in REC estimates, the magnitude of the slopes of the exposure–response relationship would be different.

Crump and colleagues extended the work of the original investigators by developing six alternative models for estimating historical measures of REC (Crump and Van Landingham 2012; Crump et al. 2015). These models test the sensitivity of the exposure, and ultimately, lung cancer risk estimates to many of the important assumptions underlying the analysis of the DEMS data and the development of the historical estimates of REC.

- REC 1 uses an alternative approach to imputing missing CO values and alternative regression models fit to

the CO and REC data, yielding $\beta = 0.3$[†] (as originally published in Crump and Van Landingham 2012).

- REC 2 is a variation on REC 1, removing the "High Period" variable for Mine H created by the original investigators to deal with what they say is an anomalous period of high exposures in that mine, but still using $\beta = 0.3$.
- REC 3 is a variation on REC 2 that addresses concerns that the variable included in the DEMS models to account for newer technology in the mines post-1990 (AdjHP+1990) prematurely anticipates the ability of newer regulations in the mines to reduce REC levels. A value of $\beta = 1$ was used in this model.
- REC 4 is a variation on REC 3 using a value of $\beta = 0.3$.
- REC 5 uses 3-year averages of CO samples post-1975 (instead of 5-year averages used by Silverman et al. [2012]), with $\beta = 0.3$.
- REC 6 estimates the REC in a given year independently of CO. This model estimates REC concentration in a given year relative to the concentration in the 1998–2001 DEMS survey using a ratio of AdjHP/CFM for a given year divided by the corresponding AdjHP/CFM during the period of the DEMS survey (see details in Appendix F).

The six alternative REC exposure estimates derived from the above models showed similar patterns over time to the REC exposure estimates of the DEMS, with some exposures being higher in some periods and lower in others relative to those in the DEMS (Crump et al. 2015).

Vermeulen and colleagues (2010a) found from their regression analysis that the $AdjHP_{1990+}$ term improved the models by accounting for an observed reduction of CO levels in the 1990's. Crump and colleagues (2015) speculated that 1990 would have been too early for a substantial infiltration of new engines into the mines and in their REC 3 model removed the term. Lower cumulative exposures observed when using this model would be expected because removing the $AdjHP_{1990+}$ factor increases the denominator (estimated CO concentration in the reference year) relative to the numerator (estimated CO concentration in the *i*th year before 1990) in *RELtrend* in Equation 2. The smaller ratio has the effect of making the REC estimates smaller in earlier years from which the cumulative exposures, lagged 15 years are developed.

Of the alternative models developed for historical REC exposures, the Panel thought the REC 6 model was particularly informative. This model does not depend on any assumptions about relationships with CO; nor does it use an adjustment for engines purchased after 1990. It depends only on the mine-specific concentrations of REC in 1998–2001 and AdjHP and CFM estimates developed by the DEMS investigators for each of the mines based on equipment records, including model numbers, ventilation records, and interviews with mining employees. Consistent with the deterministic modeling structure defined by Haney and Saseen (2000), the REC 6 model reinforces the fact that diesel equipment utilization and ventilation are the drivers of REC concentration trends over time and between mines. (See Appendix Figure F.2, which compares DEMS REC to AdjHP/CFM, and Figure F.10, which compares the DEMS REC and REC 6 estimates for the Mine E operator.)

4. The diesel exhaust emissions represented by a unit of HP is relatively constant with the exception of engines introduced in the period post-1990, for which an additional term was included in the model.

Some of the models discussed in the previous section address uncertainties about the relative trends in CO and REC arising from changes in diesel engine technology and fuels over time. The related concern raised by the engine manufacturing community, articulated in Appendix F, is the extent to which the DEMS retrospective REC exposure estimates also reflect changes in absolute levels of PM emitted. On the one hand, certification data from the U.S. Environmental Protection Agency (U.S. EPA) on PM emissions per HP-hour by model year from onroad heavy-duty diesel engines document a steady decline in emissions from 1975 to 1995, the period over which they had data (U.S. EPA 2002, shown in Figure F.9). Data from the testing of a small number of historical engines from model years 1950 to 1975 suggest that the emissions rates going back to 1950 were similar, on average, to those in 1975 (Fritz et al. 2001). On the other hand, the analysis in Appendix F suggests that the DEMS model implicitly reflects smaller ratios of REC emissions rates per unit HP-hour in the earliest periods of the mines when equipment was older (see Figure F.11), and again in the later period when an adjustment was made for post-1990 engines as already discussed.

To explore this issue further, the Panel suggested a possible approach to incorporating changing emissions patterns more directly into the REC exposure estimates. The approach, described in Appendix F, essentially relies on the simple model that depends on AdjHP/CFM and the reference REC levels in 1998–2001 (the same model developed independently by Crump et al. 2015, REC 6). In Appendix F it is illustrated with data for the mine operator in Mine E

[†] In comments to the HEI Panel on the draft report (Silverman D, personal communication), the DEMS investigators have indicated that their efforts to reproduce this value of $\beta = 0.3$ suggest Crump and van Landingham (2012) fit incorrect models. As Crump and van Landingham's (2012) analysis also involved alternative imputation methods, the Panel could not resolve this disagreement but has left the assumption as a sensitivity analysis.

and the U.S. EPA (2002) on road emissions data (Figure F.12). Crump and colleagues (in press) have subsequently extended this approach to estimate exposure for all job categories and mines and have also reestimated the odds ratios in the case–control study. Taken together these analyses suggest that exposures might have been higher than the DEMS estimates historically, in some cases by as much as a factor of 2 or more in some years (see Figure F.12) in Mine E, but in other cases by very little (e.g., Mines D and I in Crump et al. in press).

All of these alternative exposure analyses demonstrate that it is possible to test the sensitivity of the DEMS exposure analysis to a variety of different assumptions, an important objective given that historical exposure reconstruction is inherently uncertain. The challenge for reviewers and ultimately for risk assessors is to understand the relative plausibility of these different models and therefore the extent to which they can be utilized to characterize uncertainty or bias in the historical exposure estimates. Some of the alternative assumptions may seem relatively modest (e.g., $\beta = 0.58$ versus $\beta = 1$); others are potentially quite strong (e.g., the assumption explored in Appendix F that onroad technology was reflected within a few years in the engines used in mines) and somewhat conflicting with other models (e.g., REC 3, which removes the adjustment variable for post-1990 technology); others are in dispute (e.g., $\beta = 0.3$). The relative merit of these different assumptions and exposure estimates can only be fully assessed by subjecting them to the same systematic assessment and comparisons with empirical data that were conducted by the original investigators with the DEMS exposure estimates. The ultimate question is the extent to which these alternative exposure scenarios affect quantitative characterization of the REC–lung cancer exposure–response relationship, a question the Panel returns to in Section 4.4.

4.2.3 STATISTICAL ANALYSES

The investigators' approach to the analysis of the DEMS cohort followed a logical and standard progression. They first conducted exploratory external analysis of the full cohort, using SMRs, followed by extensive internal analysis of the full cohort, and finally a more detailed study of confounding variables with the questionnaire data from the nested case–control study.

4.2.3.1 Cohort Study

4.2.3.1.1. Standardized Mortality Analysis The standardized mortality analysis was appropriately done and was a reasonable first step in exploring the mortality experience in the study population. Some reviewers of the DEMS study have argued that the higher SMRs for lung cancer in the surface-only workers, where REC exposures are lower, are inconsistent with an exposure–response effect (Hesterberg et al. 2012b). However, these results should not be overinterpreted for the following reasons. First, these analyses cannot take into account any covariates (e.g., smoking, other occupational exposures) that could differ between the study cohort and the general population. Furthermore, the SMRs by worker location in this study are not necessarily comparable given that they are developed using an indirect standardization method whereby stratum-specific mortality rates in the general population are weighted according to the age- or other-specific strata in each study population (e.g., by location), and the weights may differ between populations. Overall, the results in the external analysis showing elevated risks of mortality from lung cancer are broadly consistent with those of the internal analyses of the cohort and nested case–control study discussed below. SMRs are not ideally suited for modeling an exposure–response relationship, and so the Panel chose not to focus further on these in its evaluation of the study.

4.2.3.1.2. Exposure–Response Modeling One of the most important choices the investigators had to make was how to characterize the exposure–response relationship to investigate the association with historical diesel exhaust exposures and lung cancer. Cancer risk estimation models can be divided into two broad categories: empirical models and biologically based models. Empirical models employ flexible parametric or nonparametric functional forms to describe exposure–response relationships in statistical terms. They do not explicitly consider the molecular mechanisms involved in carcinogenesis but can provide good fits to toxicological and epidemiological data, as well as reasonable estimates of cancer risk within the range of the available data. Biologically based models are derived by developing functional forms that are intended to reflect the underlying biological mechanisms involved in the process of carcinogenesis, taking into account critical biological processes such as mutation and cell proliferation.

4.2.3.1.2.1 Cox Proportional Hazards Modeling The CPH model chosen by the DEMS investigators is an example of an empirical modeling approach and one that was generally suitable for the type of data in the study.

Attfield and colleagues (2012) present the CPH regression model in terms of the hazard function, a set of k explanatory variables, and diesel exhaust in the form:

$$h(t) = \exp\left(\sum_{i=1}^{k} \beta_i x_i \; + \; \beta_{DE} x_{DE}(t)\right) h_0(t),$$

where $h_0(t)$ is the baseline hazard rate, and $h(t) = h(t;x)$ is the hazard rate for a failure at time t of a subject with explanatory variables, x_i. The variables, x_i, used in Attfield and colleagues (2012) are race/ethnicity, sex, and birth year and are time-independent. The variable $x_{DE}(t)$ is estimated diesel exposure and is allowed to vary with time. The baseline hazard function was also allowed to vary by state.

The primary measures of exposure used were cumulative exposure to REC (μg/m^3-yr) and average exposure to REC (μg/m^3). Each was evaluated at the same time as the failure event (unlagged), and up to 15 years prior to the failure event (lagged 15 years). They were included in the models either as categorical values or continuous values. When categorical, the groups were defined either by quartiles of exposure, distributed equally among the lung cancer deaths, or an expanded set of eight exposure categories. The expanded categories were logarithmically spaced in that they involved a doubling of exposure for each successive category and were the same across location worked, permitting a direct comparison by exposure level and location. Such a comparison is not possible using quartiles since the cut-points differ by location.

The most important assumption of the proportional hazards model is that the covariates have a proportional effect on the hazard function (Fisher and Lin 1999). A unit change in x_{DE} at time t changes the hazard by a multiplicative factor, $\exp(\beta_{\mathrm{DE}})$, whether that change occurs at time $t = 20$ or time $t = 60$ years. There are various methods for checking this proportionality of hazards over time. If the variable in question is time-independent, as are the x_i variables in Attfield and colleagues' model, then the proportionality assumption implies that the survivor curves over time for different levels of x_i cannot cross. This assumption can be checked as part of the analysis by estimating these curves nonparametrically, and plotting them. The investigators report that they have done such analysis, but the data were "not shown" (Attfield et al. 2012). As with any regression analysis, there are many choices for the measurement of exposure to DE, and the investigators used several different choices in their analyses.

Analysis using quartiles, or other quantiles of exposure is frequently used in epidemiology. Advantages of this approach are that it is relatively straightforward to implement and to communicate, and when a larger number of quantiles of exposure are used, it allows for some degree of nonlinearity in the exposure–response function. In a recent review of this approach, however, Bennette and Vickers (2012) have noted a number of potential disadvantages (e.g., use of unrealistic step functions in risk, loss of statistical power, potential inaccurate estimation of effects, difficulties in comparison of results across studies with different data-driven cut points, among others). See also discussions in leading textbooks in epidemiology (e.g., Rothman 2012).

Modeling exposure measurement as a continuous variable, as was done in this study, can be more useful for quantitative risk assessment, but essentially constrains the shape of the exposure–response function to be linear on the log scale under the proportional hazards model described above. Bennette and Vickers (2012) suggest nonlinear functions of exposure based on splines as a more flexible and appropriate approach to characterizing more complex exposure–response relationships, but note these are not without limita tions. These more complex functions can be useful for descriptive analysis of the exposure–response function, but cannot be easily used for prediction, and would be unreliable for prediction for exposures outside the range of the data. There are other intermediate modeling approaches that allow for some nonlinearity without being fully nonparametric (e.g., piecewise linear functions or polynomial regression) that could be considered as part of additional evaluations of alternative models for risk assessment.

Given that lung cancer is a disease that takes years to develop, time is an important factor to take account of in characterizing cancer risk. In modeling environmental exposures, the typical approach is to treat time in terms of exposure lags; whereby the most recent exposures are not considered. The lag can also be interpreted as the induction or latency period, the time between the initiation of cancer and its detection (Rothman and Greenland 1998). When exposures are unlagged, average or cumulative exposure is calculated on the basis of the full duration of exposure. In this study, both unlagged and lagged exposures were analyzed and presented. A 15-year lag was chosen for the primary exposure estimates, meaning that average or cumulative exposure was calculated from the start of exposure to the point in time 15 years before the death from lung cancer among cases, and 15 years prior to the relevant point of follow up among controls.

The investigators selected the 15-year lag for their primary analyses via a standard statistical approach. They systematically evaluated individual exposure lags ranging from 0 lag (unlagged) to 25 years at intervals of 2 years and compared changes in model deviance relative to a model that included no REC exposure, which they interpreted as a measure of model fit (Silverman et al. 2012 provides an analysis of model fit in their online supplemental material). They reported that the choice of a 15-year lag was supported in 7 of the 12 variations on exposure models they fit, although data for these were "not shown." The investigators continued to provide results for unlagged exposures as a sensitivity analysis. Although selected on the basis of statistical fit, the choice of a 15-year lag is within the range of the

lags or latency periods reported in other studies of the associations of lung cancer with exposures to other complex mixtures similar to diesel exhaust (Blot et al. 1983; Droste et al. 1999; Gustavsson et al. 2000; Morabia et al. 1992; Schoenberg et al. 1987).

4.2.3.1.3 Biologically Based Modeling: Multistage Clonal Expansion Models of Carcinogenesis Over the past few decades, a series of models have been developed that are grounded on concepts about the biological stages of carcinogenesis and can also take into account time-dependent patterns of exposure (e.g., Moolgavkar et al. 1993, 1999; National Research Council [NRC] 1993). The multistage clonal expansion model of carcinogenesis is built upon the notion that a first malignant cancer cell is formed within a tissue comprised of normal somatic cells following the occurrence of two or more critical mutations. Initiated cells that have sustained the first mutation may enjoy a selective growth advantage, resulting in an increasing population of such cells that may be transformed into a malignant cancer cell after sustaining the second critical mutation. Upon further uncontrolled division, this first cancer cell then leads to a malignant, ultimately clinically detectable, tissue mass.

Elegant mathematical descriptions of the multistage clonal expansion model of carcinogenesis have been developed based on the solution of stochastic cellular birth-death-mutation processes. Likelihood-based methods of fitting this model to toxicological and epidemiological data are also available, facilitating the application of the model in practice. Initial models were described for two mutations, and subsequently, multistage extensions of the model have also been developed. A three-stage clonal expansion model is discussed below in its application to the DEMS cohort data.

In practice, biologically based multistage clonal expansion models of carcinogenesis enjoy both strengths and weaknesses. Strengths of this approach to cancer risk modeling include the ability to interpret the model parameters in biological terms; the ability to describe complex time-dependent patterns of exposure to the agent of interest; and the theoretical generalizability of a validated biological model to exposure circumstances other than those associated with the data on which the model has been fit. In reality, because the two- and three-stage clonal expansion models depend on the estimation of parameters from a specific dataset, like empirical models, it is unclear that they are better than other models for extrapolating to other settings, particularly when the biological plausibility of the model is not well supported, as discussed below.

A potential weakness of such models is the possibility of over-parameterization. With two or more parameters required to describe each fundamental biological event (each mutational event, for example, requires estimation of the background mutation rate and the effects of exposure of the agent of interest on this mutation rate), simplifying assumptions (such as constraining the background mutation rates for different mutations to be equal), may be required for model identifiability. Fitting the model to epidemiological data may also be challenging, in that observation of incident cancer cases in population-based studies does not provide direct information on underlying mutation and cell proliferation rates. Fitting the model to cohort studies is easier than fitting it to case–control studies, as background mutation rates are not directly estimable from case–control data. Despite these limitations, biologically based models have been informative in cancer risk assessment with agents such as radon, x-radiation, and tobacco smoke (Krewski et al. 2003; Moolgavkar et al. 1993).

Moolgavkar et al. (2015) recently applied a three-stage clonal expansion (TSCE) model to the DEMS cohort data reported by Attfield and colleagues (2012). They presented their preliminary work at the HEI workshop on March 6, 2014 and in their recent publication (Moolgavkar et al. 2015). The TSCE is based on the notion that a malignant cancer cell arises from a normal somatic cell after it has sustained three critical mutations, any of which could be affected by exposure to REC. The TSCE also allows for promotion of initiated cells that have sustained the first two mutations, where promotion is defined as the increase in the net birth rate of the initiated cell population. Promotion increases cancer risk by increasing the pool of initiated cells available to undergo the third mutational event needed to complete the process of malignant conversion. After reparameterizing to ensure the identifiability of the parameters used to characterize the TSCE model, REC was seen to affect only the promotion rate of the initiated cell population, but none of the mutation rates included in the model (see Table III in Moolgavkar et al. 2015). Under this parameterization of the model, the three mutations are presumed to occur spontaneously, with the only effect of REC being to increase net birth rate of the initiated cell population.

In evaluating the biological plausibility of this model, it is worth noting that the IARC (2014) provided evidence that "diesel engine exhausts and the mechanisms by which they induce lung cancer in humans are complex, and no single mechanism appears to dominate." Key mechanisms cited by the IARC (2014) include genotoxicity (particularly DNA mutation), oxidative stress, inflammation, and cell proliferation. Other investigators have also provided possible evidence of electrophilicity (Arlt 2005), epigenetic alterations (Belinsky et al. 2002; Liu et al. 2008), immunosuppression (Bezemer et al. 2011; Diaz-Sanchez et

al. 1994), receptor-mediated effects (Furuta et al. 2008), and immortalization (Ensell et al. 1998; Shaw et al. 2011).

The TSCE appeared to provide a reasonable fit to the observed hazard function in the DEMS cohort (Moolgavkar et al. 2015, Figure 3). The investigators reported that the TSCE model "describes the observed hazard functions for the DEMS data well" and, as in the DEMS original investigators' analysis, that the "model with a 15-year lag fit the data substantially better than with no lag." The results of their analyses of the full cohort and of the ever-underground workers using the TSCE model note a "small, but statistically significant impact of REC … on the promotion of initiated cells, resulting in increased lung cancer mortality." Applying the model to specific mine types (limestone, potash, salt, and trona) produced significant results for the limestone mine ($P = 0.005$), near significant results for the trona mines ($P = 0.08$), but nonsignificant results for the potash and salt mines (see Table II, Moolgavkar et al. 2015).

The fitted model permits the exploration of risk associated with different temporal patterns of exposure. Figure 4 from Moolgavkar and colleagues (2015) illustrates how, for what appears to be an annual average exposure of 50 μg/m^3 starting at 20 years of age and continuing through age 40, the hazard function is elevated over baseline hazard rates, but slowly declines if exposure is stopped. This exposure amounts to a total cumulative exposure of 1000 μg/m^3-yrs, a level that falls within the highest quartile of exposure for ever-underground workers in DEMS. In this example, the hazard rate of the exposed individuals is predicted to approach the baseline hazard rate quite closely by age 90.

The TSCE model also permits an examination of other dose-rate effects. Moolgavkar and colleagues (2015) also examined the hazard profiles associated with a cumulative exposure of 50 μg/m^3-yr experienced at different rates starting at age 20 (5 μg/m^3 over 10 years, 2.5 μg/m^3 over 20 years, or 1 μg/m^3 over 50 years) for the entire DEMS cohort. Their results, shown in Figure 5 of their paper, suggest that for this scenario, the rate at which a given exposure is accumulated matters, and that the relative risk declines slowly once exposure is stopped (least slowly for the 1 μg/m^3 exposures over 50 years).

The Moolgavkar analysis (Moolgavkar et al. 2015) is a useful first step toward examining the sensitivity of the DEMS exposure response relationships to alternative modeling approaches. However, further work would be necessary to provide a more direct comparison with the analyses by Attfield and colleagues (2012). As noted above, and indeed in discussions at the March 2014 workshop, the TSCE model might best be interpreted as an alternative empirical model, not one that is superior by virtue of being biologically based (given the mechanistic evidence from IARC 2014 discussed above). The ability of the model to describe the dependence of risk on different temporal patterns of exposure is attractive. We note, however, that 50 μg/m^3-yr lies in the lowest quartile of cumulative exposure, lagged 15 years, for the ever-underground workers ($0 \leq 108$ μg/m^3), the reference category for the analysis (Table 4, Attfield et al. 2012). The rationale for this choice is not given in the paper, but seems to be directed at an evaluation of risk associated with ambient exposures. Since elevated risks are observed with this model even for these lower levels of exposure, it would be important to know what the results would be with higher levels of exposure observed in the DEMS. A more complete and transparent set of analyses with the TSCE model, including evaluation of a range of comparable levels of exposure in the DEMS, would provide a more direct basis for comparison of the findings with the DEMS and a more complete assessment of the time-related risks associated with exposure to REC for use in quantitative risk assessment.

4.2.3.2 Nested Case–Control Study

The Panel thought the statistical analyses were well described and appropriate for a case–control study design. The investigators' primary analysis relied on conditional logistic regression to estimate the risk of lung cancer mortality in the form of ORs. Their primary analyses included terms for exposure represented by quartile cut points for average REC intensity (μg/m^3) and cumulative REC (μg/m^3-yr) in which the ORs are estimated using the lowest level of exposure as a referent. Sensitivity analyses were conducted using quartiles of duration of REC exposure in years and the same eight expanded categories of average and cumulative REC exposure as used in the cohort study. The models also included terms to adjust for smoking, history of respiratory disease 5 or more years before date of death or reference date, and history of work for at least 10 years in an occupation with a high risk of lung cancer. These latter variables were designed to remove the effect of other occupational exposures that might also be associated with lung cancer. ORs were estimated for all subjects and separately for subjects who worked only on the surface and for those who had ever worked underground.

From a risk assessment perspective, models that represent exposure as a continuous variable can be more directly useful, as decision makers need to understand the risks associated with incremental changes in exposure. Silverman and colleagues (2012) also conducted an analysis of ORs associated with continuous REC exposures (d) using four different functional forms: 1) a log-linear model, $\mathrm{OR}(d) = \exp(\beta d)$; 2) a "power" model, $\mathrm{OR}(d) = d^{\beta}$; 3) a linear model, $\mathrm{OR}(d) = 1 + \beta d$; 4) and a linear-exponential model, $\mathrm{OR}(d) = 1 + \beta d \exp(\gamma d)$. All models were adjusted for the same set of potential confounding factors as described above (see

Figure 1 in Silverman et al. [2012] and the online supplement, with the exception of the results for the log-linear model, which was reported in the footnote to that figure to have a poor fit to the data).

The investigators have suggested that the steeper concentration–response at lower levels of REC exposure among surface-only workers and the plateaus observed at higher levels of exposure among ever-underground workers may be explained biologically by saturation of metabolic activation pathways, greater DNA repair efficiency, or greater nondifferential misclassification at higher exposure levels. As the reason for this behavior remains unclear, and since any application to quantitative risk assessment would ultimately need to rely on some model of continuous exposure or its approximation, the Panel agrees with the investigators that these results bear repeating in other studies. They could also be evaluated further with evidence from other studies as part of a quantitative risk assessment. On balance, however, the results from modeling exposure in different ways were broadly consistent and point to elevated levels of lung cancer risk with increasing exposures, with some observations reaching a plateau in the highest exposure categories.

4.2.3.3 Subgroup Analyses of the DEMS Cohort

In analyzing large complex datasets with multiple, possibly time-varying covariates for each individual subject, it is often of interest to examine subgroups that may be of particular interest with respect to the exposure circumstances, sociodemographic characteristics, or other factors that may modify the main effect of interest. However interesting the individual questions may be, the broader question concerns whether, or when, subgroup analysis of any type is valid. The goal of finding true signals in the data needs to be balanced with the need to avoid claiming false signals. The conventional approach to this is to set out the main methods of analyses in advance. On the other hand, limiting analysis to prespecified grouping risks missing important signals in the data, and in large, time-intensive and costly studies such as DEMS this is a particular concern. From Cox and Donnelly (2011):

> *"Even if the pre-specified methods have to be used it is, however, especially in major studies, crucial not to confine analysis to such procedures. This is for two rather different reasons. First careful analysis may show the initial method to be inappropriate. For example, … transformation of variables may be desirable to deal with non-linearity or with heterogeneity of variance. More important and controversially, the data, or experience gained while collecting the data, may suggest new research questions or even, in extreme cases, the abandonment of the original objectives and their replacement. The first reason, the technical inappropriateness of the original analysis, may not be particularly controversial. The second reason, a change of objectives, is more sensitive. In principle any conclusions of this kind require independent confirmatory study…. However the general point remains that, while an initial plan of analysis is highly desirable, keeping at all costs to it alone may be absurd."*

Useful discussions on approaches to subgroup testing have been provided by a number of investigators (e.g., Pocock et al. 2002; Wang et al. 2007). Wang and colleagues (2007) offer some guidelines for conducting and reporting subgroup analyses. Subgroup analyses may be guided by different analytic strategies.

1. One approach to subgroups analysis begins with an overall test of heterogeneity to determine if there are significant differences among subgroups; if heterogeneity is detected, comparisons among different subgroups may then be undertaken, using appropriate multiple comparisons methods designed to control the overall false positive rate. Wang and colleagues (2007) note that "a common mistake is to claim heterogeneity on the basis of separate tests of treatment effects within each of the levels of the baseline variable."
2. If interest focuses on the hypothesis that a particular subgroup may drive the results for the entire group, a classical *leave-out-one* influence analysis may be conducted. As the name implies, this approach involves leaving out one subgroup at a time, and examining the consistency of the findings based on the remaining subgroups to the exclusion of a particular subgroup.
3. In some cases, there may be a subgroup hypothesis that is of particular interest a priori, with testing of that hypothesis specified in advance. In this event, this may be considered as a main hypothesis of interest, rather than a hypothesis formulated based on data-driven subgroup analysis.
4. Finally, post hoc analyses guided by ongoing analyses of the dataset of interest may become of interest during the course of the analysis. While such analyses can be informative, care should be taken to control for multiple hypothesis testing, and to not over interpret the results of multiple data-driven hypothesis tests. The results of such post hoc analyses are often regarded more as hypothesis generating than as hypothesis testing.

The DEMS was designed to create a cohort of sufficient size and range of exposures to allow investigation of the

relationship between exposure to diesel exhaust and lung cancer in workers employed in nonmetal mines in the United States. Although the DEMS investigators conducted and reported results for a number of sensitivity analyses of their data (e.g., limiting the range of exposure to < 1280 and tenure exclusions in Attfield et al. 2012), the Panel focused its discussion on two decisions in particular that have received particular scrutiny in the HEI public workshop and in published critiques: 1) the analysis of the cohort by location of work; and 2) analyses of the data by mine type. The Panel considered the appropriateness of the analyses performed in these two areas and their relevance to evaluation of the studies for quantitative risk assessment.

4.2.3.3.1 Analyses of Data by Work Location The original 1997 protocol (NCI–NIOSH 1997) for the DEMS called for analysis of the complete cohort, without regard to whether or not the work location was on the surface or underground. However, the investigators found "different patterns of lung cancer mortality by location had obscured exposure–response in the complete cohort" (Attfield et al. 2012). Further analyses in the cohort were carried out in the two groups of workers separately. In the case–control study, a variable for location, in combination with variables for smoking, was introduced to the models (Silverman et al. 2012). Crump and colleagues (2015) have conducted analyses in the case–control study with the group of workers that only worked underground under the presumption that these were the workers "most heavily and consistently exposed" to diesel exhaust. Moolgavkar and colleagues conducted analyses with this subgroup in the cohort study.

The DEMS was designed to detect the effects of interest in the full cohort, so analyses of subgroups within it require both a sound rationale and careful interpretation. The investigators' decision to analyze surface-only workers separately from the ever-underground workers is an example of subgroup analysis that was motivated by the data, and in particular large disparities in the diesel exposure levels between the two groups, differences in smoking patterns, and apparent differences in lung cancer rates between surface-only and ever-underground workers. This is arguably what Cox and Donnelly (2011) call a "technical" rationale (as opposed to a change of objectives), although the distinction between these two is not entirely clear cut. These separate evaluations may also be of greater utility for meta-analyses designed to compare and analyze risk estimates over a broader range of exposures (e.g., Vermeulen et al. 2014b) and for considering the generalizability of the results to other populations with similar levels of exposure (e.g., general population versus worker populations).

The Panel found the basis for analysis of the case–control data of the workers who only worked underground to be less compelling. While the same rationale was put forward as that given for separate analysis of the surface-only and ever-underground workers (a large difference in exposures), the reality was that the cumulative exposures of the ever-underground and only-underground were of a similar magnitude. Comparison of the cumulative exposures for the ever-underground (Table IV) and only-underground (Table V) subjects for each of the alternative exposure models shows that the cumulative exposures were similar in the two groups, if not lower among those who had only worked underground. Although average area level exposures were higher underground than above ground, the lower cumulative exposures likely depend on shorter time periods spent only working underground jobs.

Analysis of lung cancer risks based on further division of a group whose exposures are comparable would be expected to reduce the precision of the estimates, and indeed this is what the results show (compare the ORs in Table IV for the ever-underground with those in Table V for only-underground in Crump et al. 2015). In addition, the mean ORs are also shifted toward the null value of no effect. Moolgavkar and colleagues (2015) provide similar findings in their analyses of the cohort data in Table S1 of their supplementary materials. Since quantitative risk assessments should rely on the data yielding the most accurate and precise estimates of risk, analysis of the more complete set of ever-underground subjects represents the more logical choice for analysis.

4.2.3.3.2 Mine by Mine Analyses There has been considerable interest in whether individual mines or mine types can explain or drive the results observed in the analyses of the cohort as a whole. In secondary analyses of the case–control data, Silverman and colleagues (2012) investigated whether the effects of REC differed by mine or ore type, focusing on the potash and trona mines where the most workers were employed (see Table 7 in that publication). Using their main models, adjusted for smoking and other covariates, they found significant effects of both average and cumulative REC on lung cancer risk were noted in both the potash and trona mines. These investigators suggested that too few workers were employed in the single salt mine and single limestone mine to support meaningful analyses of these two subgroups.

Attfield and colleagues (2012, Supplementary Table 14) also examined the association between diesel exhaust and lung cancer risk within individual mine types in the DEMS cohort study. Without adjustment for multiple testing, significant or near significant associations between cumulative

REC (excluding exposures < 1280 μg/m^3-yr) and lung cancer risk were noted for the limestone, potash and trona mines; average REC was only significant in the limestone mine, regardless of tenure exclusion. These analyses were not adjusted for smoking, as smoking data were only available for subjects in the nested case–control study analyzed by Silverman and colleagues (2012).

Using data from the DEMS nested case–control study, Crump and colleagues (2015) conducted a classical leave-out-one analysis to evaluate the potential influence of a single mine on the overall results. Lung cancer ORs associated with cumulative exposure to REC were insensitive to the exclusion of any one mine, with all ORs remaining significant, regardless of which mine was excluded from the analysis. These analyses were conducted using the original DEMS REC estimates, but with the set of variables included in Crump and colleagues' "with radon" models, which differ from those of the original investigators (see discussion in Section 4.3.2.1).

In a reanalysis of the DEMS cohort data, Moolgavkar and colleagues (2015) reported a significant association (P = 0.05) between the logarithm of REC and lung cancer risk only in the limestone mine, and in the entire cohort and marginally significant results in the trona mines; similar results were obtained using both the TSCE and CPH models. Significance of exposure–response was assessed using likelihood ratio tests within each mine type, and for the entire cohort across mine types. The investigators suggest their findings should be interpreted to show that the "exposure–response parameter for the entire cohort is driven by the limestone mine" and that the increase in lung cancer mortality "appears to be confined" to that mine (Moolgavkar et al. 2015). They conclude that the questions raised by their results mean that "the DEMS data cannot reliably be used for quantitative risk analysis."

Over the course of its evaluations, the Panel considered this analysis and the conclusions of Moolgavkar and colleagues (2015) carefully and offers a different perspective. First, despite the original design of the cohort based on all the mines, the Panel understands the interest in testing the hypothesis that there might be an unmeasured confounder associated with the mine, its operations, or ore type, for example, that might explain different lung cancer rates. Both the original investigators and other analysts (e.g., Moolgavkar et al. 2015) have pointed out that the limestone mines have operations that are quite different than those in the other mines (i.e., they use high-HP engines to move the ore laterally through tunnels that are naturally, not mechanically, ventilated). However, no hypothesis has been put forward about the nature of the resulting diesel exhaust exposures or other factors in the limestone mine that might be expected a priori to yield different REC–lung cancer results. Exposures to silica, asbestos, nondiesel PAHs, radon, and respirable dust were all comparable to, or lower than, those in other mines.

Returning to the guidelines for post hoc statistical analyses, the Panel considered Moolgavkar and colleagues' (2015) analysis of mine type using proportional hazards models (the clonal expansion modeling results were not presented in sufficient detail to analyze). In the Panel's view an appropriate test of homogeneity or "equality of exposure–response parameters across mine types" (Moolgavkar et al. 2015) had not been done using either the clonal expansion models or the proportional hazards models. Moolgavkar and colleagues (2015) conducted tests of parameter equality between the limestone mine type and the other three mine types aggregated into a single group. An appropriate test would treat all four mine types separately and independently without separating any of them out for post hoc testing.

Second, because there are four mine types, an analysis that compares limestone to the other three mine types is one of four possible selections, and it is conventional in these post hoc analyses to make a correction for multiple testing, in effect requiring much stronger evidence in comparisons that have been selected in light of the data. Moolgavkar and colleagues (2015) did not correct for multiple comparisons. Given the estimate of the study-wide exposure–response coefficient, and its standard error, it is not at all surprising that the statistical significance could be driven by one mine type.

Finally, from the standpoint of the low levels of REC exposures in the limestone mine, and the steeper exposure–response relationships at low exposures in the continuous modeling, it is not surprising that the limestone mine (Mine A) would have a strong influence on the overall exposure–response relationships in the mines. Figure 4.2 indicates that exposure levels to the loader operator were typically lower in this mine in early years of dieselization (up to 1970) than at many of the mines. In addition, the mean year of hire at this mine was 1967. The period from 1967 up to 1976 (i.e., 1967 plus the mean employment duration of 9 years) covers a period when this mine certainly had some of the lowest exposures of the eight mines. Given that probably about half the workers were employed before 1967 (based on the mean year of hire being 1967), many would have had very low exposures during this time.

Quantitative risk assessment takes place after a determination that there is a relationship between exposure and adverse health outcomes. The suitability of an epidemiological study for use in quantitative risk assessment hinges on whether or not it provides suitably accurate estimates of an exposure–response relationship. Whether or not the

exposure–response relationship is the same in subgroups of the data is not relevant; the relevant quantities are the estimate of the exposure–response curve and the estimate of its standard error. While the strength of the exposure–response relationship across all mine types is increased by including the limestone mine, it is decreased by including the trona mines. Given the larger numbers of workers in the potash and trona mines, their inclusion had substantial influence on the precision of the study-wide estimate. These findings in fact strengthen the case for the use of the entire cohort study for estimating the REC–lung cancer relationship and for quantitative risk assessment.

4.3 CONTROL FOR CONFOUNDING FACTORS

As with any observational epidemiological study, there is always the possibility of unmeasured confounding. For a study of lung cancer, the most obvious candidate for such a confounder is smoking, whether active smoking by the individual or exposure to secondhand smoke, either at home or in the workplace. Working populations may also have been exposed to other pollutants thought to be lung carcinogens, either in the current mines or in previous workplaces. This section examines the efforts taken by the DEMS investigators to address these potential confounders, both in the design of the study and in their analyses. In DEMS, individual smoking and occupational histories were obtained in the case–control study and so will be discussed in that context.

Two important questions must be answered regarding use of the DEMS results for quantitative risk analysis: 1) Do any limitations related to characterization of smoking or other occupational exposures suggest the potential for significant confounding of the REC–lung cancer association? 2) Is there plausible evidence of effect modification (for example, different levels of REC–lung cancer risk at different intensity or duration of smoking) that would need to be considered in using these results in risk assessment applications?

4.3.1 CONTROL FOR SMOKING

To evaluate the potential effect of smoking on lung cancer outcomes in the study, the investigators had first used computer-assisted telephone interviews to collect detailed questionnaire data about individual and family medical history, diet, and lifetime occupational exposures, but particularly about smoking from a subset of the cohort. More specifically, they interviewed individuals (self), or where necessary, their next of kin (proxy), for 198 lung cancer deaths (cases, all by proxy) and 562 matched controls (222 self and 340 proxy interviews). The Panel first examined descriptive statistics on smoking patterns that illustrated some differences in smoking behaviors among cases and controls and by location that warranted formal examination in the statistical models. The descriptive statistics point to higher rates and intensity of smoking among those who had only worked on the surface. For example, the percentage of subjects reporting they had "ever smoked" (i.e., smoked 100 cigarettes or more), was about 41% higher in cases than in controls (93% versus 66%) for those who worked only on the surface compared with roughly equivalent percentages between cases (73%) and controls (78%) among those who had ever worked underground. A higher percentage of cases than controls were reported to be regular smokers among surface-only workers than among ever-underground workers. Finally, a higher percentage of cases relative to controls were categorized as heavier smokers (i.e., smoking either 1 to < 2 packs per day or 2 or more packs per day) among surface-only workers (24% vs. 8%) than among ever-underground workers (21% versus 14%).

Silverman and colleagues (2012) analyzed the effect of smoking in the case–control data:

1. by estimating lung cancer ORs for categories of a combination smoking variable consisting of status (never, former, current, and unknown) and intensity (packs/day). ORs were estimated for all subjects and by work location (surface-only or ever-underground). (See Table 2 in Silverman et al. 2012.)
2. by estimating lung cancer ORs by quartile of REC exposure in the complete dataset adjusting for smoking and location using a categorical combination variable (smoking status/intensity and location). (See Table 3 in Silverman et al. 2012.)
3. by estimating lung cancer ORs in ever-underground and surface-only workers separately using the categorical combination smoking variable (status/intensity). These required some exclusion of unmatched subjects. (See Tables 4 and 5 in Silverman et al. 2012.)
4. by estimating lung cancer ORs by tertile of cumulative REC, lagged 15 years and by category of smoking intensity (packs/day) alone, adjusted for mine location. (See Table 6 in Silverman et al. 2012.)

When the complete set of cases and controls were analyzed by smoking status/intensity and adjusted for cumulative REC, the results showed statistically increased odds of lung cancer with smoking (Table 2, Silverman et al. 2012). However, they observed an apparent interaction with location. The exposure–response relationship between smoking and lung cancer among the surface-only miners was more consistent with the expected relationship, being strong and monotonically increasing with the amount smoked. It was less consistent in ever-underground miners, with apparent

attenuation for smokers of 1 to < 2 packs per day and ≥ 2 packs per day, regardless of smoking status (former/current). In their analysis of the ORs for lung cancer and tertiles of cumulative REC, lagged 15 years for different levels of smoking intensity, they found that ORs were on average higher at increasing tertiles of REC exposure for all but those who smoked 2 packs or more of cigarettes per day (Table 6, Silverman et al. 2012).

While statistically significant, these latter findings provided an indication that there might be complex interactions between either DE exposure or another characteristic of the underground mining environment and the smoking–lung cancer relationship. One explanation may simply be the higher rates of smoking among surface workers discussed earlier; among never smokers, risks after adjustment for 15-year lagged cumulative were similar in ever-underground workers to those in surface-only workers (OR = 0.90; 95% CI = 0.26 to 3.09, Table 2, Silverman et al. 2012). Speculation about the biological basis for the direction of this interaction at higher levels of exposure requires extrapolation beyond the knowledge available in this study, and the Panel chose not to evaluate this issue further here.

In any event, the investigators chose to deal with this interaction in their main analyses of the complete case–control data set by adjusting for smoking with an indicator variable representing a combination of smoking status, intensity, and location. The investigators used an indicator variable to represent each combination for each of the two mining types (surface vs. underground), smoking status (current, former, never, or unknown/occasional), and each of three smoking intensities (< 1 pack per day, 1 to < 2 packs per day, ≥ 2 packs per day among current or former smokers) in their models. All of their models also adjusted for history of respiratory disease and for history of work in high risk jobs. Their main approach is summarized in the second column of Table 4.3 (Silverman et al. 2012).

4.3.1.1 Self Versus Proxy Reports

One question that could be raised pertains to the quality of the self-versus-proxy data on smoking. Since the investigators were able to obtain smoking data from a high percentage of cases and controls or from their next of kin, the primary concern might be whether there are systematic differences in the accuracy of smoking information provided (i.e., differential recall bias by individuals versus next of kin) that could influence the study findings. For example, some evidence suggests that individuals tend to underreport their actual amount of smoking and that the misclassification of smokers as nonsmokers is greater in higher smoking categories (e.g., 0.8% to 2.8% among occasional smokers versus 6% to 15% for regular smokers [Wells et al. 1998]). In DEMS, this source of underreporting of smoking in self reports would likely affect only the controls (as all smoking for cases was collected by proxy interview) and would tend to dampen the effect of REC exposure on lung cancer in the study. However, other evidence suggests that proxies underreport smoking by the subject (Soulakova et al. 2009), and that in some cases underreporting may be differential with respect to cancer diagnosis, with next of kin underreporting the amount smoked in decedents with cancer (Steenland and Schnorr 1988). In a study like DEMS where all smoking data for cancer cases are from proxies, this kind of bias could lead to an underadjustment for smoking and an upward bias in the "true" REC–lung cancer effect.

Given that there was no way to compare self and proxy responses for the same individual, the investigators took the reasonable step of determining whether direct versus next of kin interviews in control subjects gave similar percentages for several important variables, including smoking (these comparisons are given in Silverman et al. 2012). For many comparisons of smoking categories, the results were similar between cases and controls (for example, the percentages of never, occasional, or former smokers of less than one pack per day). More sophisticated analyses could be done on the impact of differential responses by subjecting the adjustment for smoking to a probabilistic uncertainty analysis for response bias (i.e., missing data bias) and measurement error; however, the Panel thought that it was unlikely to result in major changes in the results.

4.3.1.2 Alternative Smoking Analyses by the HEI Panel

The main focus of the HEI Panel was on the investigators' approach to adjustment for smoking in the main analyses, specifically on: 1) the measure of smoking used and how it was incorporated in the models, and 2) how they evaluated the potential interaction between smoking and location.

4.3.1.2.1 Alternative Smoking Metrics The Panel wanted to see the impact of adjusting for smoking using alternative smoking metrics. The Panel's goal was to evaluate the sensitivity of the REC–lung cancer effect to different metrics, rather than to resolve the debate on what smoking metrics are most scientifically and biologically appropriate (see for example, exchanges between Peto [2012, 2013] and Lubin and Caporaso [2013] and related publications). As it turned out, the investigators had already examined duration and packyears of smoking (e.g., years) as part of their original work but had reported that the results were not sensitive to this decision (and were thus not shown in the published paper). Subsequently, the DEMS investigators published a detailed letter to the editor showing REC–lung cancer results

Table 4.3. Comparison of Silverman et al. and HEI Panel Analyses Controlling for Smoking

Published[a]	Silverman et al. 2012	Silverman et al. 2014	HEI Panel
Descriptive analyses	**Table 1:** Cigar smoking, years Pipe smoking, # of pipefuls/ week Number of smokers in participant homes **Table 2:** Status/intensity[d] by location Status intensity by case or control status	Discussed in original study	**Additional Materials 3:**[b] Cigarette smoking variables[c] by the case–control status, location of employment (surface-only/ever-underground), and proxy status self/proxy)
OR analyses	**Exposure metric[e]:** Average REC, lag 0, 15 yr Cumulative REC, lag 0, 15 yr Duration of REC exposure (yrs) **Smoking status:** Never, former, current, unknown[f]; Intensity **Interactions:** None (smoking status and work location were combined in the analysis)	**Exposure metric:** Average REC, lag 15 yr Cumulative REC, lag 15 yr **Smoking status[g]:** Status–duration; Status–packyears; Status–packs/day and duration **Interactions:** None (Smoking status and work location were combined in the analysis)	**Exposure metric:** Average REC, lag 0, 15 yr Cumulative REC, lag 0, 15 yr **Smoking status[h]:** Status–duration; Status–packs; Status–pack-years; Status (never, former, current, unknown) and Duration (as a continuous variable); Status and Packs/day (as a continuous variable); Status and pack-years (as a continuous variable) **Interactions:**[i] location of employment (ever-underground / surface-only) and duration, packs/day, and pack-years as continuous variables

[a] These indicate only the analyses that were available to the Panel for its review in the published literature, not that they were not done by the original investigators.

[b] Additional Materials 3 is available on the HEI Web site.

[c] Smoking variables included: smoking status, smoking intensity (packs per day), smoking cessation, smoking status and packs per day, smoking duration, smoking status and duration, pack-years, smoking status and pack-years, and a combination smoking–location variable. See Appendix C for details.

[d] Intensity refers to packs per day.

[e] For all adjustments, covariates, and other information about the models see the footnotes of Table 3 in Silverman et al. 2012.

[f] Unknown category includes occasional smokers.

[g] Status includes never smoker, former smoker, current smoker, unknown.

[h] Status includes never smoker, former smoker, current smoker, unknown. Analyses were also run excluding "unknown."

[i] The Panel had hoped to examine interactions with other variables but found that, when the data set was broken down by some of these categories, there were too few subjects in some categories to do so reliably.

adjusted for smoking using three other measures of intensity — packyears, duration, and packs per day + duration, again in combination variables with location as in the original models (see details in the third column of Table 4.3, Silverman et al. 2014).

4.3.1.2.2 Evaluation of Potential Interactions Between Smoking and Work Location While the Panel considered the investigators' choice of modeling with combined indicator variables to be reasonable (it has the advantage of providing estimates of risk for each combination of smoking level and location worked relative to a common referent [i.e., nonsmoking surface workers]), it is not the most intuitive modeling strategy. The Panel wanted to examine the implications for the results when relevant components of the smoking histories (packs per day, age started smoking, duration, time since cessation) were modeled separately. The Panel would then able to explore interaction effects more directly by including specific interaction terms between location and smoking.

There are multiple objectives for such an approach, specifically: 1) to examine some additional smoking variables that might not have been examined already (including continuous [e.g., cigarettes per day], rather than categorical versions); 2) to improve interpretability of the smoking variables themselves, separate from the impact of location; and 3) to examine more explicitly the interplay between smoking and location than done in the published work. Silverman and colleagues [2012] had suggested some degree of interaction between smoking and location in some of their analyses: "The addition of a variable representing the interaction of location worked and smoking to models statistically significantly improved analogous models that included smoking without location (*P* values for the likelihood ratio tests ranged from 0.011 to 0.064 for average REC intensity and cumulative REC, unlagged and lagged.)" Finally, the Panel thought such analyses might improve understanding of the utility of the DEMS data or results for quantitative risk in other settings and populations where specific smoking patterns or categories might be different from those in the mines.

The last column of Table 4.3 lists the alternative models explored by the Panel for this project. The essential difference between the Panel's and the original investigators' analyses was that the various smoking metrics were included as separate variables from location and were modeled as categorical variables. As sensitivity analyses, the Panel also modeled smoking exposure using continuous versions of the same variables and explored the impact of excluding subjects with missing or unknown smoking information. All models also included variables to adjust for history of respiratory illness and of high risk jobs as in the original analyses. Despite its original goal of modeling all variables separately, the Panel found that it also had to use smoking variables that were a combination of smoking status (never, current, ever, and unknown) and intensity (duration, packs per day, or packyears) in order to avoid overparameterization of the models. The Panel was unable to analyze the data by age at start or time since cessation of smoking, because missing data reduced the numbers of subjects and made the analyses less reliable. The Panel's modeling approaches are described in more detail in Appendix D.

4.3.1.2.3 Comparison of Results for Alternative Smoking Analyses Figures 4.3 and 4.4, based on 15-year lagged average and cumulative REC, respectively, compare the mean ORs with 95% CIs and *P* values for tests of trends for the original DEMS analyses, for the investigators' additional analyses, and for those conducted by the HEI Panel. The *P* values for trend were calculated using the same method as the original analyses (two-sided Wald test). Data for these and related analyses in Table 4.3 can be found in Appendix Table D.2.

In each figure, the first set of results comes from the DEMS model of Silverman and colleagues (2012), which adjusted (controlled) for smoking using a combination variable consisting of smoking status (never, occasional, ever, and unknown), intensity (packs/day), and location (surface-only vs. ever-underground). The second group of models is the result of alternative adjustments for smoking provided by the investigators in a recent update (Silverman et al. 2014) and include a combination variable of smoking status, location, and either smoking duration (years), packyears, or duration and packs/day. The third group of boxplots displays results for the Panel's models where smoking was included as a variable separate from location. Smoking is characterized as the combination of smoking status and either duration, packs/day, packyears, and packs/day with duration. Both the DEMS and the Panel models adjust for history of respiratory illness and history of high risk jobs as separate variables.

Comparisons of the Silverman and colleagues' analyses (2012, 2014) show that the magnitude and trend in ORs for lung cancer with increasing REC exposure were robust to different choices of smoking metrics and to whether they were included as separate or combination variables. Cumulative REC lagged for 15 years, in particular, consistently showed statistically significant increasing trends in cancer risk with increasing REC exposure in the both sets of analyses. Since inclusion of measures of smoking separately in the models from location yielded similar results to the combined indicator variables, the Panel concluded that the approach could be used equally well for quantitative risk assessment.

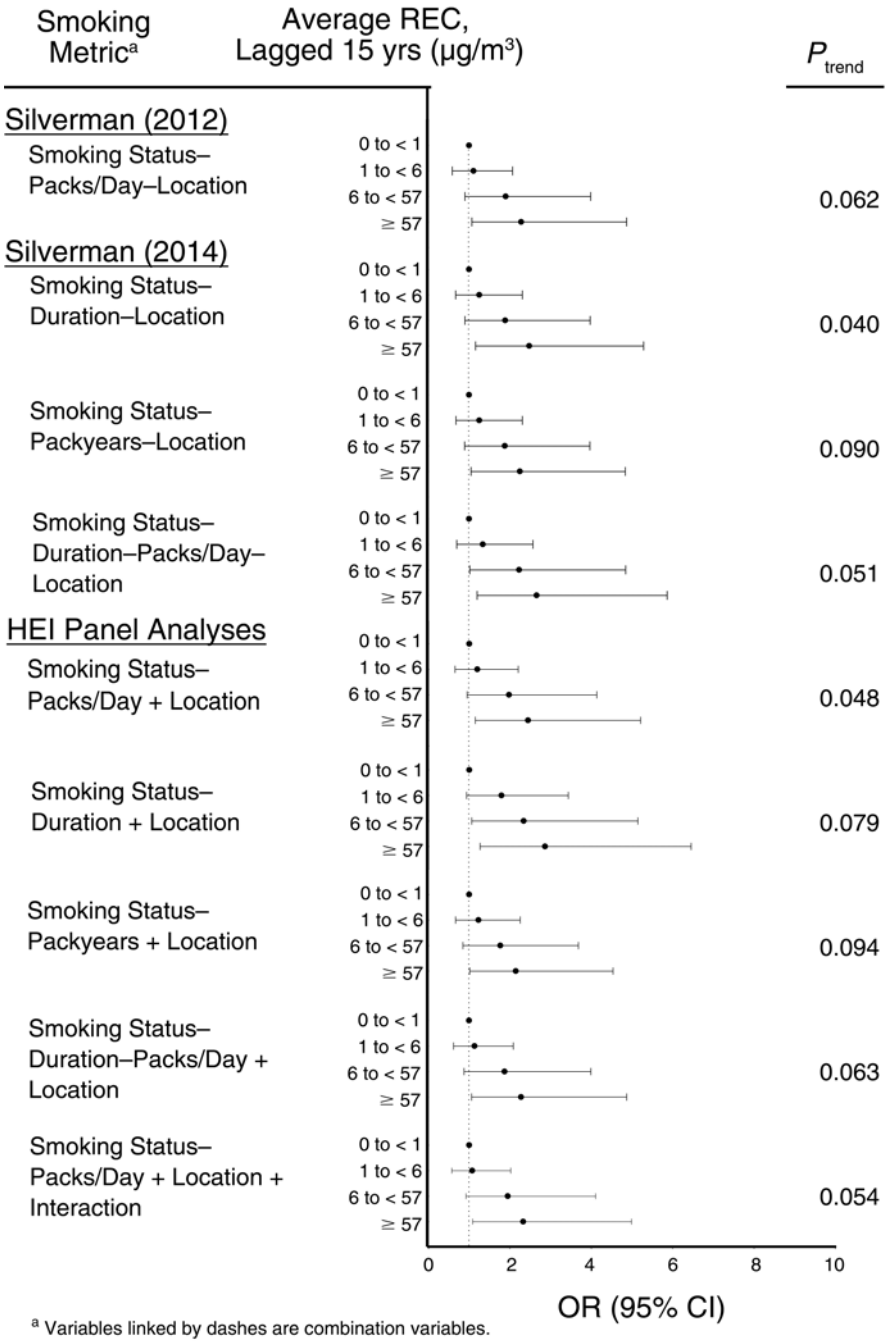

Figure 4.3. Comparison ORs for Lung Cancer and Average REC, lagged 15 years, in all subjects, using Alternative Smoking Metrics. Each box plot represents the maximum likelihood value (•) and 95% CI (whiskers) of the OR. For each smoking metric, combination variables are linked by dashes (–); individual variables included in the model are indicated by plus signs (+). The *P* values for 2-sided Wald tests of linear trend (using median ORs assigned to each subject in a quartile) are shown to the right. Both the Silverman et al. (2012, 2014) and HEI Panel models adjust for history of respiratory illness and history of high risk jobs as separate variables. Details of the models can be found in Appendix D.

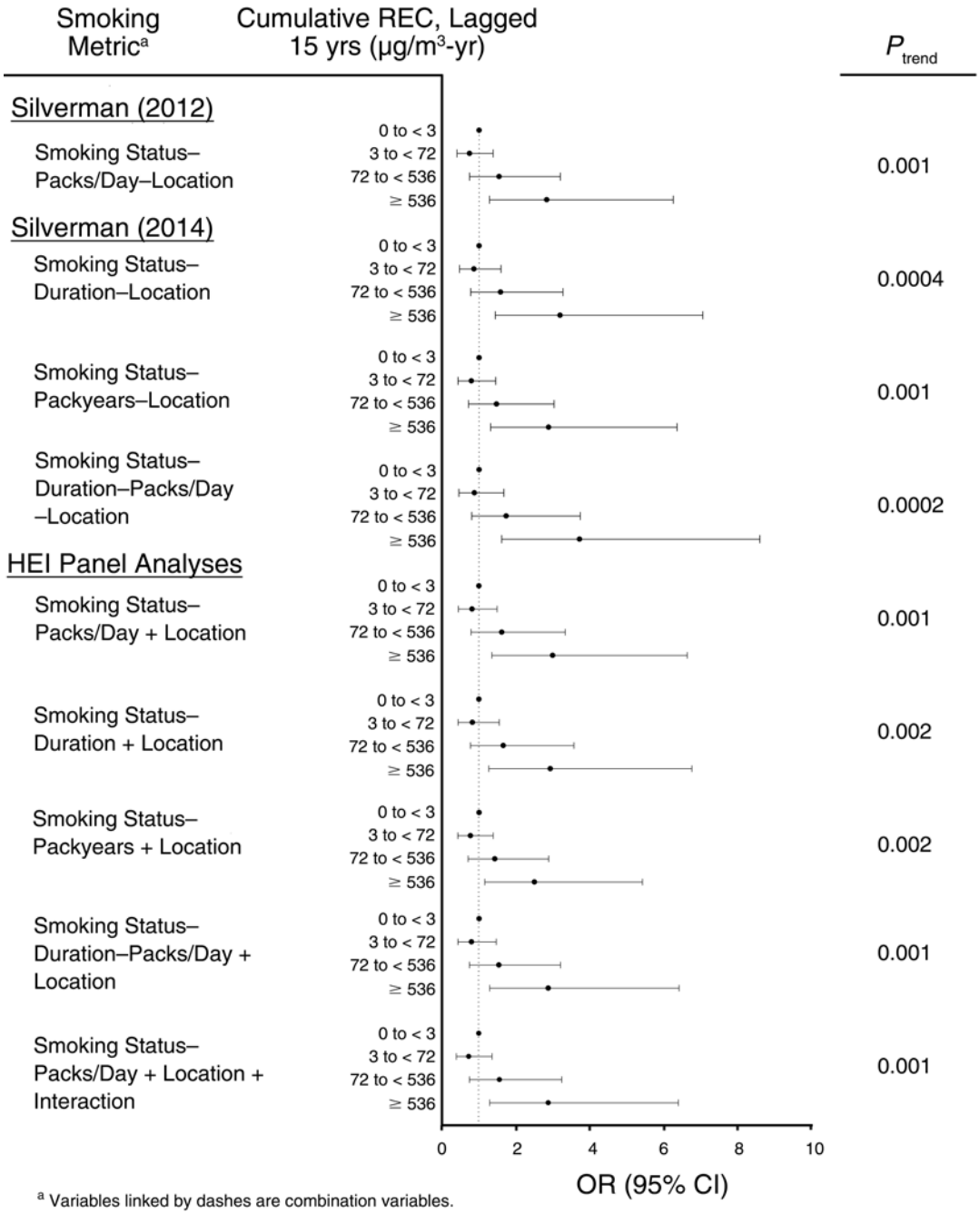

Figure 4.4. Comparison of ORs for Lung Cancer and Cumulative REC exposures, lagged 15 years, in all subjects, using Alternative Smoking Metrics. Each box plot represents the maximum likelihood value (•) and 95% CI (whiskers) of the OR. For each smoking metric, combination variables are linked by dashes (–); individual variables included in the model are indicated by plus signs (+). The *P* values for 2-sided Wald tests of linear trend (using median ORs assigned to each subject in a quartile) are shown to the right. Both the Silverman et al. (2012, 2014) and HEI Panel models also adjust for history of respiratory illness and history of high risk jobs as separate variables. Details of the models can be found in Appendix D.

The Panel's analyses exploring the interactions between smoking and work location found that the interaction terms were not significant in most models, with the exception of those that relied on smoking measured in packs per day as a continuous variable (subjects with unknown smoking data excluded), with or without smoking status as separate categorical variables. To illustrate this point, the last set of boxplots in Figures 4.3 and 4.4 shows results for the models using smoking in packs/day as a categorical variable, smoking status, and interaction terms for smoking by location. Details for the all the interaction analyses conducted and their results may be found in Appendix D and Table D.2, respectively.

For REC–lung cancer models including smoking status as well as the continuous smoking measures in pack-per-day, the parameter estimate for the interaction term was −0.64 (P = 0.012, based on Wald chi-square statistic) for average REC lagged 15 years and −0.66 (P = 0.11) for cumulative REC lagged 15 years. The results for the models indicate some residual interaction between smoking and location, where risk of lung cancer was higher among ever-underground workers than among surface-only workers. However, a crude model including a variable for location alone indicated that location itself was not a significant predictor of lung cancer risk (OR 1.041 [0.741 to 1.463]; P = 0.817, where the OR compares ever-underground workers to surface-only workers as the reference group).

Collectively, these analyses alleviated concerns that the investigators' choice of modeling approaches provided results that suffered from model selection bias. The Panel's analyses suggest that effects observed by the original investigators were robust to modeling choices both with regard to how to characterize smoking exposure and how to understand the interaction between worker location and smoking on lung cancer case status. The methods of adjusting for smoking in characterization of the REC–lung cancer association in this study were appropriate (in particular the choice of packs-per-day as a measure of smoking). The lingering suggestion of some effect modification by location or by level of REC — that the risk of lung cancer from cigarette smoking would differ for surface workers compared with underground workers — remains somewhat nonintuitive and challenging to interpret. This issue merits further evaluation in the context of the broader scientific literature.

4.3.2 CONTROL FOR CONFOUNDING BY OTHER OCCUPATIONAL EXPOSURES

The other source of potential confounding is the presence of other exposures to pollutants in the mines that have been associated with lung cancer — silica, asbestos, nondiesel PAHs, radon, and respirable dust. As discussed earlier, the DEMS investigators first dealt with the potential for confounding occupational exposures in the fundamental design of the study. They selected the eight nonmetal mines for inclusion in the study because the available data indicated that levels of all of these exposures were likely to be low. Attfield and colleagues (2012) provide a summary of the concentrations of each of these exposures for the full cohort and by work location on the surface or underground for each mine. (This summary has been reproduced in Appendix Table C.2.) Although low, the exposures to some contaminants were on average higher for ever-underground workers than for workers who worked only on the surface, so the potential for confounding of the REC effect needed to be assessed.

The DEMS investigators evaluated the effect of each of these other exposures on the REC–lung cancer relationships in both the cohort and case–control studies. For the cohort study, Attfield and colleagues (2012) reported that inclusion of cumulative exposure to silica, asbestos, nondiesel PAHs, and respirable dust [individually] in their models led to small increases ("5% overall") in the HRs for REC and lung cancer. They did report that cumulative radon exposure did have some effect on lung cancer risks in workers with long tenures in particular, an effect that could be eliminated by exclusion of those workers from their analysis. In the case–control study, Silverman and colleagues (2012) also constructed and evaluated models in which each occupational exposure was included as an additional covariate. They considered an occupational exposure to be confounding only if its inclusion in the exposure–response model led to a greater than 10% change in the resultant OR relative to that in their main model. They reported that none of the potential confounders reached this threshold and therefore included none of these variables in their final models.

4.3.2.1 HEI Panel Analyses of Adjustment for Radon

Concerns were raised at the HEI Diesel Epidemiology Workshop in March 2014 and subsequently in a publication by Crump and colleagues (2015) that this approach may not have appropriately accounted for confounding by radon. Radon is the second leading cause of lung cancer in the United States after tobacco smoke (NRC 1999). Radon is a potential confounder in this study because it is both associated with lung cancer and correlated with cumulative and, to a lesser extent, average REC. (See Appendix Table E.1, which shows that the mean radon exposures in working level month [WLM] were slightly higher for cases than for controls; and Appendix Table E.2, which shows the correlations between radon and REC exposures.) Thus, the possibility that radon may have contributed to a portion of the lung cancer burden among miners who ever worked underground needs to be considered.

The HEI Panel first conducted descriptive analyses of the radon data available to the investigators to assess the levels found both by mine and by location in the mine, as well as the levels of exposure experienced by cases and controls. To put these findings into perspective, the Panel next estimated the excess relative risk of lung cancer for radon in the mines. The Panel then reanalyzed the case–control analytic data set using the investigators' original models to estimate lung cancer ORs for diesel exhaust, adjusting for radon in various forms. The main findings are discussed below; details of the Panel analyses are presented in Appendix E.

The DEMS investigators had limited data on radon levels in the mines over the study period. Specifically, they had 28 measurements taken as part of the 1998–2001 DEMS survey and 251 measurements taken by MSHA from the 1970s to the 1990s (i.e., the MIDAS survey). Radon levels measured in each mining facility in the DEMS survey were all below the LOD, which ranged from < 0.01–0.07 working levels (WL). Of the 251 MSHA radon measurements in the MIDAS survey, 54% were below the LOD.

Table 4.4 summarizes the data available from the MIDAS survey by mine, including the percentage of measurements below the LOD, the mean detected radon values in picocuries per liter (pCi/L), and the mean radon levels in WLs. These data were made available online by NCI: *http://dceg.cancer.gov/research/what-we-study/environment/diesel-exhaust-miners-study-dems*). The percentage of nondetects varied notably from mine to mine (from 16% in Mine A [limestone] to over 80% in Mine I [trona]). For samples reported as nondetected, the investigators imputed values to those samples by dividing the LOD by $\sqrt{2}$ and by a percentage (80%) to adjust for equilibrium of radon daughters. Using these data, the DEMS investigators estimated the mean and 95% confidence limit on the underground radon level for each mine in WLs (a WL corresponds to about 200 pCi/L). As the table shows, the mean WLs were not highly variable: for the ever-underground workers, they ranged from approximately 0.01 WL for facilities A, H, I, and J to 0.02 WL for facilities B, D, E, and G (see Appendix Table C.2 for radon statistics by mine and work location).

The cumulative radon exposure level assigned to individual miners was in units of WLMs and was the product of the facility-specific mean WL, the years spent in individual jobs, taking into account whether the job was above ground or underground and the percentage of time in the job spent underground (i.e., Job Radon Level [WLM] = Job Duration [yrs] × Radon Level [WL] × Radon Category × % Underground × 2000 hours per year/170 hours per month). Radon exposures when working on the surface were assumed to be zero.

The underground radon levels in the mines were low by both occupational and environmental standards. Specifically, WLs were well below the NIOSH Recommended Exposure Level (1 WL), the MSHA standards (1 WL) and the OSHA Permissible Exposure Limit (100 pCi/L or about 0.5 WLs). MSHA further limits annual exposure to 4 WLMs;

Table 4.4. Summary of Radon Levels in the Mines

Facility	Mine Type	% Values < LOD[a]	Mean Area Concentration (pCi/L)[b]	Mean Area Ever-UG workers[c] WL[d]
A	Limestone	16	1.8	0.009
B	Potash	56	3.4	0.017
D	Potash	61	3.2	0.016
E	Salt	31	3.2	0.016
G	Trona	76	3.4	0.017
H	Trona	85	1.6	0.008
I	Trona	80	1.6	0.008
J	Potash	62	1.8	0.009

[a] Source: Calculated using NIOSH/MIDAS data.

[b] Source: Converted from Attfield 2012 working levels by pCi/L=WL*200.

[c] Attfield et al. 2012, Table 2. Provided in Appendix Table C.2.

[d] WL = a measure of exposure to alpha particle energy per liter of air to both radon and its daughters.

UG = underground; WL = working level.

the mean cumulative exposure to ever-underground workers was just over 2 WLM (2.3 for cases and 2.0 for controls, Appendix Table E.1) with an estimated upper 95% confidence limit of about 5 WLMs. The mean radon levels of 0.01 to 0.02 WLs correspond to 1.8 to 3.4 pCi/L and are also below the residential indoor action level of 4 pCi/L set by the U.S. EPA for undertaking remedial measures to reduce radon in U.S. homes.

The Panel conducted a series of analyses to evaluate directly the potential for confounding by radon in the study (an overview of the models used can be found in Appendix Table E.3). The Panel first sought to control for confounding by radon by including terms for radon exposure in the original models used by the investigators. These analyses included radon either as a categorical or as a continuous variable in the main study models that included REC as either average or cumulative exposure (lagged 15 years) as well as the other standard covariates (i.e., smoking, history of respiratory illness, and history of working in high risk jobs). The results of these analyses are presented in Appendix Table E.4).

A comparison of the results from the original investigators' main models without radon (Model 1) to those including radon as a continuous variable (Model $1R_{cont}$) suggest that radon has a modest confounding effect on the association between REC and lung cancer risk. The lung cancer ORs for average REC with adjustment for radon decline relative to those of the main model at both the third quartile (by 16.7%) and the fourth quartile (by 19.6%) of REC exposure and the trend in lung cancer risk weakens somewhat. The lung cancer ORs for cumulative REC, lagged 15 years, declined by 20% relative to the main model at the highest quartile of cumulative exposure, although there was still a positive trend in lung cancer with increasing REC exposure. The corresponding ORs for radon and lung cancer suggest an effect of radon analyzed as a continuous variable on lung cancer; in the models with cumulative REC, lagged 15 years, the OR was 1.11 (95%CI: 0.94–1.30) (Appendix Table E.5).

However, models that include both cumulative radon (measured in WLM) and REC, especially cumulative REC, are problematic in this study. The metric WLM is based on both radon concentration (measured in WL) and the duration of time spent working in the mines (months). Given that there is very little variation in radon WLs, most of the variability in individual radon exposures comes from duration of exposure, which is effectively determined by duration of underground employment (thus the correlation between cumulative radon and duration of work in the mines is high: Spearman correlation = 0.92; Pearson correlation 0.89).

Since duration of time spent in the mines also involves exposure to diesel exhaust, the concern is that WLMs are effectively a surrogate for duration of exposure to diesel exhaust, and therefore cumulative REC. Indeed, the Spearman (Pearson) correlation between cumulative REC and radon (WLM) was 0.86 (0.75) for unlagged exposures and 0.68 (0.66) for exposures lagged 15 years based on all subjects; they were slightly lower in subjects exposed to radon (Appendix Table E.2). Cumulative radon is also correlated with average REC (unlagged and lagged 15 years) among all subjects, although the Spearman and Pearson correlation coefficients did not agree as well (e.g., 0.62 and 0.35, respectively for REC lagged 15 years). The corresponding correlations between cumulative radon and average REC exposures were substantially lower (e.g., 0.13 and 0.27, respectively for REC lagged 15 years) when the analysis was restricted to those subjects exposed to radon. The Panel found that the correlations between REC and radon diminished across most quartiles of radon but particularly with average REC (Appendix Table E.2) which likely contributes to variability in ORs observed across quartiles. As there is more variation in REC than in radon exposures, cumulative REC is less correlated with duration alone (Spearman correlation 0.32; Pearson 0.35) than it is with cumulative radon.

Given these concerns, the Panel explored a number of models incorporating both duration of REC exposure and radon (see Tables E.4 and E.8). In models where duration was added to the original investigators' main models and adjusted for radon, there was a limited effect on the relationship between REC exposure and lung cancer, particularly for cumulative REC, lagged 15 years (Appendix Table E.4).

The clearest evaluation of the strong relationship between duration of exposure to REC and cumulative radon can be found in the analyses in Appendix Table E.8. The analysis repeats one by Silverman and colleagues (2012, Table 3) in which the effect of duration of REC exposure on lung cancer is assessed directly using a referent group that is unexposed to REC (i.e., "all subjects who worked surface jobs with either negligible or bystander exposure to REC, regardless of duration"). The analysis shows a positive trend in lung cancer risk with increasing duration of exposure to REC, although this was largely driven by OR in the highest quartile (2.09, 95% CI: 0.89 to 4.90). Adjusting for cumulative radon (as a continuous variable) had the effect of reducing the ORs in that quartile (1.32, 95% CI: 0.5 to 3.51) (see Appendix Table E.8). This result is what would be expected given the high correlation between duration of REC exposure and cumulative radon and provides some evidence that adjusting for radon is essentially removing some of the effect of exposure to REC.

Note that the use of the unexposed referent group has the effect of focusing the analysis on subjects for whom duration was more likely to be a measure of exposure to REC

(i.e., removing from the higher quartiles exposure subjects who might have worked for long periods, but who had no exposure to REC). When subjects are divided by quartiles of duration without regard to REC exposure, subjects with longer duration of work, but low exposure to REC are removed from the reference category and dispersed to higher quartiles of duration. This has the expected effect of diluting the relationship between duration of exposure to REC and lung cancer risk. This effect can be seen clearly in the analyses by Crump and colleagues (2015, bottom of Table II) where no relationship is found; this effect may also have obscured some of the Panel's analyses of REC–lung cancer risks in models discussed above that include duration.

Crump and colleagues (2015) created two modified REC lung cancer risk models, one "without radon" and one "with radon" with which to analyze the effect of radon. Each uses a different, but overlapping, set of covariates from one another and from that of the original investigators, so the models are not strictly comparable (see Table 4.5). They reanalyzed the DEMS case–control data using the same REC exposure estimates employed by the original investigators (referred to as DEMS REC 1). Their Table II shows that in all but the analysis with cumulative REC lagged 15 years, the magnitude and significance of the trend for the association between REC and lung cancer declined when the "with radon" model was used. For example, using the original DEMS exposure assignments for cumulative REC exposure, lagged 15 years (DEMS REC 1), Crump reported in Table II that the OR in the highest quartile was estimated to be 3.24 (1.40–7.55) using his "without radon" model and 2.46 (0.94–6.47) using the "with radon" model, a decline of 24%. Given the differences between the variables included in the models, the differences in their results cannot necessarily be attributable only to radon. In particular, the "with radon" models add two other variables that are usually included in the original investigators' main models.

Crump and colleagues (2015) also introduced an additional test for trend (T2) to the analyses of the "with" and "without radon" models. In the T2 test for trend, REC exposure is assigned as a continuous variable; in the T1 test, also used by the original investigators, the median REC exposure is assigned to each individual in a category of exposure. The two trend tests lead to different conclusions, primarily when the "with radon" model is used and in analysis of the only-underground subgroup. Taking the results for cumulative REC, lagged 15 years as an example, the categorical trend test (T1, $P = 0.006$) and continuous trend test (T2, $P = 0.06$) were both significant in the "without radon" model; in the "with radon" model, the T1 ($P = 0.02$) indicated a significant trend while the T2 test for trend did not ($P = 0.72$).

Although use of continuous variables for tests of trend are generally preferred, the tests for trend can be greatly influenced by the existence of a number of influential data points; this is one possible reason for the marked differences noted between the results of the tests. Crump and colleagues

Table 4.5. Comparison of Covariates in Silverman et al. 2012 and Crump et al. 2015 Models

Covariate	Silverman et al. 2012	Crump et al. 2015	
		Without Radon[a]	With Radon[a]
Smoking status/pks/day/location of work	X[b]		
Smoking status and packs/day		X	X
High risk job for lung cancer of more than 10 years duration (Yes, No, unknown)	X		X
History of respiratory disease 5 or more years before date of death/reference date	X	X	X
Body mass index		X	X
Smokers in residence in childhood and adulthood		X	X
Family history of lung cancer			X
Cumulative radon (WLM) as a continuous variable			X

[a] Designation given by Crump et al. 2015.

[b] X – Variables included.

WLM = working level month.

(2015) cite one example in which they found that the exclusion of only 5 of 666 controls resulted in the *P* value for the continuous T2 test to change from 0.12 to 0.02 in one analysis, with minimal change in the categorical T1 test. Crump and colleagues (2015) indicated a preference for reliance on the continuous T2 test, provided that some procedure for dealing with highly influential data points were used. However, in the Panel's view, such analyses of potential influential data points have not yet been done and it would be difficult to draw any conclusions from the T2 test results until such analyses are completed and published.

As a final step to put radon lung cancer risks in perspective, the Panel used the BEIR VI (National Academy of Sciences 1999) constant relative risk models to estimate the lifetime relative risk of lung cancer at the highest exposures reported for subjects in the mines. The BEIR VI committee expressed a preference for a simple linear model for exposures less than 50 WLM: Relative Risk = 1 + β*WLM) where β is excess relative risk per exposure and is estimated as 0.0117/WLM (95% CI: 0.002 to 0.225). The highest average exposure for the highest quartile of exposure from the case–control study was 5.08 WLM for cases and 4.81 WLM for controls. Using the NRC's constant relative risk model, a lifetime exposure of 5 WLM would correspond to a relative risk of 1.06 (95% CI: 1.01 to 2.03). This level of risk is notably lower than the lung cancer OR of 2.83 (95% CI: 1.18 to 6.26) reported by Silverman et al. (2012) at the highest quartile of cumulative REC exposure, lagged 15 years without adjustment for radon and 2.26 (95% CI: 0.94 to 5.46) when adjusted for radon as a continuous variable.

The Panel's review of the available radon data led it to conclude that the design of the study — the selection of mines that had generally low exposure to radon — was borne out by the exposure monitoring. The low levels of radon in the mines, limited detection of and variability in the radon levels, and the inability to disentangle the cumulative REC and cumulative radon in the analyses, led the Panel to conclude that simple adjustment for cumulative exposures to radon in the DEMS data set yields results of questionable validity. Sparseness of the radon data alone can lead to artifacts in the adjusted risk estimates that may go unrecognized (Sullivan and Greenland 2013), so they should be evaluated as part of decisions on whether or how to control for radon in quantitative risk assessments. While it is not possible to exclude some contribution of radon to the lung cancer risk observed in underground miners, it is implausible that the radon levels in these mines would substantially explain the associations with REC observed in this study. Given all these factors, the Panel concluded that adjustment for cumulative radon exposure was not critically important and could itself lead to unintended biases in the REC–lung cancer associations.

4.4 SENSITIVITY OF ALTERNATIVE STATISTICAL MODELS AND EXPOSURE ESTIMATES ON RISK OF LUNG CANCER IN THE CASE–CONTROL STUDY

The sensitivity analyses conducted by both the original investigators and by Crump and colleagues (2015 and in press) with alternative retrospective exposure estimates offered the Panel the opportunity to examine the extent to which those estimates influenced the strength of the associations, including trends in the ORs across quartiles of exposure and slopes of exposure–response relationships based on continuous data. Crump and colleagues implemented their alternative REC estimates in the case–control study.

Table 4.6 presents a comparison of the results of the original investigators' sensitivity analyses to those of their "primary" analyses. Attfield and colleagues' (2012) supplemental analyses found that REC estimates developed using: 1) the 5-year CO averages after 1976 and the ratio of AdjHP/CMF before 1976; 2) the alternative exponent (β) of 0.58 in *RELtrend*; or 3) median, rather than mean REC measurements to derive the 1998–2001 REC reference values (R_{kR}) each produced lower mean estimates of risk (expressed as HR per 1000 μg/m^3-yr cumulative REC exposure, 15-year lag) than did the primary analyses. The sensitivity analyses conducted by Silverman and colleagues (2014) in the case–control study showed a small reduction in the ORs for cumulative REC at the highest quartile when using the REC estimates based on either 5-year average CO values or on the use of β = 0.58, but the overall trends remained largely the same and were highly significant.

As discussed earlier, Crump and colleague's analyses of exposure covered a broader range of alternate exposure assumptions. They imputed missing or nondetected CO values using different methods than employed in the original analyses and estimated different CO:REC relationships (Crump and Van Landingham 2012). They also evaluated the impact of their alternative model that did not rely on the use of CO but instead on the underlying AdjHP/CFM relationship (REC 6). They implemented these alternative exposure estimates in the "with radon" and "without radon" models discussed earlier (see Table 4.5). They conducted these analyses in the case–control study with all subjects, with ever-underground workers and with those subjects who had worked only underground in the mines (see Crump et al. 2015). They also used conditional logistic regression analysis to derive slope factors (risk per μg/m^3-yr) for each of these analyses to illustrate how the DEMS data could be used for quantitative risk analysis. They evaluated the significance of the trends for lung cancer risk in each of their analyses using two tests for trend, T1 (categorical REC) and T2 (continuous REC), introduced in the previous section on radon.

Table 4.6. Sensitivity Analyses and Impact of Alternate Models on Risk Estimates

		Sensitivity Analyses		
	Primary Model	5-yr CO Average	REC = $CO^{0.58}$ Model	Median
Cohort Study: Attfield et al. 2012				
HR (95% CI) per 1000 μg/m³-yr cumulative REC exposure (< 1280 μg/m³-yr)	No tenure exclusion 2.79 (1.59–4.89) Excluding < 5 yrs tenure 4.06 (2.11–7.83)	No tenure exclusion 1.83 (1.00–3.35) Excluding < 5 yrs tenure 2.39 (1.20–4.76)	No tenure exclusion 1.87 (1.03–3.43) Excluding < 5 yrs tenure 2.64 (1.29–5.41)	No tenure exclusion 2.35 (1.31–4.22) Excluding < 5 yrs tenure 3.33 (1.71–6.47)
Nested Case–Control: Silverman et al. 2014				
OR (95% CI) Quartiles[a], μg/m³-yr cumulative REC exposure, 15-year lag	Q1: 1 (reference) Q2: 0.74 (0.40–1.38) Q3: 1.54 (0.74–3.20) Q4: 2.83 (1.28–6.26) P_{trend} = 0.001	Q1: 1 (reference) Q2: 0.71 (0.38–1.31) Q3: 1.73 (0.84–3.59) Q4: 2.28 (1.02–5.08) P_{trend} = 0.015	Q1: 1 (reference) Q2: 0.73 (0.39–1.37) Q3: 1.48 (0.72–3.05) Q4: 2.34 (1.08–5.10) P_{trend} = 0.004	Q1: 1 (reference) Q2: 0.71 (0.38–1.32) Q3: 1.67 (0.81–3.45) Q4: 3.03 (1.35–6.79) P_{trend} = 0.001

[a] Quartile exposure ranges are as follows; Q1: 0 to < 3, Q2: 3 to < 72, Q3: 72 to < 536, Q4: ≥ 536.

Sources: Attfield et al. 2012 (Supplementary Table 13); Silverman et al. 2014.

In its review, the Panel focused on the results of Crump and colleagues' (2015) alternative REC analyses using their models "without radon" adjustment in the full set of case–controls and ever-underground workers. For reasons discussed at length in the radon analyses, the Panel thought that analyses adjusting for radon, in particular cumulative radon, were problematic even in the more complete data sets. Given potentially important differences in the covariates included in the "with radon" and "without radon" models, the Panel also found it difficult to attribute differences in results from those models to radon alone. As discussed in the section on subgroup analyses, the Panel thought that the further stratification of the ever-underground workers to only-underground was not well justified by differences in exposure and only had the predictable effect of diminishing the statistical power and the precision of the estimates as a consequence of reduced sample size.

Based on data from Crump and colleagues' paper (2015), Figures 4.5 (for all-subjects) and 4.6 (for ever-underground subjects) plot the mean ORs (•) and 95% CIs (whiskers) and to the right, the analyses for slope and trend. In both figures, the first three sets of boxplots are based on the original DEMS REC estimate (DEMS REC 1) and two of the original investigators' alternative DEMS metrics shown in Table 4.6 (DEMS REC 2 based on 5-year average CO after 1976 and the ratio of AdjHP/CFM before 1976, DEMS REC 3 based on using β = 0.58 in *RELtrend*).

These results demonstrate a high degree of robustness in the association of REC with lung cancer to their alternative estimates of exposure to REC. Although there are notable differences in the magnitude of the quartiles between models, the figures show risks of a similar magnitude and similar trends in ORs by quartile across all alternative REC models, with broader CIs in the ever-underground group. The analyses with the REC 6 model are particularly noteworthy because they show a clear association between lung cancer and the simple indicator of REC based on AdjHP/CFM in the mines, one that does not rely on CO.

In the analyses with all subjects, the T1 slopes are similar to or lower than that estimated using the DEMS REC 1 estimate but vary by a factor of at most 2 and all (except REC 3) are statistically significant; the T2 slopes are all consistently lower than those predicted using the T1 test, but again reflect mostly significant trends. In the ever-underground

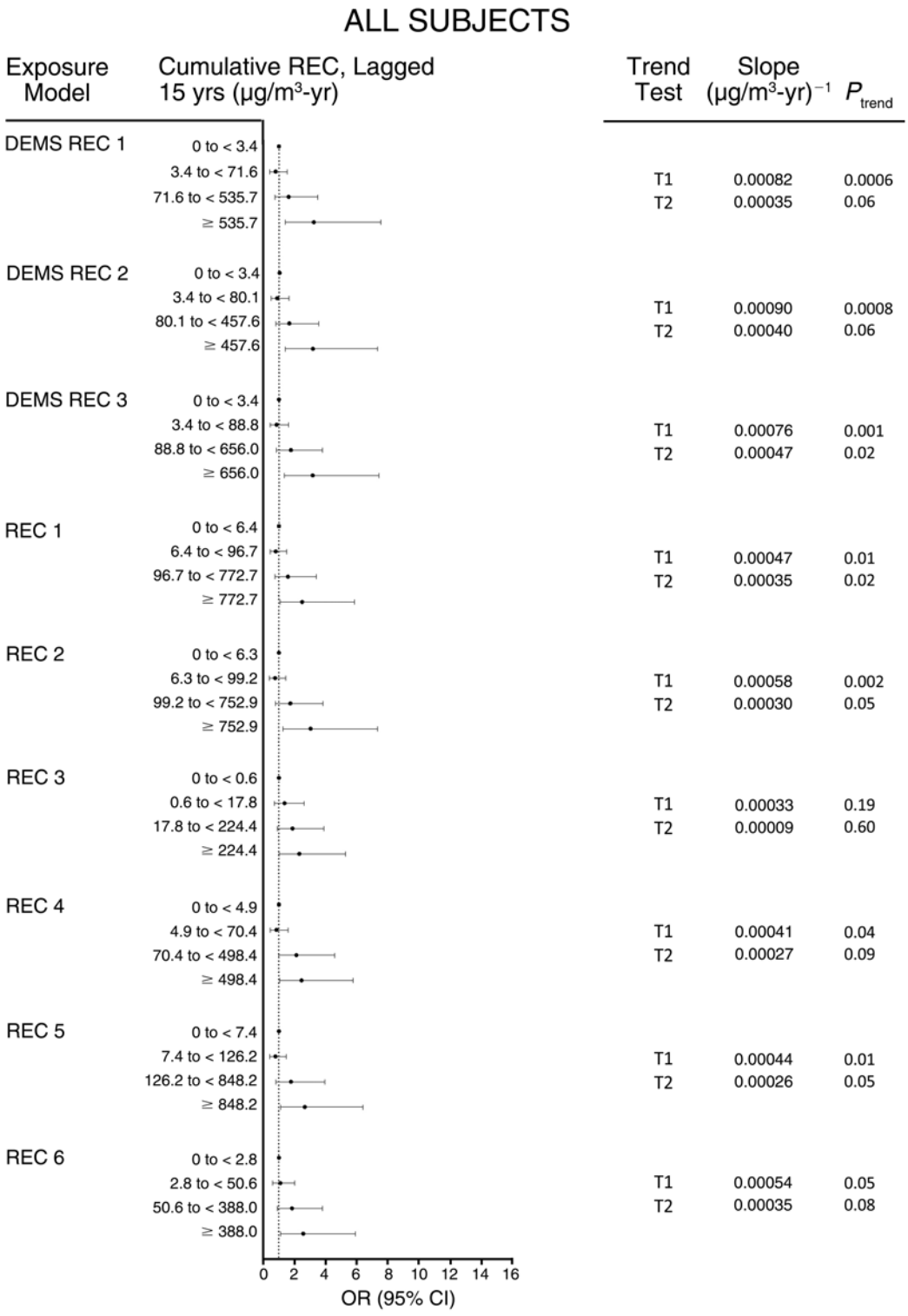

Figure 4.5. Comparison of ORs for lung cancer and cumulative REC, lagged 15 years using alternative REC exposure models, without adjustment for radon, in all subjects. In the table on the right, the T1 trend test was similar to that of the original investigators but assigned each subject the average REC exposure in each quartile, as opposed to medians used by Silverman et al. (2012). The alternative T2 trend test used each individual subject's estimated REC exposure. (Source: Crump et al. 2015, Table III.)

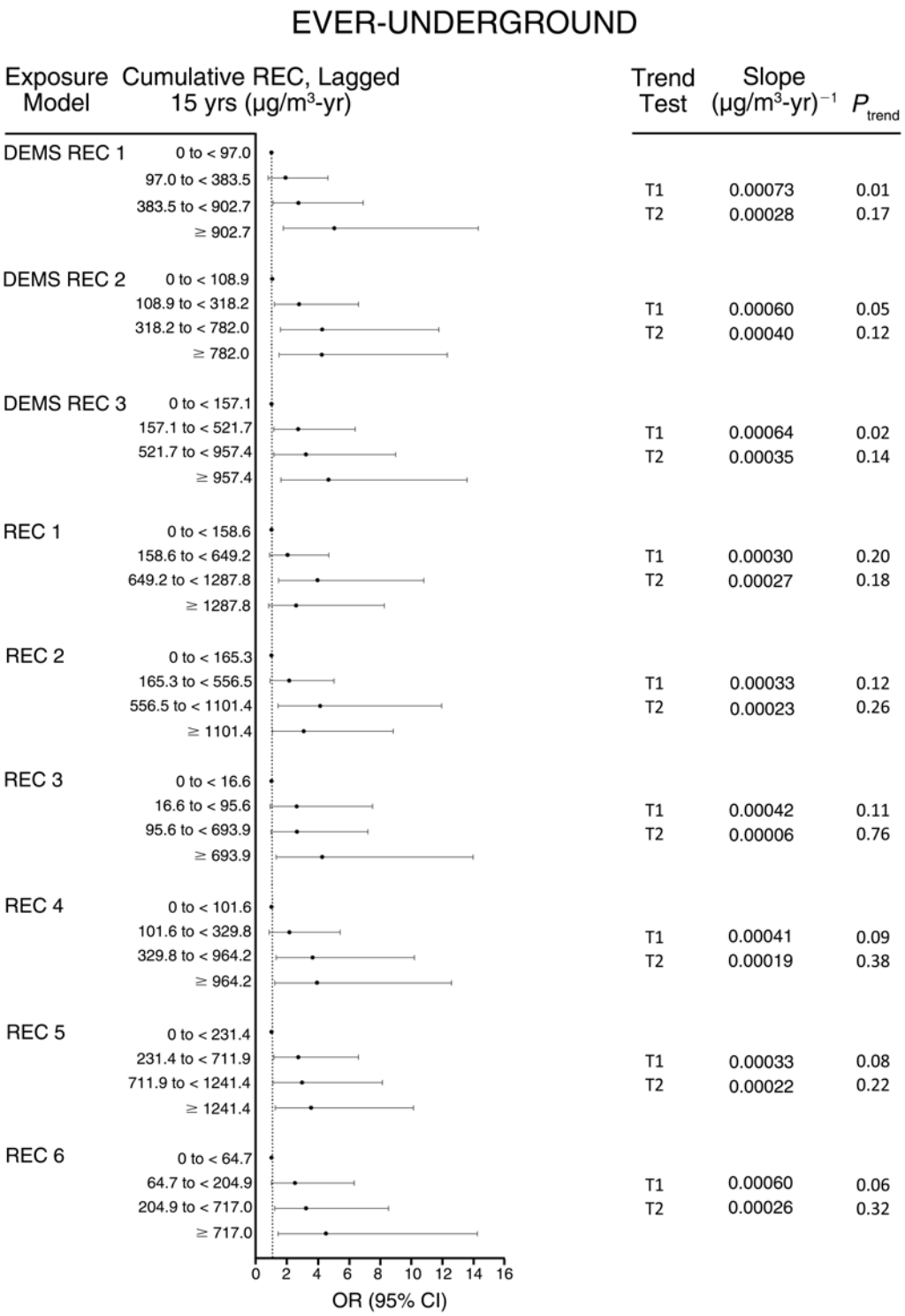

Figure 4.6. Comparison of ORs for lung cancer and cumulative REC, lagged 15 years using alternative REC models, without adjustment for radon, in ever-underground workers. In the table on the right, the T1 trend test was similar to that of the original investigators but used the average REC exposure in each quartile, as opposed to medians used by Silverman et al. (2012). The alternative T2 trend test used each individual subject's estimated REC exposure. (Source: Crump et al. 2015, Table IV.)

analyses, the T1 results (slopes and trend tests) are consistent with those with all subjects, although the levels of significance are somewhat lower.

The T2 tests yielded similar or lower slopes compared to T1, but the *P* values for trend are all larger and nonsignificant. The reasons for the differences in the trend tests applied were not given; as discussed earlier, the presence of influential data points in the continuous analyses is one possible explanation (Crump et al. 2015). Another reason may be that the T2 test used individual exposures and fits a linear model to the log OR, just as in a regular logistic regression with continuous exposures. However, Silverman and colleagues (2012) noted in their evaluation of continuous exposure data that the log-linear models did not fit the data well. Whether the T2 trends are less pronounced or nonsignificant because there is in fact no trend or because the trend is not log-linear should be clarified in future analyses. At this juncture, these analyses do not alter the Panel's conclusions about the basic robustness of the REC–lung cancer relationship against the rich suite of alternative REC exposure estimates developed by the original investigators and by Crump and colleagues (2015 and in press).

As discussed in the earlier section on exposure assessment, the Panel could not resolve the extent to which these alternative approaches to estimating REC had addressed when or if documented reductions in diesel engine emissions per unit of brake HP-hour might have been reflected in the mines over time (e.g., for onroad engines in U.S. EPA 2002). The Panel thought it unlikely that a different trend in historical emissions over time would change the relative ranking of exposures assigned across subjects, and thus would not necessarily undermine the basic association between REC and lung cancer. In fact, this view is supported by the robustness of the association to alternative exposure assumptions by both Silverman and colleagues (2014) and Crump and colleagues (2015 and in press). However, these analyses also suggest that biases in historical exposures could affect the magnitude of the estimated slope of the exposure–response relationship.

4.5 CONCLUSIONS

In its evaluation of the DEMS and its potential use for quantitative risk analysis, the HEI Panel considered a broad set of attributes for the design, conduct and oversight of a study that affects the basic integrity or potential for bias in the data collected and of the analyses that depend on them. These attributes were discussed in detail in Chapter 2 and broadly involve the strength of the study design, the integrity and quality of the methods used to collect data on exposure and health outcomes — including confounding factors — the appropriateness of the statistical analyses, the extent to which alternative assumptions and other uncertainties have been explored, and the processes followed for oversight and peer review of the study. The Panel also examined in greater detail three issues of particular importance to understanding the main results of the study: the potential for confounding by smoking; the potential for confounding by radon exposure in the mines; and the effect of error in the measurement and modeling of exposure on estimated health effects.

The Panel's overall assessment was that the DEMS was designed, conducted, overseen and evaluated according to high standards of scientific research and that its data can be used to support quantitative risk analyses, including sensitivity and uncertainty analyses. The study was carefully designed to test the hypothesis of an association between long-term exposure to diesel exhaust in the mines and lung cancer, while providing data with which to evaluate and control for potentially important occupational confounders. A nested case–control design was also used to evaluate and adjust for smoking and for other risk factors for lung cancer (including history of chronic respiratory disease or history of time spent in other occupations associated with a higher risk of lung cancer). The results of the cohort and case–control analyses, despite the absence of control for smoking in the cohort study, were broadly consistent with an increasing risk of lung cancer in relation to exposure to REC. Both the Panel and other investigators (Crump et al. 2015; Crump et al. in press; Moolgavkar et al. 2015) have successfully replicated the main results reported by the original investigators. The association between REC exposure and lung cancer risk has been shown to be robust in numerous sensitivity analyses using alternative assumptions, statistical models, and alternative REC exposure estimates by the original investigators and by others (Crump et al. 2015; Crump et al. in press; Moolgavkar et al. 2015). Consequently, much important groundwork has been laid for the use of these studies to develop quantitative risk estimates and to characterize the level of confidence or uncertainty in the results.

As with any retrospective occupational health study, the need to develop historical estimates of exposure contributes potentially important uncertainties that are inherent to retrospective epidemiological investigations. While the existence of the association between diesel exposure and lung cancer in DEMS is robust to alternative exposure metrics, the magnitude of the resulting exposure–response function (the key value for any risk assessment application) does vary with alternative assumptions. An additional limitation relates to the differential risk of lung cancer from cigarette smoking between surface and underground workers, which

is challenging to interpret and may indicate additional uncertainties in the pooled exposure–response relationship.

There are still many decisions to be made in how the data and specific results of this study might be further analyzed or adapted for use in various quantitative risk assessments for different populations. The results are clearly more directly generalizable to work environments where the characteristics of the populations, diesel engine technology, fuels, and other exposures are similar to those in the mines studied. However, the lower exposures observed in the DEMS are approaching levels in urban air. Ultimately, it is important to recognize that the DEMS would not likely be the sole basis for any quantitative risk assessment, but would need to be interpreted more fully and applied within the context of the broader scientific literature on diesel exhaust and health, exposure–response modeling, and the emerging literature on emissions from newer technology engines.

Chapter 5

Summary and Conclusions

[J]udgments as to the level of uncertainty to be tolerated are not scientific but rather reflective of the policy-making process. Here, there should be substantial, continuing dialogue between scientists and policy makers. (The 2002 HEI Diesel Epidemiology Working Group)

5.1 SUMMARY

5.1.1 INTRODUCTION

This report has provided the HEI Diesel Epidemiology Panel's review of two studies of exposure to diesel exhaust and the risk of lung cancer: the Diesel Exhaust in Miners Study (DEMS*) conducted by investigators led by Drs. Debra Silverman and Dr. Michael Attfield and their colleagues at the National Cancer Institute and the National Institute for Occupational Safety and Health, respectively, and a study of workers employed in the trucking industry (the Truckers study) conducted by Dr. Eric Garshick and his colleagues at the Veterans Administration Hospital and Harvard University. These two studies, in combination with the full body of evidence on diesel exhaust, contributed to the International Agency for Research on Cancer's (IARC) decision to designate diesel exhaust as an IARC Group 1, or known human carcinogen. This decision, building as it does on an assessment of the broader scientific evidence, establishes the first step in a risk assessment process discussed in Chapter 2, the identification of a potential hazard.

In response to requests from HEI sponsors, the Panel was charged with evaluating the two studies, their strengths and limitations, and the extent to which their data and results could now support the next step, a quantitative characterization of the lung cancer risks associated with diesel exhaust, or a quantitative risk assessment. The charge, detailed in Chapter 1, did not include comprehensive reanalyses of the studies, development of exposure–response relationships for regulatory use, or estimation of potential risks to other occupational or general populations.

Within the broader structure of quantitative risk assessment, the Panel focused on the potential value of the studies for development of the quantitative exposure–response relationship between diesel exhaust and lung cancer. The Panel evaluated the studies with respect to earlier HEI Panels' research recommendations to address limitations of previous epidemiological research (Table 5.1), and with respect to the attributes of well-designed epidemiological studies that make them useful for quantitative risk assessment. These attributes pertain to the overall process by which the study was conducted, the strengths of the study design to estimate the exposure–response function, including the control for potential confounders, the appropriateness of the overall analytical approach, the quality of the outcome and exposure assessments, the strength of the statistical analyses, the robustness of the analytic methods and results to alternative assumptions, and the characterization of uncertainties at various steps in the analyses.

As part of its evaluation, the Panel spent more time with the DEMS data because of the opportunity to examine the robustness of the lung cancer risk estimates to the two major factors of concern — the impact of potential confounding exposures and the potential uncertainties in exposure assignment. The Panel understood that the ultimate decisions about which data or results to use for quantitative risk assessments, or how particular policies should take into account remaining uncertainties, were beyond its scope.

5.1.2 THE TRUCKERS STUDY: LUNG CANCER AND ELEMENTAL CARBON EXPOSURE IN THE TRUCKING INDUSTRY

The study by Garshick and colleagues (2012a), the focus of the HEI Panel's evaluation in Chapter 3, is the culmination of decades of work investigating a number of health outcomes in association with employment in the trucking industry. This study specifically examined the risk of lung cancer in relation to quantitative estimates of personal exposure to submicron elemental carbon (SEC). Several publications led up to this study, laying the groundwork for the development of individual-level exposure estimates and the subsequent epidemiological analyses (Davis et al. 2006, 2007, 2009, 2011; Garshick et al. 2008; Jain et al. 2006; Laden et al. 2007, Sheesley et al. 2008, 2009; Smith et al. 2006). The investigators found weak associations and evidence of trends in hazard ratios for cumulative SEC and lung cancer; those associations strengthened when adjusted for duration of work, a proxy for a healthy worker survivor bias. The findings were strongest when subjects in the mechanics job category were excluded from the analysis, a category whose exposures the investigators judged were

* A list of abbreviations and other terms appears at the end of the Special Report.

Table 5.1. Research Needs for Quantitative Risk Assessment: Overview of Progress in DEMS and Truckers Studies[a]

Research Needs for QRA / Specifically	DEMS	Truckers
Better measures of exposure		
• Measures of diesel constituents.	√	√
• Of particular importance are the selection and validation of a chemical marker of exposure to the complex mix of diesel exhaust emissions.	√	√
• Specific biomarkers of diesel exposures, health outcomes, and susceptibility are needed.	X	X
Better models of exposure		
• Exposure models may include data from personal monitors, area monitors placed where diesel exposure is likely to occur, and current and historical data regarding emission sources.	√	√
• In any such modeling effort, the effects of environmental tobacco smoke should be removed as completely as possible.	√	X
• Reliable estimates of past emissions and of factors affecting historical exposures in a range of settings are needed to improve the characterization of uncertainties, both quantitative and qualitative, in historical models of exposures.	√	√
Design needs for new studies of exposure–response		
• Exposures should be adequately and accurately characterized with respect to magnitude, frequency, and duration, rather than solely by duration of employment.	√	√
• The exposures considered should be close to levels of regulatory concern, including a range of exposures to provide a base for understanding the relation between exposure and health effects.	√	√
• Errors and uncertainties in exposure measurements should be quantified where possible;	√	√-
• These should be fully reported to users, and taken into account in both power calculations and exposure response analyses.	√	√-
• Cigarette smoking is a potent risk factor for lung cancer, and it must be controlled for in any study of risk factors for this disease.	√	X
• Smoking histories obtained for a cohort study subset that uses a case–control or case–cohort design will strengthen the interpretation of results.	√	X

[a] Sources: HEI 1999, 2003; HEI Diesel Epidemiology Expert Panel 1999; HEI Diesel Epidemiology Working Group 2002.

√ = the study addressed the issue: X = the study did not or could not address the issue.

subject to more exposure measurement error than other job categories.

The Truckers study was designed, conducted, and independently reviewed according to high standards of scientific research. In light of the research needs identified from evaluations of earlier epidemiological studies (Table 5.1) and the characteristics desirable in studies that are intended to provide the basis for quantitative risk analysis, the Panel thought the study had a number of specific strengths:

- The study was designed for and conducted in a large cohort of 31,135 workers employed in trucking facilities geographically distributed across the United States. The follow-up period and process for identifying cases of lung cancer were adequate.
- The investigators provided a well-reasoned justification for their selection of SEC as a measure of exposure. They found that it better predicted work exposure levels and their source apportionment analyses identified diesel engines as a primary, although not exclusive, source of SEC in a subset of their trucking

terminals. SEC had fewer interferences from other sources of combustion products than other particulate matter (PM) components they considered (organic carbon or total carbon), including cigarette smoke.

- The retrospective exposure assessment was conceptually and statistically sound. It relied on a statistically-designed exposure monitoring survey in U.S. trucking terminals, detailed job history and work practice records, and a creative, state-of-the-art structural equation modeling approach to estimating job-specific SEC exposures. It also estimated historical trends in those exposures using regional coefficient of haze measurements, a reasonable surrogate for particulate elemental carbon (EC). The investigators were able to validate some components of their exposure model, giving some insights to the sensitivity of their model estimates to their key assumptions.
- The SEC predictions by job category span a range that both overlaps with that of DEMS and includes concentrations relevant to ambient levels. Mean SEC ranged from 1.8 $\mu g/m^3$ for clerks to 40.8 $\mu g/m^3$ for dock workers in 1971–1980; it ranged from 0.8 to 24.7 $\mu g/m^3$ for the same groups two decades later.
- The statistical analyses followed a logical and well-established sequence beginning with standardized mortality ratios that identified a modest elevation in lung cancer risk and leading up to the proportional hazards modeling analyses. The proportional hazards modeling appropriately stratified by decade of age at entry, calendar year, and decade of hire and also adjusted for race, census region, and duration of work. The investigators adjusted for duration of work to account for the healthy worker survivor bias observed in their data.
- The investigators conducted sensitivity analyses evaluating different exposure metrics (average and cumulative SEC, based on both categorical and continuous measurements) for the full cohort and for the cohort excluding mechanics.

As in any epidemiological study, the Truckers study has limitations, with resultant uncertainties, that warrant consideration in the interpretation and application of study results to quantitative risk assessment for diesel exhaust. Some notable uncertainties that emerged from the Panel's evaluation were that:

- As is often the case for retrospective exposure reconstruction, the investigators had little independent data with which to validate their predictions of SEC exposure at various steps in their analysis. For example, their use of coefficient of haze to capture and reflect temporal trends in background EC levels was an appropriate step, but was based on one region of the country (New Jersey). The representativeness of New Jersey data for other parts of the country where trucking terminals were located has not been explored.
- SEC may not be entirely attributable to diesel exhaust in this study, given the presence of other combustion or fuel sources. Although the supporting analyses conducted for this study, as well as other scientific literature, point to diesel exhaust as a major contributor to EC concentrations, and to SEC in particular, other fuel sources contribute to varying degrees by location, and possibly over time.
- Analyses conducted in an earlier study by Garshick and colleagues (2008) in which they used job-level smoking rates to adjust job-related lung cancer rates provide useful insights, but the investigators of the 2012 study could not obtain individual-level smoking data so were unable to control directly for smoking in their analyses.
- The investigators found weak associations between cumulative SEC exposure with lung cancer in both the full cohort and in the cohort excluding mechanics. These associations were stronger and more consistent with a trend when the models were adjusted for duration of employment. While the target of this adjustment, the healthy worker survivor bias, is a concern in occupational epidemiological studies, the science on the role of duration of employment in such analyses remains in an unsettled state. The adjustment for duration in this study creates some challenges for interpretation of the results and for their comparison with those of other studies lacking such adjustment.

5.1.3 THE DIESEL EXHAUST IN MINERS STUDY (DEMS)

The DEMS was designed to study associations between diesel exhaust, measured as respirable elemental carbon (REC), and lung cancer in a cohort of 12,315 workers from eight nonmetal mines in the United States. The Panel's review of the DEMS focused on the analyses of the cohort conducted by Attfield and colleagues (2012), on the nested case–control study by Silverman and colleagues (2012), and on the related series of five publications that laid out the details and results of the retrospective exposure analysis (Coble et al. 2010; Stewart et al. 2010; 2012; Vermeulen et al. 2010a,b). The results of the cohort and the case–control studies were broadly consistent and found an increasing risk of lung cancer in relation to increasing cumulative exposure to REC.

In its evaluation of DEMS, the Panel had the opportunity to conduct several analyses with the data sets from the study as well as to examine a number of recent critiques and

analyses conducted by other scientists (in particular, Crump et al. 2015; Crump et al. in press; Moolgavkar et al. 2015). Overall, the Panel thought that the process by which DEMS was designed, conducted, independently overseen, and peer reviewed met high standards of scientific research. Considering the research recommendations for epidemiological studies of diesel exhaust (Table 5.1) and the attributes of epidemiological studies that support quantitative risk assessment, the Panel concluded that the DEMS demonstrated a number of inherent strengths:

- The study was carefully designed with sufficient statistical power and relevant data on covariates to test the hypothesis of an association between long-term exposure to diesel exhaust in the mines and lung cancer in the cohort of mine workers. The eight mines were specifically chosen based on data demonstrating the use of diesel engines during the study period, and the decision to study them together rather than individually was well-justified.
- The approach to health outcome assessment was of high quality; the lung cancer diagnoses were ascertained by laboratory pathology reports where available, and the assignment of health outcomes was blind to assignment of exposure.
- The study design, data collection instruments, and analytical approach all included strategies for controlling for potential occupational and other confounders for lung cancer. These included: selection of mines expected to have low levels of occupational carcinogens (radon, polycyclic aromatic hydrocarbons (PAH), silica, asbestos, respirable dust); measurement data to confirm levels of these carcinogens; and a nested case–control design that included questionnaire data on individual-level smoking histories, occupational histories, and several other risk factors for lung cancer; and statistical analyses that explored the impact of all of these potential confounders on lung cancer risk.
- The choice of REC as a marker for exposure to diesel exhaust in the mines was well-justified.
- The DEMS retrospective exposure assessment was logically constructed, thorough in its collection and assessment of available sources of data, and incorporated state-of-the-art methods to develop quantitative estimates of personal exposure to REC for the full period of the study. To the extent possible, the investigators confirmed or justified the decisions they made at several stages in the development of their models, using independent approaches or data where available.
- The analytical approach to the analysis of the exposure and lung cancer data followed a logical and standard progression beginning with standardized mortality ratio analyses, and proceeding with extensive analyses using both categorical and continuous exposures to REC in the cohort and in the nested case–control study. The Cox proportional hazards models were an appropriate empirical modeling choice for the type of data and hypotheses tested in this study. The investigators also fit several continuous models to their data, which provided additional ways to characterize possible exposure–response relationships that may be useful for quantitative risk assessment.
- The Panel thought the decision to analyze the full cohort with adjustment for work location (i.e., surface-only or ever-underground) or in subgroups by location was theoretically and statistically sound.
- The investigators also conducted numerous informative analyses of the sensitivity of their findings to alternative assumptions about exposure metrics, to alternative approaches to modeling relationships between diesel exhaust exposure and lung cancer, and to adjusting for confounding factors.

In its own analyses of the DEMS data, the HEI Panel focused on understanding and evaluating the sensitivity of the main findings of the case–control study to alternative approaches to adjusting for the two most important potential confounders, smoking and radon. Coupled with its evaluation of the original and additional analyses on smoking and other factors provided by the DEMS investigators (Silverman et al. 2014), the Panel concluded:

- The DEMS nested case–control study findings of an increased risk of lung cancer with increasing cumulative exposure to REC in the full cohort were robust to alternative approaches to adjusting for smoking. The differential increase in risk of lung cancer from cigarette smoking between surface and ever-underground workers, as well as some of the other differences in associations between these two groups, are challenging to interpret and may merit further exploration in applications to quantitative risk assessment.
- The Panel's assessment of radon in this study left it with a high level of confidence that radon does not substantially confound the study's results. To understand whether radon could have substantially influenced the REC–lung cancer relationship, the Panel examined the radon measurement data from the mines, conducted several analyses of its own, and thoroughly considered several additional sensitivity analyses of radon by Crump and colleagues (2015). The levels of radon were low by both occupational and residential radon criteria, the percentage of non-detectable radon measurements was large in most mines, and high correlations between cumulative REC

and cumulative radon (determined largely by duration of work in the mines) made them difficult to disentangle in the analyses. While analyses showed some sensitivity of the REC–lung cancer associations to different modeling approaches, the impact in the most appropriate analyses suggest that radon is not a major confounder, that adjustment is not critically important and could itself lead to unintended biases in the results.

- The basic association of lung cancer with diesel exhaust exposure was essentially robust to alternative modeling approaches in both the DEMS cohort and case–control studies. In the DEMS cohort data, Moolgavkar and colleagues (2015) applied the three-stage clonal expansion (TSCE) model, a model intended to represent biological processes relevant to carcinogenesis and to take into account time-dependent patterns in exposure and risk. The model predicted elevated lung cancer risks associated with different temporal patterns of exposure to REC that were attenuated with age after exposure ended, particularly at the lowest cumulative exposures (50 $\mu g/m^3$-yr). At the highest exposures comparable to those in the cohort study, attenuation did not reach background for several decades after the end of exposure. However, the Panel suggested some caution in interpreting the dose rate effects biologically, given that the TSCE model was not entirely successful in representing the carcinogenic mechanisms associated with diesel exhaust that have been identified in the scientific literature.
- With the case–control data, Crump and colleagues (2015) fit the same statistical models as the original investigators but selected sets of confounding variables that differed both from those in the original models and from one another. That is, the "with radon" and "without radon" models differed by more than the radon variable. The Panel noted that basic results from the "without radon" models, which eliminate the issues with control for radon discussed above, were similar to those of the DEMS main models.

Despite the many strong characteristics of the DEMS, there remain areas of uncertainty, most of which involve the retrospective exposure assessment and its impact on estimates of exposure–response functions. Many of the limitations of the retrospective exposure reconstruction are what make such reconstruction necessary in the first place:

- Few direct measurements of EC or other metrics specific to diesel exhaust over the full period were covered by the study, necessitating reliance on a combination of measurements from a survey in 1998–2001 and historical data on other contaminants (in particular, carbon monoxide [CO]) to estimate historical trends and levels of REC concentrations in the mines. For workers in surface jobs, very limited historical monitoring data were available to characterize exposures to diesel exhaust; consequently job-specific exposures estimated from the 1998–2001 DEMS survey were assumed to be constant back to the start of dieselization.

As discussed in detail in Chapter 4, several important questions have been raised about the validity of the retrospective exposure assessment in this study: methods for imputing missing measurements; the choice of CO with which to model trends in airborne contaminants in the mines, the use of horsepower (HP) and ventilation as the primary predictors of CO concentrations; the correlations between CO and REC in emissions; and temporal changes in diesel engine technology, fuels and their implications for the concentrations of diesel exhaust in the mines. The Panel agreed that these are potentially important sources of uncertainty in the exposure estimates and therefore in the exposure–response relationships that might be derived from the study.

Many of these issues have now been extensively explored, both by the original investigators in their own sensitivity analyses, by Crump and van Landingham (2012) in their analyses of the exposure assessment, and by Crump and colleagues (2015 and in press) in their exploration of the implications of the differences in exposure estimates for the exposure–response relationships in the case–control study. They demonstrated sensitivity in the REC–lung cancer odds ratios and in the slopes of the exposure–response relationships to their alternative statistical and exposure models. The variability in results was considerable in some cases, but in the Panel's view of the most relevant analyses, the variability was smaller and the results still demonstrated a clear, significant association between REC and lung cancer risk. The associations remained even in the alternative models (REC 6 alone and the version including alternative trends in PM emissions rates) that did not rely on the HP–CO–REC relationships that were used in the original investigators' main models. Further refinement of the exposure model using alternative trends in emissions rates may provide further understanding of the uncertainties in risk estimates, for example, that might arise from uncertainties about the timeline for the introduction of newer engine technology into the mines.

The testing of the original results, first by the original investigators and then by both the HEI Panel and by Drs. Crump and Moolgavkar and their colleagues, provides confidence in the integrity of the DEMS reported results. They also show that the quality of the DEMS data that has been made available makes possible further exploration of the data to evaluate numerous analytic choices and to understand their implications for interpretation of the results.

5.2 DISCUSSION AND RECOMMENDATIONS

5.2.1 DISCUSSION

In the Panel's view, both the Truckers study and the DEMS made considerable progress toward addressing a number of the major limitations that had been identified in previous epidemiological studies of diesel exhaust and lung cancer (see Table 5.1). These related particularly to the need for more specific metrics, models, and ultimately quantitative estimates of exposures to diesel exhaust. They both also demonstrated many of the attributes of high quality epidemiological studies that scientists and regulators value in evidence used to support quantitative risk assessments. The Panel concluded that the Truckers study and the DEMS in particular provide valuable new information that advances our understanding of the quantitative relationship between exposure to diesel exhaust experienced by the workers in those studies and their risk of lung cancer. The Panel also concluded that the Truckers study had greater uncertainty in the exposure–response relationship given the mix of sources, the more limited evidence of a trend in exposure–response without adjustment for duration of employment, the correlation between cumulative exposure and duration of employment, and the inability to directly control for tobacco smoke exposure. As is true in most occupational epidemiological studies, the findings of these studies are readily generalizable to workers in other populations exposed to similar concentrations of diesel exhaust, emitted from comparable older engines, over comparable periods of time.

The Panel was also asked to consider whether data from these studies might also be used to quantify lung cancer risk in general populations exposed to diesel exhaust at lower concentrations with different temporal patterns. The difference in exposures, along with differences in patterns of exposure over a lifetime, can raise questions about whether similar mechanisms of toxicity can be assumed. However, the broad and overlapping ranges of exposures to SEC and REC in these studies mitigates to a considerable extent concern about their generalizability to ambient levels. Although each explores higher exposures than observed in ambient environments, exposures in both studies include low concentrations of EC. In the Truckers study, job-specific SEC levels ranged from 1.8 $\mu g/m^3$ (clerks exposed to background levels) to 40.8 $\mu g/m^3$ (dockworkers using diesel equipment) (in 1971–1980) and from 0.8 to 24.7 $\mu g/m^3$ for the same groups in later years (1991–2000) (Davis et al. 2011). In DEMS, the average REC exposures over all facilities ranged from 1.7 $\mu g/m^3$ for surface-only workers to 128.2 $\mu g/m^3$ for the ever-underground workers (Attfield et al. 2012).

The low end of the range of exposures in each of the studies is very close to the levels of EC that have been reported in ambient air in the United States. Recent studies from HEI's National Particle Component Toxicity program have found mean ambient $PM_{2.5}$ EC (EC PM $\leq$ 2.5 µm in aerodynamic diameter) concentrations ranging from 0.26 $\mu g/m^3$ (East Lansing, MI) to 1.2 $\mu g/m^3$ in New York City (Lippmann et al. 2013) and from 0.8 to 2.2 $\mu g/m^3$ at home-outdoor sites in St. Paul, Minnesota, and in New York City, respectively (Vedal et al. 2013). The fourth in the Multiple Air Toxics Exposure Study (MATES) series studying the south coast air basin of California found $PM_{2.5}$ EC levels ranging from 0.9 to 1.4 $\mu g/m^3$ across 10 monitoring sites. The measurements from the studies all used thermo-optical methods for analyzing EC but are not strictly comparable as they all are based on slightly different size fractions of PM. In addition, as discussed in the context of the Truckers study, the SEC measured may reflect contributions from other vehicle and sources of EC and so may not solely represent exposures to diesel exhaust.

The approach ultimately taken to modeling the exposure–response relationship between diesel exhaust (as REC or SEC) and lung cancer risk is a choice to be made as part of the quantitative risk assessment process. The DEMS and the Truckers study provide a number of alternative assumptions and approaches that could be considered in this regard. Various investigators have already explored a number of models, both categorical and continuous, for fitting the full range of exposures in both the DEMS cohort and the case–control studies that provide insight into some of the uncertainties arising from model selection. For example, Moolgavkar and colleagues (2015) fit an alternative model, the TSCE model, with the DEMS cohort data and explored the impact on lung cancer risk of alternative patterns of cumulating exposure over time.

When multiple studies exist that offer different estimates of exposure–response, meta-analytic techniques exist to combine the information they provide while taking into account their relative strengths. The recent meta-analysis by Vermeulen and colleagues (2014b) demonstrated one such approach. This work sought to characterize the exposure–response relationship for diesel exhaust and lung cancer by fitting log linear models to the varied risk estimates from DEMS, the Truckers study, and an earlier study of truckers by Steenland and colleagues (1998). Figure 5.1, (from Vermeulen et al. 2014b), plots the relative risks at different levels of cumulative exposure from each of the studies, estimated at the exposure lags representing the best fit to the data in each of the studies. The figure suggests general compatibility of the published exposure–response results for the DEMS and the Truckers study, as well as for the earlier study of trucking workers by Steenland and colleagues.

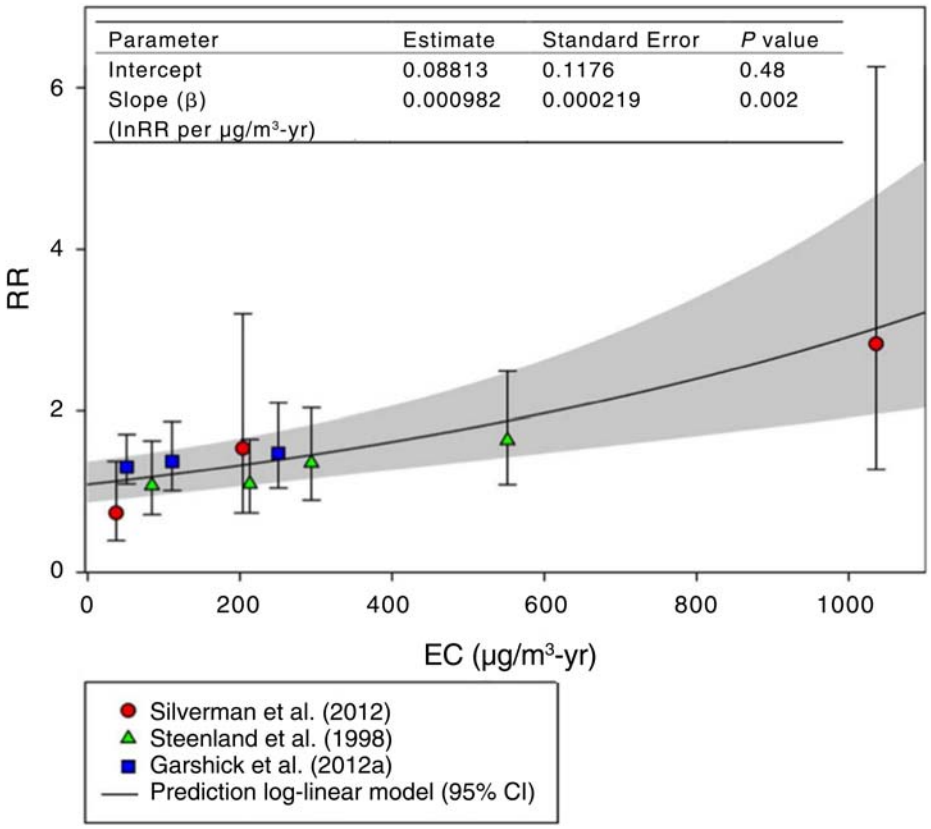

Figure 5.1. Relative risks were estimated using exposures lagged 15 years in Silverman et al. (2012) and 5 years in both the Garshick et al. (2012a and Steenland et al. (1998), based on the best model fit in each study. The authors presented sensitivity analyses to lag choices in supplemental material, available online. Elemental carbon was measured as REC in DEMS, as SEC in Garshick et al. (2012a), and as EC in Steenland et al. (1998). Source: Vermeulen et al. 2014b.

Vermeulen and colleagues' decision to use the results from the models that best fit the data in the individual studies is a well-accepted approach in the absence of the ability to pool and standardize the underlying data. Nevertheless, both the investigators and other analysts have conducted sensitivity analyses to their choice of results and other assumption. The relative merits of using consistent lags from all studies in the meta-analysis, regardless of their appropriateness in individual studies, has been evaluated by Crump (2014) and debated in response by Vermeulen and colleagues (2014b). Morfeld and Spallek (2015) reported sensitivity of the meta-regression estimates to selected alternative assumptions about analytical methods, choice of study, and the choice of results from the studies. In the Panel's view, these sensitivity analyses have not undermined the basic findings and utility of the original study but provide a useful basis for a more systematic evaluation of meta-analytic choices.

Some other potential issues were not explored, such as harmonizing the SEC and REC metrics used in the different studies. The Panel thought it unlikely that the differences in particle size distributions between the metrics would lead to substantial differences in the effects of exposures to EC between the studies, given the similarities in lung deposition for particles in these size ranges (Kreyling et al. 2006). Data on both REC and SEC do exist in the DEMS that might be used to adjust one metric for the other so that a risk assessment could rely on a common exposure metric.

The ideal solution, if feasible, would be to pool the primary raw data from these studies, giving careful consideration to these various issues. However, based on past experience with pooled studies of residential radon, it is not necessarily clear that a pooled analysis would provide results dramatically different from the present meta-analysis reported by Vermeulen and colleagues (2014a); an early meta-analysis of residential radon studies conducted

by Lubin and Boice (1997) produced results that were compatible with a later pooled analysis conducted by Krewski and colleagues (2005). There would also be considerable challenges in reconciling covariates and establishing an internally consistent dataset.

All of these efforts demonstrate that the DEMS and the Truckers study have provided new sources of data with which to explore relationships between exposure to diesel exhaust and lung cancer risk in human populations. While there remains debate or uncertainty about what the 'right' models are or the predictions that follow from them, that in and of itself does not mean that these studies and their data are not useful. It is unrealistic to expect that individual results would be universally applicable or that all of the issues could be anticipated for extrapolating the results of the studies to other populations, time periods, and exposure conditions, including different diesel exhaust technologies. What is important is that they allow exploration and communication of the nature and magnitude of those uncertainties.

5.2.2 RECOMMENDATIONS ON THE VALUE OF ADDITIONAL STUDIES AND/OR ANALYSIS IN THESE DATA SETS

The Panel was asked to consider the usefulness of extending or conducting further analyses of existing data sets and for the design of new studies that would provide a stronger basis for risk assessment. The Panel had no further recommendations for major analyses that would need to be done before it could come to a conclusion about the use of these studies for quantitative risk assessment.

Similarly, the Panel thought it would be difficult to identify alternative research designs that would substantially improve on these two studies in the foreseeable future. As discussed above, the Truckers study and the DEMS had incorporated or embodied many of the earlier recommendations made by earlier HEI Diesel Epidemiology Panels (Table 5.1). Some of the major uncertainties in the studies arise from factors largely beyond the control of these investigators — and likely any future investigators — most notably the absence of or only partial historical exposure monitoring and other records to develop more accurate and precise estimates of exposure. Even if another well-designed prospective occupational cohort study were to be initiated today, with improvements such as detailed personal exposure monitoring for individual workers and follow up of each worker beginning at hire or at first exposure to reduce concerns about healthy worker survivor bias, it would take decades for results to become available. Further, with the dramatic reduction in exposure to diesel emissions due to the use of new diesel technologies and cleaner fuels, in order to have adequate statistical precision in the estimated exposure–response function, the number of workers that would need to be studied would likely be larger than that included in the currently available studies.

Given the number of questions raised about the elements of exposure assessment, the Panel thought it could be useful to develop a more explicit framework or model for exposure measurement error as part of the quantitative risk assessment process. Such a framework could serve as a more systematic basis for discussing and communicating how different types of error might affect the risk estimates and thus help in identifying where additional sensitivity analyses might be most useful.

The Panel also saw merit in the initiation of exposure monitoring programs to track trends in exposure to diesel emissions in the future. Data from such programs could be useful for better estimation of future exposure reductions and for evaluating concomitant reductions in human lung cancer risk while avoiding the need for the kinds of historical reconstructions of exposure that have received so much criticism in these and other occupational epidemiological studies. Monitoring programs could be targeted at both occupational groups that continue to be exposed to diesel emissions and to specific populations exposed at ambient levels in the general population. The series of four MATES studies that have been conducted since 1987 in southern California have been used effectively to track trends air pollution sources and levels and to evaluate the impact of regulatory actions to improve air quality and health (South Coast Air Quality Management District [SCAQMD] 2014).

5.2.3 CONSIDERATIONS FOR FUTURE QUANTITATIVE RISK ASSESSMENTS OF DIESEL EXHAUST

The Panel recognizes that its evaluation of the Truckers study and the DEMS is only one step in a more comprehensive risk assessment process for both characterization of the exposure–response relationship and its application in different risk management settings. The National Research Council (NRC) risk assessment–risk management paradigm introduced at the outset of this report (Figure 2.1) makes it clear that these steps are informed not only by a broad set of evidence but by the particular decision that must be made and its regulatory context.

It is unlikely, for example, that a single study or statistical model will provide the sole basis for all characterizations of the exposure–response relationship for diesel exhaust and lung cancer. Based on projected changes in the U.S. EPA Integrated Risk Information System program (U.S. EPA 2013),

recently reviewed favorably by the NRC (2014), exposure–response relationships will be described by a range or distribution of plausible models and model results, based on consideration of other relevant toxicologic, mechanistic, or other evidence. For example, in its recent review of diesel exhaust literature, the IARC examined the available mechanistic data and noted that diesel exhaust demonstrates a number of classical markers of genotoxicity, which is often thought to suggest a linear relationship between exposure and response. At the same time, the IARC identified effects of diesel exhaust on cell proliferation, which may contribute to nonlinearities in the exposure–response curve at higher levels of exposure. Such information could be taken into account in the evaluation of alternative approaches to characterization of the diesel exhaust–lung cancer exposure–response relationship. Bayesian methods, including Bayesian model averaging for informing model selection, have also been proposed for examining and combining results from different modeling approaches (NRC 2014).

Additional considerations in translating the results from these studies to other target populations include:

- generalizability of risk estimates from these predominantly healthy male, Caucasian workers to subpopulations thought to be more susceptible to the effects of exposure to diesel exhaust (e.g., children, elderly people, and those with preexisting comorbidities).
- differences in patterns of exposure at work (e.g., higher exposures, 40 hours a week for most of a working lifetime) compared with patterns more relevant to different occupations, or to the general population (e.g., lower exposures, possibly throughout the day or over a lifetime) and implications for risk. The TSCE model proposed by Moolgavkar and colleagues (2015) represents one such approach.

Chapter 1 of this report identified the other major factors that need to be considered in the use of any exposure–response relationships developed from these studies: the major changes in diesel fuels, engines and after-treatment technologies that have occurred since these studies were conducted, and the implications those changes have for ambient concentrations and composition of diesel emissions and the risk associated with them. These include:

- 99% reductions in PM mass emissions from 2007 and 2010 heavy-duty diesel engines relative to 1998 emissions standards.
- Reduction in EC's role as the predominant component of diesel PM from about 70% by mass in 2004 to 13% in the 2007 and 16% in 2010 engines. Coupled with the reductions in diesel particulate mass, the emissions of EC have dropped by 99% (HEI 2015).
- Substantial reductions in emissions of PAHs, nitroPAHs, metals and other compounds of that have been of concern due to their toxic or carcinogenic properties — about 80% for 2007 engines and 99% for 2010 engines relative to 2004 technology engines. (Khalek et al. 2011, 2015).
- Evidence comparing results from the California SCAQMD MATES IV study and the 2005 MATES III study that ambient diesel PM levels have dropped by about 70% between 2005 and 2014 (SCAQMD 2014).
- From the same studies, average PM_{10} EC measurements were 25% lower and $PM_{2.5}$ EC measurements were 35% lower (SCAQMD 2014).
- A study of chronic exposure of rodents to new technology diesel emissions from 2007 technology engines, found no evidence of carcinogenicity (McDonald et al. 2015). These rodents were exposed to much lower levels of EC and carcinogens (e.g., the PAHs, benzo-a-pyrene and benzo-e-pyrene, dioxin) than in previous studies given the improvements in technology described above.

The Truckers study and the DEMS both involved exposure predominantly to older diesel engines, particularly given the emphasis on exposures that occurred 10 to 15 years before the mortality from lung cancer. Consequently, the exposure–response relationships derived from these studies are most relevant to occupational and ambient settings where similar engine technology is in use or where the transition to newer fuels and engine technologies is less advanced. Complete turnover of the onroad heavy-duty diesel engine fleet from older to newer technology may take one to two decades in the United States and other developed countries (International Institute for Applied Systems Analysis 2012). It may take longer in occupational settings and in developing countries, where the rate of turnover historically has lagged for a number of reasons (International Council on Clean Transportation 2014). Some major developing countries (e.g., China), however, are accelerating the introduction of these new fuels and technologies. Risk assessments will need to consider data on fleet composition and turnover in assessing the contribution to emissions and ambient levels of EC from diesel engines.

5.3 CONCLUSIONS

The HEI Panel found that the epidemiological information that has accrued since the previous HEI panel reported on this issue in 1999 is both relevant and informative. The

occupational studies of nonmetal miners and workers in the trucking industry represent useful contributions by investigators who have worked carefully over extended periods of time to recreate historical exposure profiles and to describe exposure–response relationships between diesel exhaust and human lung cancer. Overall, these studies made considerable progress toward addressing the deficiencies that HEI had identified in the utility of earlier epidemiological research studies of diesel exhaust.

In undertaking its charge, the HEI Panel placed its detailed review of the Truckers study and the DEMS within the broader context of scientific research and the policy decisions that depend on it. Well-designed and executed epidemiological studies provide an important basis for decisions about the hazards of, and the exposure–response relationship associated with, exposures to particular agents, especially when compared to the far more common alternative, namely the use of animal studies to predict human health risk.

The detailed evaluations of these studies by IARC, the HEI Panel and other analysts lay the groundwork for a systematic characterization of the exposure–response relationship and associated uncertainties. In addition, the Panel has identified the analytical challenges that should be confronted in extrapolating the results from these studies to different populations and time periods, particularly given the rapid changes in diesel technology and its deployment around the world. The Panel concluded that the results and data from both the Truckers study and the DEMS can be usefully applied in quantitative risk assessments. The uncertainties within each study should be considered in any attempts to derive an exposure–response relationship.

References

Applebaum KM, Malloy EJ, Eisen EA. 2011. Left truncation, susceptibility, and bias in occupational cohort studies. Epidemiology 22:599–606.

Arlt VM. 2005. 3-Nitrobenzanthrone, a potential human cancer hazard in diesel exhaust and urban air pollution: a review of the evidence. Mutagenesis 20(6):399–410.

Attfield MD, Schleiff PL, Lubin JH, Blair A, Stewart PA, Vermeulen R, et al. 2012. The diesel exhaust in miners study: A cohort mortality study with emphasis on lung cancer. J Natl Cancer Inst 104:869–883.

Belinsky SA, Snow SS, Nikula KJ, Finch GL, Tellez CS, Palmisano WA. 2002. Aberrant CpG island methylation of the p16(INK4a) and estrogen receptor genes in rat lung tumors induced by particulate carcinogens. Carcinogenesis 23(2):335–339.

Bennette C, Vickers A. 2012. Against quantiles: Categorization of continuous variables in epidemiologic research, and its discontents. BMC Med Res Methodol 12:21.

Bezemer GF, Bauer SM, Oberdörster G, Breysse PN, Pieters RH, Georas SN, Williams MA. 2011. Activation of pulmonary dendritic cells and Th2-type inflammatory responses on instillation of engineered, environmental diesel emission source or ambient air pollutant particles in vivo. J Innate Immun 3(2):150–166. doi: 10.1159/000321725.

Birch ME, Cary RA. 1996. Elemental carbon-based method for occupational monitoring of particulate diesel exhaust: Methodology and exposure issues. Analyst 121:1183–1190.

Birch ME, Noll JD. 2004. Submicrometer elemental carbon as a selective measure of diesel particulate matter in coal mines. J Environ Monit 6:799–806.

Blot WJ, Brown LM, Pottern LM, Stone BJ, Fraumeni JF Jr. 1983. Lung cancer among long-term steel workers. Am J Epidemiol 117:706–716.

Boffetta P. 2012. Re: The Diesel Exhaust in Miners Study: A nested case–control study of lung cancer and diesel exhaust and a cohort mortality study with emphasis on lung cancer. J Natl Cancer Inst 104:1842–1843.

Borak J, Bunn WB, Chase GR, Hall TA, Head HJ, Hesterberg TW, et al. 2011. Comments on the diesel exhaust in miners study. Ann Occup Hyg 55:339–342.

Breslow NE, Day NE, Halvorsen KT, Prentice RL, Sabai C. 1978. Estimation of multiple relative risk functions in matched case-control studies. Am J Epidemiol108:299–307.

Bunn WB 3rd, Valberg PA, Slavin TJ, Lapin CA. 2002. What is new in diesel. Int Arch Occup Environ Health 75 Suppl:S122–132.

Cass GR, Conklin MH, Shah JJ, Huntzicker JJ, Macias ES. 1984. Elemental carbon concentrations: Estimation of an historical data base. Atmos Environ A 18: 153–162. ISSN 0004-6981.

Centers for Disease Control and Prevention. 2005. CDC WONDER. Available: http://wonder.cdc.gov/ [accessed 19 August 2005].

Chaloupka K, Harper N, Krishnan V, Santostefano M, Rodriguez LV, Safe S. 1993. Synergistic activity of polynuclear aromatic hydrocarbon mixtures as aryl hydrocarbon (ah) receptor agonists. Chemico-biological interactions 89:141–158.

Clark NN, Jarrett RP, Atkinson CM. 1999. Field measurements of particulate matter emissions, carbon monoxide, and exhaust opacity from heavy-duty diesel vehicles. J Air Waste Manage Assoc 49:76–84.

Coble JB, Stewart PA, Vermeulen R, Yereb D, Stanevich R, Blair A, et al. 2010. The Diesel Exhaust in Miners Study: II. Exposure monitoring surveys and development of exposure groups. Ann Occup Hyg 54:747–761.

Cohen HJ, Borak J, Hall T, Sirianni G, Chemerynski S. 2002. Exposure of miners to diesel exhaust particulates in underground nonmetal mines. Am Ind Hyg Assoc J 651–658.

CONCAWE. 2012. A global and historical perspective on the exposure characteristics of traditional and new technology diesel exhaust: Impact of technology on the physical and chemical characteristics of diesel exhaust emissions. Brussels, Belgium:CONCAWE.

Cox DR, Donnelly CA. 2011. Principles of Applied Statistics. Cambridge, U.K.:Cambridge University Press.

Crump K. 2014. Meta-Analysis of Lung Cancer Risk from Exposure to Diesel Exhaust: Study Limitations. Environ Health Perspect 122(9). *http://dx.doi.org/10.1289/ehp.1408482*.

Crump K, Van Landingham C. 2012. Evaluation of an exposure assessment used in epidemiological studies of diesel exhaust and lung cancer in underground mines. Crit Rev Toxicol 42:599–612.

Crump KS, Van Landingham C, McClellan R. In press. Influence of alternative exposure estimates in DEMS miners study: diesel exhaust and lung cancer. Risk Anal.

Crump KS, Van Landingham C, Moolgavkar SH, McClellan R. 2015. Reanalysis of the DEMS nested case-control study of lung cancer and diesel exhaust: suitability for quantitative risk analysis. Risk Anal 35(4):676–700. doi: 10.1111/risa. 12371.

Davis ME, Hart JE, Laden F, Garshick E, Smith TJ. 2011. A retrospective assessment of occupational exposure to elemental carbon in the U.S. trucking industry. Environ Health Perspect 119:997–1002.

Davis ME, Laden F, Hart JE, Garshick E, Blicharz A, Smith TJ. 2009. Predicting changes in PM exposure over time at U.S. Trucking terminals using structural equation modeling techniques. J Occup Environ Hyg 6:396–403.

Davis ME, Laden F, Hart JE, Garshick E, Smith TJ. 2010. Economic activity and trends in ambient air pollution. Environ Health Perspect 118:614–619.

Davis ME, Smith TJ, Laden F, Hart JE, Blicharz AP, Reaser P, et al. 2007. Driver exposure to combustion particles in the U.S. Trucking industry. J Occup Environ Hyg 4:848–854.

Davis ME, Smith TJ, Laden F, Hart JE, Ryan LM, Garshick E. 2006. Modeling particle exposure in U.S. Trucking terminals. Environ Sci Technol 40:4226–4232.

Department of Labor, Mine Safety and Health Administration. 2005. 70 FR 32868: Diesel particulate matter exposure of underground metal and nonmetal miners. Federal Register volume 70, issue 107.

Diaz-Sanchez D, Dotson AR, Takenaka H, Saxon A. 1994. Diesel exhaust particles induce local IgE production in vivo and alter the pattern of IgE messenger RNA isoforms. J Clin Invest 94(4):1417–1425.

Doll R. 1985. Occupational cancer: a hazard for epidemiologists. Int J Epidemiol 14:22–31.

Dosemeci M, Chen JQ, Hearl F, Chen RG, McCawley M, Wu Z, et al. 1993. Estimating historical exposure to silica among mine and pottery workers in the People's Republic of China. Am J Ind Med 24:55–66.

Droste JH, Weyler JJ, Van Meerbeeck JP, Vermeire PA, van Sprundel MP. 1999. Occupational risk factors of lung cancer: A hospital based case-control study. Occup Environ Med 56:322–327.

Fann N, Bell ML, Walker K, Hubbell B. 2011. Improving the linkages between air pollution epidemiology and quantitative risk assessment. Environ Health Perspect 119:1671–1675.

Federal Focus. 1995. Principles for evaluating epidemiologic data in regulatory risk assessment. Washington D.C.:Federal Focus.

Ensell MX, Whong WZ, Heng ZC, Nath J, Ong T. 1998. Activation of pulmonary dendritic cells and Th2-type inflammatory responses on instillation of engineered, environmental diesel emission source or ambient air pollutant particles in vivo. Mutat Res 412(3):283–291.

Fertuck KC, Matthews JB, Zacharewski TR. 2001. Hydroxylated benzo[a]pyrene metabolites are responsible for in vitro estrogen receptor-mediated gene expression induced by benzo[a]pyrene, but do not elicit uterotrophic effects in vivo. Toxicol Sci 59:231–240.

Fisher LD, Lin DY. 1999. Time-dependent covariates in the cox proportional-hazards regression model. Annual Review of Public Health 20:145–157.

Flanders WD, Eldridge RC, McClellan W. 2014. A nearly unavoidable mechanism for collider bias with index-event studies. Epidemiology 25:762–764.

Friesen MC, Coble JB, Lu W, Shu XO, Ji BT, Xue S, Portengen L, et al. 2012. Combining a Job-Exposure Matrix with Exposure Measurements to Assess Occupational Exposure to Benzene in a Population Cohort in Shanghai, China. Ann Occup Hyg 56 (1): 80–91. doi:10.1093/annhyg/mer080.

Friesen MC, Macnab YC, Marion SA, Demers PA, Davies HW, Teschke K. 2006. Mixed models and empirical Bayes estimation for retrospective exposure assessment of dust exposures in Canadian sawmills. Ann Occup Hyg 50:281–288.

Fritz SG, Bailey CR, Scarbro CA, Somers JH. 2001. Heavy-duty diesel truck in-use emission test program for model years 1950 through 1975. 0148-7191.SAE Technical Paper.

Furuta C, Noda S, Li C, Suzuki AK, Taneda S, Watanabe G, et al. 2008. Nitrophenols isolated from diesel exhaust particles regulate steroidogenic gene expression and steroid synthesis in the human h295r adrenocortical cell line. Toxicol Appl Pharmacol 229:109–120.

Gamble JF, Nicolich MJ, Boffetta P. 2012. Lung cancer and diesel exhaust: An updated critical review of the occupational epidemiology literature. Crit Rev Toxicol 42:549–598.

Garshick E, Laden F, Hart JE, Davis ME, Eisen EA, Smith TJ. 2012a. Lung cancer and elemental carbon exposure in trucking industry workers. Environ Health Perspect 120:1301–1306.

Garshick E, Laden F, Hart JE, Davis ME, Eisen EA, Smith TJ. 2012b. Cox Model Setup: Garshick et al. Respond. Environ Health Perspect 120:A382–A383. *http://dx.doi.org/10.1289/ehp.1205573R* [Online 1 October 2012].

Garshick E, Laden F, Hart JE, Rosner B, Davis ME, Eisen EA, et al. 2008. Lung cancer and vehicle exhaust in trucking industry workers. Environ Health Perspect 116:1327–1332.

Garshick E, Schenker MB, Munoz A, Segal M, Smith TJ, Woskie SR, et al. 1987. A case-control study of lung cancer and diesel exhaust exposure in railroad workers. Am Rev Resp Dis 135:1242–1248.

Garshick E, Schenker MB, Munoz A, Segal M, Smith TJ, Woskie SR, et al. 1988. A retrospective cohort study of lung cancer and diesel exhaust exposure in railroad workers. Am Rev Resp Dis 137:820–825.

Glass DC, Adams GG, Manuell RW, Bisby JA. 2000. Retrospective exposure assessment for benzene in the Australian petroleum industry. Ann Occup Hyg 44:301–320.

Greenland S. 1986. Partial and marginal matching in case-control studies. In: Moolgavkar SH, Prentice RL (eds). Modern statistical methods in chronic disease epidemiology. New York, NY:Wiley & Sons 35–49.

Greenland S. 1997. Re: 'Estimating relative risk functions in case-control studies using a nonparametric logistic regression' (letter). Am J Epidemiol 146:883–884.

Greenland S. 2008. Introduction to regression modeling. In: Modern Epidemiology. Third edition. (Rothman KJ, Greenland S, Lash TL. eds.) pp 434–435. Philadelphia:Lippincott-Raven.

Greenland S, Lash TL. 2008. In: Modern Epidemiology. Third edition. (Rothman KJ, Greenland S, Lash TL. eds.) pp 283–284. Philadelphia:Lippincott-Raven

Gustavsson P, Jakobsson R, Nyberg F, Pershagen G, Jarup L, Scheele P. 2000. Occupational exposure and lung cancer risk: A population-based case-referent study in Sweden. Am J Epidemiol 152:32–40.

Hammond SK, Hines CJ, Hallock MF, Woskie SR, Abdollahzadeh S, Iden CR, et al. 1995. Tiered exposure-assessment strategy in the semiconductor health study. Am J Ind Med 28:661–680.

Haney HA, Saseen GP. 2000. Estimation of diesel particulate concentrations in underground mines. Mining Engineering 52:60–64.

Health Effects Institute. 1999. Diesel Workshop: Building a Research Strategy to Improve Risk Assessment. Communication 7. Cambridge, MA:Health Effects Institute.

Health Effects Institute. 2000. Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality. A Special Report of the Institute's Particle Epidemiology Reanalysis Project. Cambridge, MA:Health Effects Institute.

Health Effects Institute. 2003. Research on diesel exhaust and other particles. Program Summary. Cambridge, MA: Health Effects Institute.

Health Effects Institute. 2007. Mobile-Source Air Toxics: A Critical Review of the Literature on Exposure and Health Effects. Special Report 16. Boston, MA:Health Effects Institute.

Health Effects Institute, Subcommittee on Emerging Fuels and Technology. 2011. The Future of Vehicle Fuels and Technologies: Anticipating Health Benefits and Challenges. Communication 16. Boston, MA:Health Effects Institute.

Health Effects Institute. 2015. Executive Summary. The Advanced Collaborative Emissions Study (ACES). Boston, MA:Health Effects Institute.

HEI Diesel Epidemiology Expert Panel. 1999. Diesel Emissions and Lung Cancer: Epidemiology and Quantitative Risk Assessment. Special Report. Cambridge, MA:Health Effects Institute.

HEI Diesel Epidemiology Working Group. 2002. Research Directions to Improve Estimates of Human Exposure and Risk from Diesel Exhaust. Cambridge, MA:Health Effects Institute.

HEI Diesel Working Group. 1995. Diesel Exhaust: Critical Analysis of Emissions, Exposure, and Health Effects. Cambridge, MA:Health Effects Institute.

Hesterberg TW, Bunn WB, McClellan RO, Hart GA, Lapin CA. 2005. Carcinogenicity studies of diesel engine exhausts in laboratory animals: a review of past studies and a discussion of future research needs. Crit Rev Toxicol 35:379–411.

Hesterberg TW, Bunn WB 3rd, Chase GR, Valberg PA, Slavin TJ, Lapin CA, et al. 2006. A critical assessment of studies on the carcinogenic potential of diesel exhaust. Crit Rev Toxicol 36:727–776.

Hesterberg TW, Long CM, Bunn WB, Lapin CA, McClellan RO, Valberg PA. 2012a. Health effects research and regulation

of diesel exhaust: An historical overview focused on lung cancer risk. Inhal Toxicol 24 Suppl 1:1–45.

Hesterberg TW, Long CM, Valberg PA. 2012b. Re: The Diesel Exhaust in Miners Study: A nested case–control study of lung cancer and diesel exhaust and a cohort mortality study with emphasis on lung cancer. J Natl Cancer Inst 104:1841.

Hornung RW, Greife AL, Stayner LT, Steenland NK, Herrick RF, Elliott LJ, et al. 1994. Statistical model for prediction of retrospective exposure to ethylene oxide in an occupational mortality study. Am J Ind Med 25:825–836.

Institute of Medicine. 2013. Environmental decisions in the face of uncertainty. Washington, DC:National Academies Press.

International Agency for Research on Cancer. 1989. Diesel and Gasoline Engine Exhausts and Some Nitroarenes. Vol. 46. Monographs on the Evaluation of Carcinogenic Risks to Humans. World Health Organization, International Agency for Research on Cancer, Lyon, France.

International Agency for Research on Cancer. 2012. IARC: Diesel engine exhaust carcinogenic. Press Release 213. Lyon, France:IARC. Available: *www.iarc.fr/en/media-centre/pr/2012/pdfs/pr213_E.pdf*.

International Agency for Research on Cancer. 2014. Diesel and Gasoline Engine Exhausts and Some Nitroarenes. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Volume 105. Lyon, France:IARC. Available: *http://monographs.iarc.fr/ENG/Monographs/vol105/mono105.pdf*.

International Council on Clean Transportation. 2014. The State of Clean Transport Policy: A 2014 synthesis of vehicle and fuel policy developments. Washington, DC:ICCT. Available: *www.theicct.org/state-of-clean-transport-policy-2014*.

International Energy Agency. 2012. World energy outlook 2012. Paris, France:IEA.

International Institute for Applied Systems Analysis. 2012. The potential for further controls of emissions from mobile sources in Europe. TSAP Report #4 Version 2.0. Schloss Laxenburg, Austria:IIASA.

Jain NB, Hart JE, Smith TJ, Garshick E, Laden F. 2006. Smoking behavior in trucking industry workers. Am J Ind Med 49:1013–1020.

Kauppinen T, Toikkanen J, Pukkala E. 1998. "From Cross-Tabulations to Multipurpose Exposure Information Systems: A New Job-Exposure Matrix." *American Journal of Industrial Medicine* 33 (4): 409–17.

Khalek IA, Blanks MG, Merritt PM, Zielinska B. 2015. Regulated and unregulated emissions from modern 2010 emissions-compliant heavy-duty on-highway diesel engines. J Air Waste Manag Assoc 65(8):987–1001.

Khalek IA, Bougher TL, Merritt PM, Zielinska B. 2011. Regulated and unregulated emissions from highway heavy-duty diesel engines complying with U.S. Environmental Protection Agency 2007 emissions standards. J Air Waste Manag Assoc 61:427–442.

Kirchstetter TW, Agular J, Tonse S, Fairley D, Novakov T. 2008. Black carbon concentrations and diesel vehicle emission factors derived from coefficient of haze measurements in California: 1967–2003. Atmos Environ 42:480–491.

Krewski D, Lubin JH, Zielinski JM, Alavanja M, Catalan VS, Field RW, et al. 2005. Residential radon and risk of lung cancer: a combined analysis of 7 North American case-control studies. Epidemiology 16(2):137–145.

Krewski D, Wigle D, Clayson DB, Howe GR. 1990. Role of epidemiology in health risk assessment. Recent Results Cancer Res 120:1–24.

Krewski D, Zielinski JM, Hazelton WD, Garner MJ, Moolgavkar SH. 2003. The use of biologically based cancer risk models in radiation epidemiology. Radiat Prot Dosimetry 104:367–376.

Kreyling W, Semmler-Behnke M, Möller W. 2006. Health implications of nanoparticles. J Nanopart Res 8:543–562.

Laden F, Hart JE, Smith TJ, Davis ME, Garshick E. 2007. Cause-specific mortality in the unionized U.S. Trucking industry. Environ Health Perspect 115:1192–1196.

Langholz B, Goldstein L. 1996. Risk set sampling in epidemiologic cohort studies. Stat Sci 11:35–53.

Lavoue J, Pintos J, Van Tongeren M, Kincl L, Richardson L, Kauppinen T, et al. 2012. Comparison of exposure estimates in the Finnish job-exposure matrix FINJEM with a JEM derived from expert assessments performed in Montreal. Occup Environ Med 69:465–471.

Li J, Kanju P, Patterson M, Chew WL, Cho SH, Gilmour I, et al. 2011. Trpv4-mediated calcium influx into human bronchial epithelia upon exposure to diesel exhaust particles. Environ Health Perspect 119:784–793.

Lippmann M, Chen L-C, Gordon T, Ito K, Thurston GD. 2013. National Particle Component Toxicity (NPACT) Initiative: Integrated Epidemiologic and Toxicologic Studies of the Health Effects of Particulate Matter Components. Research Report 177. Boston, MA:Health Effects Institute.

Liu J, Ballaney M, Al-alem U, Quan C, Jin X, Perera F, Chen LC, Miller RL. 2008. Combined inhaled diesel exhaust particles and allergen exposure alter methylation of T helper genes and IgE production in vivo. Toxicol Sci 102(1):76–81.

Loomis D, Guha N, Straif K. 2014. The contribution of occupational studies to the identification of carcinogens in the IARC monographs and how to further improve it. Occup Environ Med 71 Suppl 1:A41.

Lubin JH, Boice JD Jr. 1997. Lung cancer risk from residential radon: Meta-analysis of eight epidemiologic studies. J Natl Cancer Inst 89:49–57.

Lubin JH, Caporaso NE. 2013. Misunderstandings in the misconception on the use of pack-years in analysis of smoking. Br J Cancer 108:1218–1220.

McClellan RO. 2012. Re: The diesel exhaust in miners study: A nested case-control study of lung cancer and diesel exhaust, a cohort mortality study with emphasis on lung cancer, and the problem with diesel. J Natl Cancer Inst 104:1843–1845; author reply 1847–1849.

McClellan RO, Hesterberg TW, Wall JC. 2012. Evaluation of carcinogenic hazard of diesel engine exhaust needs to consider revolutionary changes in diesel technology. Regul Toxicol Pharmacol 63:225–258.

McDonald JD, Doyle-Eisele M, Seagrave J, Gigliotti AP, Chow J, Zielinska B, et al. 2015. Part 1. Assessment of carcinogenicity and biologic responses in rats after lifetime inhalation of new-technology diesel exhaust in the ACES bioassay. In: Advanced Collaborative Emissions Study (ACES) Lifetime Cancer and Non-Cancer Assessment in Rats Exposed to New-Technology Diesel Exhaust. Research Report 184. Boston, MA:Health Effects Institute.

Moolgavkar SH, Chang ET, Luebeck G, Lau EC, Watson HN, Crump KS, et al. 2015. Diesel engine exhaust and lung cancer mortality – time-related factors in exposure and risk. Risk Analysis: doi: 10.1111/risa.12315 [Online 13 February 2015]. Available: *http://dx.doi.org/10.1111/risa.12315.*

Moolgavkar SH, Krewski D, Zeise L, Cardis E, Moller H, eds. 1999. Quantitative Estimation and Prediction of Human Cancer Risks. Lyon, France:IARC Press.

Moolgavkar SH, Luebeck EG, Krewski D, Zielinski JM. 1993. Radon, cigarette smoke, and lung cancer: A re-analysis of the Colorado plateau uranium miners' data. Epidemiology 4:204–217.

Morabia A, Markowitz S, Garibaldi K, Wynder EL. 1992. Lung cancer and occupation: Results of a multicentre case-control study. Br J Ind Med 49:721–727.

Morfeld P. 2012a. Cox Model Setup May Lead to Erroneous Conclusions. Environ Health Perspect 120:a382–a382. *http://dx.doi.org/10.1289/ehp.1205573* [Online 1 October 2012].

Morfeld P. 2012b. Diesel exhaust in miners study: How to understand the findings? J Occup Med Toxicol 7:10.

Morfeld P, Erren TC. 2012. Exposures to diesel motor exhaust and lung cancer: Are findings of a large pooled analysis really consistent? Am J Respir Crit Care Med 185:104; author reply 106–107.

Morfeld P, Spallek M. 2015. Diesel engine exhaust and lung cancer risks — evaluation of the meta-analysis by Vermeulen et al. 2014. J Occup Med Toxicol 10:31. doi: 10.1186/s12995-015-0073-6. Available: *www.occup-med.com/content/10/1/31*.

Morgan MG, Henrion M. 1990. Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis. New York, NY:Cambridge University Press.

National Academy of Sciences. 1999. Health Risks of Exposure to Radon: BEIR VI. Committee on Health Risks of Exposure to Radon, Board on Radiation Effects Research, Commission on Life Sciences, National Research Council. Washington, DC:National Academies Press.

National Cancer Institute–National Institute for Occupational Safety and Health. 1997. A cohort mortality study with a nested case-control study of lung cancer and diesel exhaust among non-metal miners (NCI–NIOSH Interagency Project Protocol, 16 September 1997). [Source: Debra Silverman, Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI Shady Grove, Room 6E142, Rockville, MD 20892; (240) 276–7174; *silvermd@exchange.nih.gov.*].

National Institute of Occupational Safety and Health. 1988. Carcinogenic Effects of Exposure to Diesel Exhaust. NIOSH Current Intelligence Bulletin. Department of Health and Human Services (NIOSH) Publication No. 88-116. Atlanta, GA:Centers for Disease Control.

National Institute for Occupational Safety and Health. 1998. NIOSH manual of analytical methods. Atlanta, GA:Centers for Disease Control.

National Research Council. 1983. Risk assessment in the federal government: Managing the process. Washington, DC:National Academies Press.

National Research Council. 1993. Issues in Risk Assessment. Washington, DC:National Academies Press. Available: *www.nap.edu/catalog/2078/issues-in-risk-assessment*.

National Research Council. 1994. Science and Judgement in Risk Assessment. Committee on Risk Assessment of Hazardous Air Pollutants. Washington, DC:National Academies Press.

National Research Council. 1996. Understanding Risk: Informing Decisions in a Democratic Society. Committee on Risk Characterization. Washington, DC: National Academies Press.

National Research Council. 1999. Health Effects of Exposure to Radon: BEIR VI. Washington, DC:National Academies Press. Available: *www.nap.edu/openbook.php?isbn=0309056454*.

National Research Council. 2009. Science and decisions: Advancing Risk Assessment. Committee on Improving Risk Analysis Approaches Used by the U.S. EPA. Washington, DC:National Academies Press.

National Research Council. 2014. Review of EPA's Integrated Risk Information System (IRIS) Process. Washington, DC:National Academies Press.

National Research Council, Health Effects Panel of the Diesel Impacts Study Committee. 1981. Health Effects of Exposure to Diesel Exhaust: Impacts of Diesel-Powered Light-Duty Vehicles. Washington, DC:National Academies Press.

National Toxicology Program. 2011. Report on Carcinogens, Twelfth Edition. Research Triangle Park:U.S. Department of Health and Human Services, Public Health Services.

National Toxicology Program, Office of Health Assessment and Translation. 2013. Draft OHAT approach for systematic review and evidence integration for literature-based health assessments. Washington, DC:National Toxicology Program.

Office of Environmental Health Hazard Assessment. 1998. Health Risk Assessment for Diesel Exhaust, February 1998. Sacramento, CA:California Environmental Protection Agency.

Owens CV Jr., Lambright C, Cardon M, Gray LE Jr., Gullett BK, Wilson VS. 2006. Detection of androgenic activity in emissions from diesel fuel and biomass combustion. Society for Environmental Toxicology and Chemistry (SETAC) 25:2123–2131.

Peto J. 2012. That the effects of smoking should be measured in pack-years: Misconceptions 4. Br J Cancer 107:406–407.

Peto J. 2013. Reply: Misunderstandings in the misconception on the use of pack-years in analysis of smoking. Br J Cancer 108:1221.

Pocock SJ, Assmann SE, Enos LE, Kasten LE. 2002. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Statist Med 21:2917–2930. doi: 10.1002/sim.1296.

Pomroy WH, Saseen GP. 2008. MSHA's DPM rule for metal and nonmetal mines. In: 12th U.S./North American Mine Ventilation Symposium 2008 (Wallace, KG, ed). Clovis, CA:Mine Ventilation Services, 455–462. Available: *www.mvsengineering.com/index.php/products/97.*

Preston D, Shimizu D, Pierce DA, Suyama A, Mabuchi K. 2003. Studies of Mortality of Atomic Bomb Survivors. Report 13: Solid Cancer and Noncancer Disease Mortality: 1950–1997. Radiation Research 160 (4): 381–407.

Preston DL, Ron E, Tokuoka S, Funamoto S, Nishi N, Soda M, Mabuchi K, Kodama K. 2007. Solid Cancer Incidence in Atomic Bomb Survivors: 1958–1998. Radiation Research 168 (1):1–64. doi:10.1667/RR0763.1.

Riddle SG, Robert MA, Jakober CA, Hannigan MP, Kleeman MJ. 2008. Size-resolved source apportionment of airborne particle mass in a roadside environment. Environ Sci Technol 42:6580–6586.

Rothman K, Greenland S. 1998. Modern Epidemiology. Philadelphia, PA:Lippincott-Raven.

Rothman KJ. 2012. Epidemiology: An Introduction: New York, NY:Oxford University Press.

Sawada H, Kodama K, Shimizu Y, Kato H. 1986. Adult Health Study Report 6: Results of Six Examination Cycles, 1968–80, Hiroshima and Nagasaki. Hiroshima: Radiation Effects Research Foundation.

Schoenberg JB, Stemhagen A, Mason TJ, Patterson J, Bill J, Altman R. 1987. Occupation and lung cancer risk among New Jersey white males. J Natl Cancer Instit 79:13–21.

Schwartz J. 2002. The use of epidemiology in environmental risk assessment. Hum Ecol Risk Assess 8:1253–1265.

Seixas NS, Moulton LH, Robins TG, Rice CH, Attfield MD, Zellers ET. 1991. Estimation of cumulative exposures for the National Study of Coal Workers' Pneumoconiosis. Appl Occup Eviron Hyg 6:1032–1041.

Shaw CA, Robertson S, Miller MR, Duffin R, Tabor CM, Donaldson K, Newby DE, Hadoke PW. 2011. Diesel

exhaust particulate-exposed macrophages cause marked endothelial cell activation. Am J Respir Cell Mol Biol 44(6):840–851. doi: 10.1165/rcmb.2010-0011OC.

Sheesley RJ, Schauer JJ, Garshick E, Laden F, Smith TJ, Blicharz AP, et al. 2009. Tracking personal exposure to particulate diesel exhaust in a diesel freight terminal using organic tracer analysis. J Expo Sci Environ Epidemiol 19:172–186.

Sheesley RJ, Schauer JJ, Smith TJ, Garshick E, Laden F, Marr LC, et al. 2008. Assessment of diesel particulate matter exposure in the workplace: Freight terminals. J Environ Monit 10:305–314.

Silverman DT, Attfield MD. 2012. Response to: 'The Diesel Exhaust in Miners Study: A nested case-control study of lung cancer and diesel exhaust' and 'The Diesel Exhaust in Miners Study: A cohort mortality study with emphasis on lung cancer.' J Natl Cancer Inst 104:1848–1849.

Silverman DT, Attfield M. 2013. Letter to the Editor; 'Diesel exhaust in miners study: how to understand the findings? Peter Morfeld.' J Occup Med Tox Epub. *www .occup-med.com/content/7/1/10/comments#1392696*.

Silverman DT, Lubin JH, Blair AE, Vermeulen R, Stewart PA, Schleiff PL, et al. 2014. Re: The diesel exhaust in miners study (DEMS): a nested case-control study of lung cancer and diesel exhaust. J Natl Cancer Inst 106:ii.

Silverman DT, Samanic CM, Lubin JH, Blair AE, Stewart PA, Vermeulen R, et al. 2012. The diesel exhaust in miners study: a nested case-control study of lung cancer and diesel exhaust. J Natl Cancer Inst 104:(11):855–868.

Smith TJ, Davis ME, Reaser P, Natkin J, Hart JE, Laden F, et al. 2006. Overview of particulate exposures in the U.S. trucking industry. J Environ Monit 8:711–720.

Soulakova J, Davis WW, Hartman A, Gibson J. 2009. The impact of survey and response modes on current smoking prevalence estimates using TUS-CPS:1992–2003. Surv Res Methods 3(3):123–137.

South Coast Air Quality Management District. 2014. Multiple Air Toxics Exposure Study in the South Coast Air Basin (MATES IV). Draft Report. Diamond Bar, California:SCAQMD.

Spallek MF, Morfeld P. 2012. Re: The problem with diesel. J Natl Cancer Inst 104:1845–1846.

Stayner L, Smith R, Bailer J, Luebeck EG, Moolgavkar SH. 1995. Modeling epidemiologic studies of occupational cohorts for the quantitative assessment of carcinogenic hazards. Am J Ind Med 27:155–170.

Stanevich RS, Hintz P, Yereb D, Dosemeci M, Silverman DT. 1997. Elemental carbon levels at a potash mine. Appl Occup Environ Hyg 12:1009–1012.

Steenland K, Deddens J, Stayner L. 1998. Diesel exhaust and lung cancer in the trucking industry: Exposure-response analyses and risk assessment. Am J Ind Med 34:220–228.

Steenland K, Greenland S. 2004. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. Am J Epidemiol 160:384–392.

Steenland K, Schnorr T. 1988. Availability and accuracy of cancer and smoking data obtained from next of kin for decedents in a retrospective cohort study. J Occup Med. 30(4):348–353.

Steenland K, Silverman D, Hornung R. 1990. Case-control study of lung cancer and truck driving in the teamsters union. Am J Public Health 80:670–674.

Steenland K, Silverman D, Zaebst D. 1992. Exposure to diesel exhaust in the trucking industry and possible relationships with lung cancer. Am J Ind Med 21:887–890.

Stewart PA, Blair A. 1994. Women in the formaldehyde industry: Their exposures and their jobs. J Occup Med 36:918–923.

Stewart PA, Coble JB, Vermeulen R, Blair A, Lubin J, Attfield M, et al. 2011. Response to Borak et al. 2010 on the Diesel Exhaust in Miners Study. Ann Occup Hyg 55:343.

Stewart PA, Coble JB, Vermeulen R, Schleiff P, Blair A, Lubin J, et al. 2010. The Diesel Exhaust in Miners Study: I. Overview of the exposure assessment process. Ann Occup Hyg 54:728–746.

Stewart PA, Schairer C, Blair A. 1990. Comparison of jobs, exposures, and mortality risks for short-term and long-term workers. J Occup Med 32:703–708.

Stewart PA, Vermeulen R, Coble JB, Blair A, Schleiff P, Lubin JH, et al. 2012. The Diesel Exhaust in Miners Study: V. Evaluation of the exposure assessment methods. Ann Occup Hyg 56:389–400.

Sullivan SG, Greenland S. 2013. Bayesian regression in SAS software. Int J Epidemiol 42:308–317.

Sutton GW, Weems GW, Schutz LA, Trabant GC. 1979. Summary report of the environmental results of the MSHA and NIOSH silica/diesel exhaust study. In: Industrial Hygiene for Mining and Tunneling: Proceedings of a Topical Symposium, November 6 and 7, 1978. Denver, Colorado.

Cincinnati, OH:American Conference of Governmental Industrial Hygienists.

Tse L-A, Yu IT-S. 2012. Re: The Diesel Exhaust in Miners Study: A nested case–control study of lung cancer and diesel exhaust. J Natl Cancer Inst 104:1843.

Turner MC, Wigle DT, Krewski D. 2010. Residential pesticides and childhood leukemia: A systematic review and meta-analysis. Environ Health Perspect 118:33–41.

U.S. Environmental Protection Agency. 2000. Control of Air Pollution from New Motor Vehicles: Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements; Final Rule. 40 CFR, Parts 80, 85, and 86.

U.S. Environmental Protection Agency. 2001. Control of Air Pollution from New Motor Vehicles: Heavy-Duty Engine and Vehicle Standards and Highway Diesel Fuel Sulfur Control Requirements; Final Rule. 40 CFR, Parts 69, 80, and 86.

U.S. Environmental Protection Agency. 2002. Health assessment document for diesel engine exhaust. EPA/600/8-90/057F. Washington, DC:National Center for Environmental Assessment.

U.S. Environmental Protection Agency. 2005. Guidelines for carcinogen risk assessment. EPA/630/P-03/001F. Washington, DC:U.S. EPA.

U.S. Environmental Protection Agency. 2013. Part 1. Status of Implementation of Recommendations. Materials Submitted to the National Research Council, by Integrated Risk Information System Program, U.S. Environmental Protection Agency, January 30, 2013 [online]. Available: *www.epa.gov/Iris/pdfs/IRIS%20Program%20Materials%20to%20NRC_Part%201.pdf* [accessed Nov. 13, 2013].

Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. 2007. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology 18(6):805–835.

Vedal S, Campen MJ, McDonald JD, Kaufman JD, Larson TV, Sampson PD, Sheppard L, Simpson CD, Szpiro AA. 2013. National Particle Component Toxicity (NPACT) Initiative Report on Cardiovascular Effects. Research Report 178. Boston, MA:Health Effects Institute.

Vermeulen R, Coble JB, Lubin JH, Portengen L, Blair A, Attfield MD, et al. 2010a. The Diesel Exhaust in Miners Study: IV. Estimating historical exposures to diesel exhaust in underground non-metal mining facilities. Ann Occup Hyg 54:774–788.

Vermeulen R, Coble JB, Yereb D, Lubin JH, Blair A, Portengen L, et al. 2010b. The Diesel Exhaust in Miners Study: III. Interrelations between respirable elemental carbon and gaseous and particulate components of diesel exhaust derived from area sampling in underground non-metal mining facilities. Ann Occup Hyg 54:762–773.

Vermeulen R, Portengen L, Silverman DT, Garshick E, Steenland K. 2014a. Meta-analysis of lung cancer risk from exposure to diesel exhaust: Vermeulen et al. respond. Environ Health Perspect 122:A230-1. doi: 10.1289/ehp.1408428R.

Vermeulen R, Silverman DT, Garshick E, Vlaanderen J, Portengen L, Steenland K. 2014b. Exposure–response estimates for diesel engine exhaust and lung cancer mortality based on data from three occupational cohorts. Environ Health Perspect 122:172–177. doi: 10.1289/ehp.1306880.

Vlaanderen J, Vermeulen R, Heederik D, Kromhout H. 2008. Guidelines to evaluate human observational studies for quantitative risk assessment. Environ Health Perspect 116:1700–1705.

Vrabie CM, Jonker MT, Murk AJ. 2009. Specific in vitro toxicity of crude and refined petroleum products. 1. Aryl hydrocarbon receptor-mediated responses. Environmental toxicology and chemistry / SETAC 28:1995–2003.

Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. 2007. Statistics in medicine: Reporting of subgroup analyses in clinical trials. N Engl J Med 357:2189–2194.

Wells AJ, English PB, Posner SF, Wagenknecht LE, Perez-Stable EJ. 1998. Review. Misclassification rates for current smokers misclassified as nonsmokers. Am J Public Health 88(10):1503–1509.

Wolff GT, Stroup CM, Stroup DP. 1983. The coefficient of haze as a measure of particulate elemental carbon. J Air Pollut Control Assoc 33:746–750.

Williams P, Paustenbach D, Balzer JL, Mangold C. 2007. Retrospective exposure assessment of airborne asbestos related to skilled craftsmen at a petroleum refinery in Beaumont, Texas (1940–2006). J Toxicol Environ Health A 70:1076–1107.

World Health Organization. 1996. Diesel Fuel and Exhaust Emissions: International Program on Chemical Safety. Geneva, Switzerland:World Health Organization.

World Health Organization. 2005. Evaluation and use of epidemiological evidence for environmental health risk assessment. EUR/00/5020369. Copenhagen, Denmark: WHO Regional Office for Europe.

Yanowitz J, McCormick RL, Graboski MS. 2000. In-use emissions from heavy-duty diesel vehicles. Environ Sci Technol 34:729–740.

Zaebst DD, Clapp DE, Blade LM, Marlow DA, Steenland K, Hornung RW, et al. 1991. Quantitative determination of trucking industry workers' exposures to diesel exhaust particles. Am Ind Hyg Assoc 52:529–541.

Zielinska B, Samy S, McDonald JD, Seagrave J. 2010. Atmospheric transformation of diesel emissions. Research Report 147. Boston, MA:Health Effects Institute.

Zuurbier M, Hoek G, Oldenwening M, Lenters V, Meliefste K, van den Hazel P, et al. 2010. Commuters' exposure to particulate matter air pollution is affected by mode of transport, fuel type, and route. Environ Health Perspect 118:783–789.

Appendices

APPENDIX A: Related HEI Publications on Diesel Exhaust

Number	Title	Principal Investigator	Date
Communications			
17	Advanced Collaborative Emissions Study (ACES) Phase 3A: Characterization of U.S. 2007-Compliant Diesel Engine and Exposure System Operation	J. Mauderly	2012
16	The Future of Vehicle Fuels and Technologies: Anticipating Health Benefits and Challenges	Health Effects Institute	2011
10	Improving Estimates of Diesel and Other Emissions for Epidemiologic Studies	Health Effects Institute	2003
9	Evaluation of Human Health Risk from Cerium Added to Diesel Fuel	Health Effects Institute	2001
7	Diesel Workshop: Building a Research Strategy to Improve Risk Assessment	Health Effects Institute	1999
Program Summaries			
	Research on Diesel Exhaust and Other Particles	Health Effects Institute	2003
	Research on Diesel Exhaust	Health Effects Institute	1999
Research Reports			
184	Advanced Collaborative Emissions Study (ACES): Lifetime Cancer and Non-Cancer Assessment in Rats Exposed to New-Technology Diesel Exhaust.	J. McDonald, J. Bemis, L. Hallberg, D. Conklin and M. Kong	2015
166	Advanced Collaborative Emissions Study (ACES) Subchronic Exposure Results: Biologic Responses in Rats and Mice and Assessment of Genotoxicity	J. McDonald, J. Bemis, L. Hallberg, D. Conklin and M. Kong	2012
165	Allergic Inflammation in the Human Lower Respiratory Tract Affected by Exposure to Diesel Exhaust	M. Riedl	2012
160	Personal and Ambient Exposures to Air Toxics in Camden, New Jersey	P.J. Lioy	2011
159	Role of Neprilysin in Airway Inflammation Induced by Diesel Exhaust Emissions	S. Wong	2011
158	Air Toxics Exposure from Vehicle Emissions at a U.S. Border Crossing: Buffalo Peace Bridge Study	J.D. Spengler	2011
156	Concentrations of Air Toxics in Motor Vehicle–Dominated Environments	E.M. Fujita	2011
151	Pulmonary Effects of Inhaled Diesel Exhaust in Young and Old Mice: A Pilot Project	D. Laskin	2010
147	Atmospheric Transformation of Diesel Emissions	B. Zielinska	2010
145	Effects of Concentrated Ambient Particles and Diesel Engine Exhaust on Allergic Airway Disease in Brown Norway Rats	J. Harkema	2009
138	Health Effects of Real-World Exposure to Diesel Exhaust in Persons with Asthma	J. Zhang	2009

Table continues next page

APPENDIX A (*continued*). Related HEI Publications on Diesel Exhaust

Number	Title	Principal Investigator	Date
Research Reports (*continued*)			
134	Black-Pigmented Material in Airway Macrophages from Healthy Children: Association with Lung Function and Modeled PM_{10}	J. Grigg	2008
129	Particle Size and Composition Related to Adverse Health Effects in Aged, Sensitive Rats	F.F. Hahn	2005
128	Neurogenic Responses in Rat Lungs After Nose-Only Exposure to Diesel Exhaust	M. Witten	2005
124	Particulate Air Pollution and Nonfatal Cardiac Events		2005
	Part I. Air Pollution, Personal Activities, and Onset of Myocardial Infarction in a Case–Crossover Study	A. Peters	
	Part II. Association of Air Pollution with Confirmed Arrhythmias Recorded by Implanted Defibrillators	D. Dockery	
126	Effects of Exposure to Ultrafine Carbon Particles in Healthy Subjects and Subjects with Asthma	M.W. Frampton	2004
118	Controlled Exposures of Healthy and Asthmatic Volunteers to Concentrated Ambient Particles in Metropolitan Los Angeles	H. Gong Jr.	2003
112	Health Effects of Acute Exposure to Air Pollution	S.T. Holgate	2003
110	Particle Characteristics Responsible for Effects on Human Lung Epithelial Cells	A.E. Aust	2002
107	Emissions from Diesel and Gasoline Engines Measured in Highway Tunnels	A. Gertler	2002
76	Characterization of Fuel and After-Treatment Device Effects on Diesel Emissions	S. Bagley	1996
68-II	Pulmonary Toxicity of Inhaled Diesel Exhaust and Carbon Black in Chronically Exposed Rats. Part II: DNA Damage	K. Randerath	1995
72	DNA Adduct Formation and T-Lymphocyte Mutation Induction in F344 Rats Implanted with Tumorigenic Doses of 1,6-Dinitropyrene	F. Beland	1995
68-III	Pulmonary Toxicity of Inhaled Diesel Exhaust and Carbon Black in Chronically Exposed Rats. Part III: Examination of Possible Target Genes	S. Belinsky	1995
64	Biomonitoring of Nitropolynuclear Aromatic Hydrocarbons via Protein and DNA Adducts	K. El-Bayoumy	1994
68-I	Pulmonary Toxicity of Inhaled Diesel Exhaust and Carbon Black in Chronically Exposed Rats. Part I: Neoplastic and Nonneoplastic Lung Lesions	J. Mauderly	1994
66	The Effects of Copollutants on the Metabolism and DNA Binding of Carcinogens	P. Howard	1994
61	Methods Development Toward the Measurement of Polyaromatic Hydrocarbon–DNA Adducts by Mass Spectrometry	R. Giese	1993
56	Characterization of Particle- and Vapor-Phase Organic Fraction Emissions of a Heavy-Duty Diesel Engine Equipped with a Particle Trap and Regeneration Controls	S. Bagley	1993
55	Mutations Induced by 1-Nitrosopyrene and Related Compounds During DNA Recombination by These Compounds	V. Maher	1993
45	The Effects of Exercise on Dose and Dose Distribution of Inhaled Automotive Pollutants	M. Kleinman	1991
46	Role of Ring Oxidation in the Metabolic Activation of 1-Nitropyrene	F. Beland	1991
40	Retention Modeling of Diesel Exhaust Particles in Rats and Humans	C.P. Yu	1991

Table continues next page

APPENDIX A (*continued*). Related HEI Publications on Diesel Exhaust

Number	Title	Principal Investigator	Date
Research Reports (*continued*)			
33	Markers of Exposure to Diesel Exhaust in Railroad Workers	M. Schenker	1990
32	Respiratory Carcinogenesis of Nitroaromatics	R. Moon	1990
37	Oxidant Effects on Rat and Human Lung Proteinase Inhibitors	D. Johnson	1990
34	Metabolic Activation of Nitropyrenes and Diesel Particulate Extracts	A. Jeffrey	1990
26	Investigation of a Potential Cotumorigenic Effect of the Dioxides of Nitrogen and Sulfur, and of Diesel-Engine Exhaust, on the Respiratory Tract of Syrian Golden Hamsters	U. Heinrich	1989
31	DNA Binding by 1-Nitropyrene and Dinitropyrenes in Vitro and in Vivo: Effects of Nitroreductase Induction	F. Beland	1989
30	Influence of Experimental Pulmonary Emphysema on Toxicological Effects from Inhaled Nitrogen Dioxide and Diesel Exhaust	J. Mauderly	1989
16	Metabolism and Biological Effects of Nitropyrene and Related Compounds	C. King	1988
19	Factors Affecting Possible Carcinogenicity of Inhaled Nitropyrene Aerosols	R. Wolff	1988
17	Studies on the Metabolism and Biological Effects of Nitropyrene and Related Nitro-polycyclic Aromatic Compounds in Diploid Human Fibroblasts	V. Maher	1988
5	An Investigation into the Effect of a Ceramic Particle Trap on the Chemical Mutagens in Diesel Exhaust	S. Bagley	1987
8	Effects of Inhaled Nitrogen Dioxide and Diesel Exhaust on Developing Lung	J. Mauderly	1987
10	Predictive Models for Disposition of Inhaled Diesel Exhaust Particles in Humans and Laboratory Species	C.P. Yu	1987
7	DNA Adducts of Nitropyrene Detected by Specific Antibodies	J. Groopman	1987
4	The Metabolic Activation and DNA Adducts of Dinitropyrenes	F. Beland	1986
2	Disposition and Metabolism of Free and Particle-Associated Nitropyrenes After Inhalation	J. Bond	1986
Special Reports			
17	Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects		2010
	Research Directions to Improve Estimates of Human Exposure and Risk from Diesel Exhaust	Health Effects Institute	2002
	Diesel Emissions and Lung Cancer: Epidemiology and Quantitative Risk Assessment	Health Effects Institute	1999
	Diesel Exhaust: Critical Analysis of Emissions, Exposure, and Health Effects	Health Effects Institute	1995
	Potential Health Effects of Manganese in Emissions from Trap-Equipped Diesel Vehicles	Health Effects Institute	1988
HEI Perspectives			
	Understanding the Health Effects of Components of the Particulate Matter Mix: Progress and Next Steps		2002
HEI Program Summaries			
	Research on Diesel Exhaust and Other Particles		2003

APPENDIX B: DIESEL EPIDEMIOLOGY EXPERT WORKSHOP AGENDA, SPEAKERS, ATTENDEE LIST

The Health Effects Institute

Final Workshop Agenda

Diesel Exhaust, Lung Cancer and Quantitative Risk Assessment

Terrace Room, Park Plaza Hotel, Boston MA

Thursday, March 6, 2014

7:00 AM	**Registration and Continental Breakfast (until 8:15)**	Terrace Room
8:30 AM	Welcome, Project History, Introductions	Greenbaum
8:40 AM	Panel Charge, Workshop Objectives and Ground Rules	Krewski
8:55 AM	**Introduction to The Diesel Exhaust in Miners Study**	Silverman
9:05 AM	**Overview of the Exposure Assessment Process**	Stewart
9:20 AM	**Estimating Historical Exposures to Diesel Exhaust in Underground Non-metal Mines**	Vermeulen
9:40 AM	**A Cohort Study with Emphasis on Lung Cancer**	Attfield
10:00 AM	**A Nested Case-Control Study of Lung Cancer and Diesel Exhaust**	Silverman
10:20 AM	BREAK	
10:35 AM	Questions from the HEI Diesel Epidemiology Panel	Panel
11:15 AM	Input to the Panel from Audience	Audience
11:55 AM	**Diesel Trucker Study Presentation**	Garshick
12:20 PM	Questions from the HEI Diesel Epidemiology Panel	Panel
12:35-1:30 PM	LUNCH	

Afternoon Program:

1:35 PM	**Meta-exposure-response-modeling of diesel engine exhaust and lung cancer mortality**	Vermeulen
1:50 PM	Questions from the HEI Diesel Epidemiology Panel	
2:05 PM	Input to the Panel from Audience on both Truckers and Vermeulen study	
2:30 PM	**Reanalyses of DEMS exposure estimates**	Crump
2:40 PM	**Reanalyses of DEMS cohort data**	Moolgavkar
3:05 PM	**Reanalyses of DEMS case-control data**	Boffetta
3:15 PM	Questions from the HEI Diesel Epidemiology Panel	Panel
3:55 PM	BREAK	
4:10 PM	Input to the Panel from Audience	Audience
4:50 PM	**US EPA Integrated Risk Information System (IRIS): Use of Epidemiologic Evidence**	Cogliano
5:05 PM	**Diesel Engine Emissions And Risk Assessment At NIOSH**	Park
5:20 PM	**Science, Uncertainty, and Decision-making: Lessons from the National Academy of Sciences**	Rodricks
5:35 PM	Questions	Panel and Audience
5:50 PM	Next steps	Walker
6:00 PM	Adjourn	

Speakers and Affiliations

Michael Attfield
National Institute for Occupational Health and Safety (retired)

Paolo Boffetta*
Institute for Translational Epidemiology, Mount Sinai School of Medicine

Vincent Cogliano
Director of the Integrated Risk Information System, United States Environmental Protection Agency

Kenny Crump*
Private Consultant

Eric Garshick
VA Boston Healthcare System, Associate Professor of Medicine, Harvard Medical School, Channing Division of Network Medicine, Brigham and Women's Hospital

Daniel Greenbaum
President, Health Effects Institute

Daniel Krewski
Chair, HEI Diesel Epidemiology Panel; Director, McLaughlin Centre for Population Health Risk Assessment, University of Ottawa

Suresh Moolgavkar*
Principal Scientist, Director of the Center for Epidemiology, Biostatistics, and Computational Biology, Exponent Inc.

Robert Park
Risk Evaluation Branch, Education and Information Division, National Institute for Occupational Safety and Health

Joseph Rodricks
Principal, Environ International Corporation; Co-author, NAS reports, *Science and Decisions* (2009); *Environmental Decisions in the Face of Uncertainty* (2013)

Debra Silverman
Chief, Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute

Patricia Stewart
Stewart Exposure Assessments, LLC, National Cancer Institute (retired), Division of Cancer Epidemiology and Genetics

Roel Vermeulen
Associate Professor of Molecular Epidemiology and Risk Assessment, Institute for Risk Assessment Sciences (IRAS), Utrecht University

Katherine Walker
Project Manager, Health Effects Institute

* The work of these investigators has been coordinated by the Truck and Engine Manufacturers Association (EMA) on behalf of the American Petroleum Institute (API), European Automobile Manufacturers Association (ACEA), American Trucking Association (ATA), International Organization of Motor Vehicle Manufacturers (OICA), Alliance of Automobile Manufacturers (Alliance), European Research Group on Environment and Health in the Transport Sector (EUGT), Association of Equipment Manufacturers (AEM), Association of American Railroads (AAR), and European Association of Internal Combustion Engine Manufacturers (EUROMOT).

Attendee List

First Name	Last Name	Company Name
Kate	Adams	Health Effects Institute
George	Allen	NE States for Coordinated Air Use Management
Michael	Attfield	National Inst. for Occupational Safety and Health
Chad	Bailey	U.S. Environmental Protection Agency
Steve	Berry	Volvo Group Trucks
Traci	Bethea	Slone Epidemiology Center
Paolo	Boffetta	Mount Sinai School of Medicine
Hanna	Boogaard	Health Effects Institute
Jonathan	Borak	Yale University
William	Bunn	Independent Consultant
Igor	Burstyn	Drexel University
Joel	Carr	Unifor
Bob	Carreau	IAMGOLD Corporation
Henry	Chajet	Jackson Lewis LLP
Farah	Chowdhury	ENVIRON International Corporation
Vincent	Cogliano	U.S. Environmental Protection Agency
Aaron	Cohen	Health Effects Institute
Susan	Collet	Toyota Motor Engineering & Manufacturing, NA
Bruce	Copley	ExxonMobil Biomedical Sciences Inc.
Daniel	Costa	U.S. Environmental Protection Agency
Maria	Costantini	Health Effects Institute
Edmund	Crouch	Camp Dresser & McKee Smith
Kenny	Crump	ICF Kaiser
David	Damico	Burns White LLC
Paul	Demers	Cancer Care Ontario
Marika	Egyed	Health Canada
Ellen	Eisen	University of California–Berkeley
Heidi	Erickson	Chevron Energy Technology Company
Howard	Feldman	American Petroleum Institute
David	Foster	University of Wisconsin–Madison
Tim	French	Engine Manufacturers Association
Eric	Garshick	VA Boston Health Care System, Channing Lab
Edward	Green	Crowell & Moring LLP
Laura	Green	Camp Dresser & McKee Smith
Dan	Greenbaum	Health Effects Institute
Paul	Greening	Association des Constructeurs Europeens d'Auto
Jaime	Hart	Harvard University
Kevin	Hedges	Independent Researcher
Uwe	Heinrich	Fraunhofer-Institut für Toxikologie und Aerosolfor
Carol	Henry	George Washington University
Thomas	Hesterberg	Center for Toxicology & Environmental Health

Table continues next page

Attendee List (*continued*)

First Name	Last Name	Company Name
Paul	Hewett	Exposure Assessment Solutions Inc.
Tom	Jayne	BNSF Railway Company
Stuart	Johnson	Volkswagen of America Inc.
Debra	Kaden	ENVIRON International Corporation
Farah	Kassam	GoldCorp Inc.
Joel	Kaufman	University of Washington
Joanne	Kim	Cancer Care Ontario
Takahiro	Koseki	Isuzu Manufacturing Services of America Inc.
Daniel	Krewski	University of Ottawa
Svitlana	Kroll	Southwest Research Institute
Francine	Laden	Brigham & Women's Hospital
Bill	Lamson	Lamson, Dugan & Murray LLP
Jonathan	Levy	Boston University
Chris	Long	Gradient Corporation
Roger	McClellan	Independent Consultant
Amy	McCool	OCI Enterprises Inc.
Jeffrey	Moninger	Mine Safety & Health Administration
Suresh	Moolgavkar	Exponent
David	Morgott	Pennsport Consulting LLC
Nick	Moustakas	Health Effects Institute
Robert	O'Keefe	Health Effects Institute
Robert	Park	National Institute for Occupational Safety and Health
Mel	Peffers	U.S. Environmental Protection Agency
Hilary	Polk	Health Effects Institute
Charles	Poole	University of North Carolina–Chapel Hill
Lutzen	Portengen	University of Utrecht
Jacqueline	Presedo	Health Effects Institute
Nancy	Reid	University of Toronto
Kevin	Reiss	John Deere Product Engineering Center
Reginald	Richards	U.S. Department of Labor, Mine Safety and Health Admin
Charles	Ris	U.S. Environmental Protection Agency
Joseph	Rodricks	ENVIRON Corporation
Arlean	Rohde	CONCAWE
Craig	Rood	OCI Wyoming LP
Michael	Rush	Association of American Railroads
Jason	Sacks	U.S. Environmental Protection Agency
Daniel	Saphire	Association of American Railroads
Joseph	Sawin	Cummins Inc.
Allen	Schaeffer	Diesel Technology Forum
Patricia	Schleiff	Centers for Disease Control and Prevention
Melissa	Seaton	National Institute for Occupational Safety and Health

Table continues next page

Attendee List (*continued*)

First Name	Last Name	Company Name
Rashid	Shaikh	Health Effects Institute
Lianne	Sheppard	University of Washington–Seattle
Yuanli	Shi	University of Ottawa
Debra	Silverman	National Cancer Institute
Nakia	Simon	Chrysler LLC
Thomas	Smith	Harvard T.H. Chan School of Public Health
Christine	Sofge	National Institute for Occupational Safety and Health
Jamie	Song	Manufacturers of Emission Controls Association
Michael	Spallek	EUGT e.v.
Kyle	Steenland	Emory University
Patricia	Stewart	Stewart Exposure Assessments LLC
Geoffrey	Sunshine	Health Effects Institute
Deborah	Tomko	Mine Safety & Health Administration
Peter	Valberg	Gradient Corporation
Annemoon	van Erp	Health Effects Institute
Martie	van Tongeren	Institute of Occupational Medicine
Roel	Vermuelen	University of Utrecht
Gregory	Wagner	Centers for Disease Control and Prevention/NIOSH
Rich	Wagner	Cummins Inc.
Katy	Walker	Health Effects Institute
Tim	Wallington	Ford Motor Company
Chris	Walters	CNH Industrial
Judy	Wendt-Hess	Shell Oil Company, Shell Health — Americas
Matthew	Winings	Cummins Inc.
George	Wolff	Air Improvement Resource Inc.
Nagarajkumar	Yenugadhati	University of Ottawa
Lu	Yu	Phillips 66
Wig	Zamore	Somerville Transportation for Equity Partnership

APPENDIX C: SUMMARY OF OTHER OCCUPATIONAL EXPOSURES IN THE MINES

Table C.1. Number of Area and Personal DE-Related Measurements by Agent for the Eight Mining Facilities

	Survey[a]												
	MIDAS 1976–2001		DEMS 1998–2001		MESA/BoM 1976–1977		Feasibility Study 1994		Other 1954–1996		All Surveys		
Agent[b]	Area[c]	Personal[c]	Area	Personal	Area	Personal	Area	Personal	Area	Personal	Area	Personal	Total
CO	9,746	46	208	0	1,099	0	25	0	46	0	11,124	46	11,170
CO_2	8,234	15	390	0	961	0	17	0	49	0	9,651	15	9,666
NO	45	0	381	995	24	0	42	69	9	0	501	1,064	1,565
NO_2	4,288	38	387	1,031	252	646	42	69	76	11	5,045	1,795	6,840
TD	1	782	215	0	161	667	32	0	69	703	478	2,152	2,630
RD	0	324	209	2	99	0	31	0	158	178	497	504	1,001
SD	0	0	121	0	0	0	69	0	20	0	210	0	210
TEC	0	0	224	0	0	0	0	0	0	0	224	0	224
REC	0	0	216	1,156	0	0	0	69	12	4	228	1,229	1,457
SEC	0	0	209	0	0	0	0	0	0	0	209	0	209
TOC	0	0	224	0	0	0	0	0	0	0	224	0	224
ROC	0	0	221	1,151	0	0	0	0	0	0	221	1,151	1,372
SOC	0	0	207	0	0	0	0	0	0	0	207	0	207
DPM/ SCD	0	0	212	0	0	0	0	0	180	102	392	102	494
Total	22,314	1,205	3,424	4,335	2,596	1,313	258	207	619	998	29,211	8,058	37,269

[a] Surveys: the MSHA MIDAS (1976–2001); the DEMS (1991–2001) (Coble et al. 2010; Vermeulen et al. 2010b); the MESA/BoM (1976–1977) (Sutton et al. 1979); the feasibility study for the DEMS in Facility B (1994) (Stanevich et al. 1997); compliance visits by the State of New Mexico, MSHA hard copy reports, and the mining facilities (1954–1996).

[b] DPM, diesel particulate matter; RD, respirable dust; ROC, respirable organic carbon; SCD, submicron combustible dust; SD, submicron dust; SEC, Submicron elemental carbon; SOC, submicron organic carbon; TD, total dust; TEC, total elemental carbon; TOC, total organic carbon.

[c] Area measurements; personal measurements. The number includes both full-shift and short-term measurements.

Reprinted from Stewart et al. 2010, Table 2 by permission of Oxford University Press.

* A list of abbreviations and other terms appears at the end of the Special Report.

Table C.2 (*continues across 2-page spread*). Mean and 95% Confidence Interval of Exposures to Respirable Elemental Carbon, Silica, Asbestos, Non-Diesel Poly-aromatic Hydrocarbons, Radon, and Respirable Dust by Facility, by Worker Location Within Facility and Over All Facilities[a]

Variable / Worker Subgroup[b]	Facility Exposures, Mean (95% Cl)			
	Limestone	Potash		
	A	B	D	J
REC (µg/m³)				
Complete cohort	45.3 (41.4–49.2)	181.3 (172.1–190.5)	92.9 (87.9–97.9)	96.3 (92.0–100.7)
Ever-underground[c]	78.1 (72.0–84.2)	216.1 (207.0–225.2)	150.2 (143.8–156.6)	122.7 (118.1–127.2)
Surface-only[d]	2.5 (2.4–2.5)	2.0 (1.9–2.1)	0.9 (0.8–0.9)	1.0 (0.9–1.0)
Silica[e]				
Complete cohort	0.01 (0.00–0.01)	0.96 (0.94–0.97)	0.88 (0.86–0.89)	0.88 (0.87–0.90)
Ever-underground	0.004 (0.001–0.006)	0.98 (0.97–0.99)	0.94 (0.93–0.95)	0.95 (0.95–0.96)
Surface-only	0.005 (0.003–0.008)	0.85 (0.82–0.89)	0.73 (0.70–0.75)	0.68 (0.65–0.71)
Asbestos[e]				
Complete cohort	0.19 (0.17–0.21)	0.31 (0.29–0.34)	0.23 (0.21–0.25)	0.14 (0.13–0.16)
Ever-underground	0.16 (0.14–0.18)	0.34 (0.31–0.37)	0.19 (0.17–0.21)	0.14 (0.13–0.16)
Surface-only	0.16 (0.13–0.18)	0.13 (0.09–0.16)	0.25 (0.23–0.28)	0.10 (0.08–0.12)
Non-DE PAHs[f]				
Complete cohort	0.15 (0.14–0.16)	0.31 (0.29–0.34)	0.23 (0.22–0.25)	0.17 (0.16–0.19)
Ever-underground	0.14 (0.12–0.16)	0.34 (0.31–0.37)	0.19 (0.17–0.21)	0.17 (0.15–0.18)
Surface-only	0.10 (0.09–0.12)	0.13 (0.09–0.16)	0.25 (0.23–0.28)	0.13 (0.11–0.16)
Radon, WL[g]				
Complete cohort	0.005 (0.005–0.006)	0.014 (0.014–0.015)	0.010 (0.010–0.010)	0.007 (0.007–0.007)
Ever-underground	0.009 (0.009–0.010)	0.017 (0.016–0.017)	0.016 (0.016–0.017)	0.009 (0.009–0.009)
Surface-only	0	0	0	0
Respirable Dust (mg/m³)				
Complete cohort	0.89 (0.87–0.91)	2.66 (2.58–2.74)	1.29 (1.26–1.31)	2.63 (2.56–2.70)
Ever-underground	1.02 (0.99–1.05)	3.06 (2.99–3.12)	1.65 (1.62–1.67)	3.16 (3.09–3.22)
Surface-only	0.71 (0.70–0.73)	0.65 (0.64–0.67)	0.76 (0.74–0.78)	0.75 (0.73–0.77)

[a] Facilities coded according to industrial hygiene reports (22-25). Jobs involving work in both surface and underground locations were prorated by fraction of time spent underground in years.

[b] CI = confidence intervals; non-DE PAHs = non-diesel poly-aromatic hydrocarbons; REC = respirable elemental carbon.

[c] Workers categorized as ever-underground after first going underground (even if surface later).

[d] Workers categorized as surface only until first going underground (if ever).

[e] Semiquantitative exposure categories coded on a relative scale (0, 1, and 2).

[f] Non-DE PAHs categorized as present or absent (0 and 1).

[g] The concentration of radon daughters is measured in units of WL, which is a measure of the potential alpha particle energy per liter of air. One WL of radon daughters corresponds to approximately 200 pCi/L of radon in a typical indoor environment.

Source: Reprinted from Attfield et al. 2012 by permission of Oxford University Press.

Table C.2 (*continued across 2-page spread*). Mean and 95% Confidence Interval of Exposures to Respirable Elemental Carbon, Silica, Asbestos, Non-Diesel Poly-aromatic Hydrocarbons, Radon, and Respirable Dust by Facility, by Worker Location Within Facility and Over All Facilities[a]

Variable / Worker Subgroup[b]	Facility Exposures, Mean (95% Cl)				
	Salt (halite)	Trona			All
	E	G	H	I	
REC (µg/m³)					
Complete cohort	155.2 (145.8–164.5)	79.3 (73.5–85.1)	78.4 (75.4–81.4)	65 (62.2–67.8)	87.0 (85.2–88.8)
Ever-underground[c]	170.5 (161.2–179.8)	152.1 (144.7–159.4)	105.6 (102.6–108.6)	100.7 (97.5–104.0)	128.2 (126.1–130.3)
Surface-only[d]	3.2 (3.1–3.3)	2.1 (2.1–2.1)	1.6 (1.5–1.6)	1.1 (1.1–1.2)	1.7 (1.6–1.7)
Silica[e]					
Complete cohort	0.003 (0.00–0.01)	1.72 (1.69–1.76)	1.76 (1.74–1.78)	1.73 (1.71–1.76)	1.11 (1.10–1.13)
Ever-underground	0.003 (0.00–0.008)	1.91 (1.88–1.93)	1.84 (1.82–1.86)	1.87 (1.85–1.89)	1.18 (1.17–1.20)
Surface-only	0	1.58 (1.53–1.63)	1.55 (1.50–1.59)	1.48 (1.43–1.52)	0.88 (0.85–0.90)
Asbestos[e]					
Complete cohort	0.16 (0.13–0.19)	0.30 (0.28–0.33)	0.27 (0.25–0.28)	0.29 (0.27–0.30)	0.24 (0.23–0.25)
Ever-underground	0.17 (0.14–0.21)	0.25 (0.21–0.28)	0.23 (0.21–0.25)	0.27 (0.25–0.29)	0.22 (0.21–0.23)
Surface-only	0.06 (0.03–0.09)	0.27 (0.24–0.31)	0.31 (0.28–0.35)	0.29 (0.26–0.32)	0.22 (0.21–0.23)
Non-DE PAHs[f]					
Complete cohort	0.16 (0.13–0.19)	0.26 (0.24–0.28)	0.24 (0.23–0.26)	0.27 (0.25–0.28)	0.23 (0.22–0.23)
Ever-underground	0.17 (0.14–0.21)	0.23 (0.20–0.27)	0.22 (0.20–0.23)	0.26 (0.24–0.28)	0.21 (0.21–0.22)
Surface-only	0.06 (0.03–0.09)	0.21 (0.19–0.24)	0.27 (0.24–0.30)	0.26 (0.23–0.28)	0.19 (0.18–0.20)
Radon, WL[g]					
Complete cohort	0.014 (0.014–0.015)	0.009 (0.008–0.009)	0.006 (0.005–0.006)	0.005 (0.005–0.005)	0.008 (0.008–0.008)
Ever-underground	0.016 (0.015–0.016)	0.017 (0.016–0.017)	0.008 (0.008–0.008)	0.008 (0.008–0.008)	0.011 (0.011–0.012)
Surface-only	0	0	0	0	0
Respirable Dust (mg/m³)					
Complete cohort	1.06 (1.03–1.09)	1.55 (1.49–1.61)	1.53 (1.50–1.56)	1.07 (1.04–1.09)	1.51 (1.50–1.53)
Ever-underground	1.10 (1.08–1.12)	2.34 (2.27–2.40)	1.84 (1.81–1.87)	1.42 (1.40–1.44)	1.93 (1.91–1.95)
Surface-only	0.53 (0.48–0.58)	0.77 (0.75–0.79)	0.68 (0.67–0.70)	0.45 (0.44–0.46)	0.67 (0.67–0.68)

[a] Facilities coded according to industrial hygiene reports (22-25). Jobs involving work in both surface and underground locations were prorated by fraction of time spent underground in years.

[b] CI = confidence intervals; non-DE PAHs = non-diesel poly-aromatic hydrocarbons; REC = respirable elemental carbon.

[c] Workers categorized as ever-underground after first going underground (even if surface later).

[d] Workers categorized as surface only until first going underground (if ever).

[e] Semiquantitative exposure categories coded on a relative scale (0, 1, and 2).

[f] Non-DE PAHs categorized as present or absent (0 and 1).

[g] The concentration of radon daughters is measured in units of WL, which is a measure of the potential alpha particle energy per liter of air. One WL of radon daughters corresponds to approximately 200 pCi/L of radon in a typical indoor environment.

Source: Reprinted from Attfield et al. 2012 by permission of Oxford University Press.

APPENDIX D. PANEL ANALYSES: ALTERNATIVE CONTROL FOR SMOKING IN THE CASE–CONTROL STUDY

CONTENTS

1. INTRODUCTION

Smoking is the major cause of lung cancer (U.S. Department of Health and Human Services 2014). In the Diesel Exhaust in Miners Study (DEMS*) nested case–control study individual smoking information was considered as a confounding variable that constituted smoking status, intensity of smoking (packs/day) and location of employment (ever-underground or surface-only employment) as a single combination variable. The objective of the present supplemental smoking analysis was to evaluate the influence of different smoking parameters (i.e., including smoking status, duration, packs/day, and packyears) in various combinations, on the relationship between diesel exhaust and lung cancer mortality. A detailed description of various smoking variables and different analyses performed are provided in the Methodology section of this appendix. All of the results are either tabulated in the Results section or provided in Additional Materials 3 (available on the HEI Web site). Finally, some key findings are noted in the Summary section.

2. METHODOLOGY

All analyses conducted here were based on the DEMS raw study data provided by the National Cancer Institute. In this section, smoking variables used by the original investigators are described, along with additional smoking variables derived from the DEMS raw study data used in the HEI Panel's analyses. The methods used in various descriptive analyses as well as conditional logistic regression analyses are also described. All statistical analyses were performed using SAS version 9.4 (SAS Institute. Cary, NC).

* A list of abbreviations and other terms appears at the end of the Special Report.

2.1 Smoking Variables

2.1.1 Original Smoking Variable In the original analyses published by the DEMS investigators (Silverman et al. 2012), smoking was characterized by means of a combination variable that consisted of smoking status, packs/day smoked, and location of employment as a single categorical variable. The 16 categories of this combination smoking variable were:

- surface work only; never smoker;
- surface work only; unknown; occasional smoker;
- surface work only; former smoker; < 1 pack/day;
- surface work only; former smoker; 1 to < 2 packs/day;
- surface work only; former smoker; ≥ 2 packs/day;
- surface work only; current smoker; < 1 pack/day;
- surface work only; current smoker; 1 to < 2 packs/day;
- surface work only; current smoker; ≥ 2 packs/day;
- ever underground work; never smoker;
- ever underground work; unknown; occasional smoker;
- ever underground work; former smoker; < 1 pack/day;
- ever underground work; former smoker; 1 to < 2 packs/day;
- ever underground work; former smoker; ≥ 2 packs/day;
- ever underground work; current smoker; < 1 pack/day;
- ever underground work; current smoker; 1 to < 2 packs/day; and
- ever underground work; current smoker; ≥ 2 packs/day.

2.1.2 New Smoking Variables The following indicators of smoking derived from the raw DEMS data were considered for use in the present analyses. Combination variables were also used in the analyses to avoid over-parameterization.

a. *Smoking status* (never, former, current, unknown): Age at quitting smoking was used to determine smoking status. If a subject (case or matched control) quit smoking two years prior to the date of death of the case, then subjects who had ever smoked were defined

as former smokers; otherwise, the subjects were designated as current smokers. Occasional smokers were included in the 'unknown' category in all analyses. Status is always categorical.

b. *Smoking intensity* (packs/day): Packs/day were calculated based on a pack size comprised of 20 cigarettes. When only a range of packs smoked was available (this occurred for 51 subjects), the midpoint of the range limits was assigned as the packs/day. (For example, subjects smoking 1–2 packs/day were assigned 1.5 packs/day as their smoking intensity.) Packs/day was used both as a continuous and as a categorical variable (never smokers; < 1 pack/day; 1 to < 2 packs/day; 2 or more packs/day; unknown).

c. *Smoking duration*: Smoking duration was derived based on the difference between age started smoking and age quit smoking. Smoking duration for current smokers was lagged by 2 years prior to the reference date (the date of death/interview for the case and matched controls). Since information on age started smoking was not available for more than a quarter of the subjects, the smoking duration variable was subject to the same degree of *missingness* as the age started smoking variable. Smoking duration was used as both a continuous and a categorical variable (never smokers; smoked 2 to 39 years; smoked for 40 or more years; unknown); tertiles were used because quartile categories were too sparse, resulting in some cells with fewer than two subjects.

d. *Packyears*: The packyears variable was defined as the product of duration and packs/day. (About a quarter of all subjects were missing this information as well.) The packyears variable was used both as a continuous and a categorical variable (never smokers; 0.5 to 26 packyears; 27 to 54 packyears; 55 or more packyears; unknown).

e. *Smoking status and duration as a combined categorical variable* (*status–duration*): For the purpose of the present analyses, a combination variable was created based on smoking status (never/former/current smokers) and tertiles of duration. This variable was referred as status–duration (never smokers; former smokers 2 to 39 years; former smokers ≥ 40 years; current smokers 2 to 39 years; current smokers ≥ 40 years; unknown).

f. *Smoking status and packs/day as a combined categorical variable* (*status–packs/day*): A combination variable was created based on smoking status and packs/day as a categorical variable. This variable was referred as status–packs/day (nonsmokers; former smokers < 1 pack/day; former smokers 1 to < 2 packs/day; former smokers ≥ 2 packs/day; current smokers < 1 pack/day; current smokers 1 to < 2 packs/day; current smokers ≥ 2 packs/day; unknown).

g. *Smoking status and packyears as a combined categorical variable* (*status–packyears*): A combination variable was created based on smoking status and categorical packyears. This variable was referred as status–packyears (never smokers; former smokers 0.5 to 26 packyears; former smokers 27 to 54 packyears; former smokers ≥ 55 packyears; current smokers 0.5 to 26 packyears; current smokers 27 to 54 packyears; current smokers ≥ 55 packyears; unknown).

h. *Proxy status* (*self/proxy*): This variable indicates whether smoking information was reported by the study participant or by a proxy respondent. (Smoking information for all of the cases was necessarily reported by proxy respondents.)

i. *Duration of smoking cessation*: The duration of smoking cessation was not used in current analysis due to the large amount of missing information on age started smoking (this information was missing for more than a fourth of the study participants); data on intermittent smoking cessation was available for very few subjects.

2.2 New Analyses Performed

The new smoking variables described above were used in the present analyses, and the results were compared with those reported by the original investigators.

2.2.1 Descriptive Statistics Descriptive statistics for all smoking variables were calculated by case and control status, proxy status (whether the questionnaire was answered by proxies, such as spouse, friends and relatives, or self-reported by the subjects), and location of employment in the mine (ever-underground; surface-only).

In addition, descriptive statistics were also calculated by both location and case–control status (surface-only cases; surface-only controls; ever-underground cases; ever-underground controls).

2.2.2 Conditional Logistic Regression Analyses Conditional logistic regression analyses were performed to evaluate the relationship between diesel exhaust exposure and lung cancer mortality, with exposure to diesel exhaust measured in terms of the concentration of respirable elemental carbon (REC), using the full suite of smoking variables described. The models used in the present analyses, compared with those of the original investigators, are summarized in the matrix provided in Table D.1.

2.2.2.1 Analytical Variables

a. *REC*: The main exposure variable was REC. Separate analyses were performed for various REC metrics, including cumulative REC, average REC, cumulative REC lagged 15 years from the reference date, and average REC lagged 15 years from the reference date.

b. *Common covariates*: The common covariates used in all the models were 'history of respiratory disease 5 or more years before date of death/reference date', and 'history of a high-risk job for lung cancer for at least 10 years'. In what follows, these common covariates were denoted simply as 'covariates'.

c. *Smoking variable*: The analyses differed with respect to how the smoking variables were used to adjust for the possible confounding effect of smoking on the relationship between REC and lung cancer mortality. Specifically, the following smoking variables were used in the analyses.
   - Original investigators' smoking combination variable with location.
   - Status–duration as a categorical variable.
   - Status–packs as a categorical variable.
   - Status–packyears as a categorical variable.
   - Smoking status as a categorical variable.
   - Smoking duration as a continuous and as a categorical variable.
   - Packs/day as a continuous and as a categorical variable.
   - Packyears as a continuous and as a categorical variable.

d. *A location variable* (ever-underground or surface-only) was used as a separate categorical variable in all HEI analyses to adjust for its possible confounding effect, unlike the original investigators' model that incorporated location as a combination variable with smoking status and intensity.

2.2.2.2 Analytical Models Various models were assessed based on the nature of the smoking variable (smoking status, duration, packs/day or packyears), the type of variable (continuous or categorical), or a combination of the above variables. These models were all of the general form:

> *Lung cancer mortality = REC metric + smoking variable + covariates + location,*

where REC was described in four ways: as average or cumulative REC (unlagged) or average or cumulative REC, lagged 15 years.

2.2.2.2.1 SILVERMAN ET AL. (2012) MODELS The original investigator models included smoking as a combination variable comprised of smoking status, packs/day and location (Silverman et al. 2012). These models are of the general form:

> *Lung cancer mortality = REC metric + smoking combination variable + covariates.*

2.2.2.2.2 MODELS BASED ON NEW SMOKING DURATION VARIABLES

a. Using the combination variable 'status–duration' as a categorical variable:

   Lung cancer mortality = REC + status–duration (never smoker; former 2–39 yrs; former > 40 yrs; current 2–39 yrs; current > 40 yrs; unknown) + location (ever-underground; surface-only) + covariates.

b. Using smoking duration as a continuous variable along with smoking status as a categorical variable:

   Lung cancer mortality = REC + smoking status (never; former; current; unknown) + smoking duration continuous variable + location (ever-underground; surface-only) + covariates.

c. Secondary analyses based on smoking duration alone.

   As supplemental analyses based on smoking duration involved models excluding the smoking status variable, no combination variables were used.

   - Using duration only as a continuous variable:

     Lung cancer mortality = REC + smoking duration (continuous) + location (ever-underground; surface-only) + covariates.

   - Using duration only as a categorical variable:

     Lung cancer mortality = REC + smoking duration (never smokers; smoked 2 to 39 years; smoked for 40 or more years; unknown duration) + location (ever-underground; surface-only) + covariates.

2.2.2.2.3 MODELS BASED ON NEW SMOKING INTENSITY (PACKS/DAY) VARIABLES:

a. Using the combination variable 'status–packs/day' as a categorical variable:

   Lung cancer mortality = REC + status–packs/day (never smokers; former: < 1 pack/day; former 1 to < 2 packs/day; former 2 or more packs/day; current

< 1 pack/day; current 1 to < 2 packs/day; current 2 or more packs/day; unknown) + location (ever-underground; surface-only) + covariates.

b. Using smoking packs/day as a single continuous variable along with smoking status as a separate categorical variable:

Lung cancer mortality = REC + smoking status (never; former; current; unknown) + smoking packs/day (continuous) + location (ever-underground; surface-only) + covariates.

c. Secondary analyses based on packs/day alone (without smoking status):

- Using packs/day only as a continuous variable:

Lung cancer mortality = REC + smoking packs/day (continuous) + location (ever-underground; surface-only) + covariates.

- Using packs/day only as a categorical variable:

Lung cancer mortality = REC + smoking packs/day (never smokers; < 1 pack/day; 1 to < 2 packs/day; 2 or more packs/day; unknown) + location (ever-underground; surface-only) + covariates.

2.2.2.2.4 MODELS BASED ON NEW SMOKING PACKYEARS VARIABLES:

a. Using the combination variable, *status–packyears*, as a categorical variable:

Lung cancer mortality = REC + status–packyears (never smokers; former smoker 0.5 to 26 packyears; former smoker 27 to 54 packyears; former smoker 55 packyears or more; current smoker 0.5 to 26 packyears; current smoker 27 to 54 packyears; current smoker 55 packyears or more; unknown) + location (ever-underground; surface-only) + covariates.

b. Using smoking packyears as a continuous variable along with smoking status as a categorical variable:

Lung cancer mortality = REC + smoking status (never; former; current; unknown) + smoking packyears (continuous) + location (ever-underground; surface-only) + covariates.

c. Secondary analyses based on packyears alone:

- Using packyears only as a continuous variable:

Lung cancer mortality = REC + smoking packyears (continuous) + location (ever-underground; surface-only) + covariates.

- Using packyears only as a categorical variable:

Lung cancer mortality = REC + smoking packyears categorical variable (never smokers; 0.5 to 26 packyears; 27 to 54 packyears; 55 or more packyears; unknown) + location (ever-underground; surface-only) + covariates.

2.2.2.2.5 MODELS BASED ON SMOKING STATUS, DURATION, AND PACKS/DAY AS SEPARATE VARIABLES:

a. Using smoking status, smoking duration and packs/day, each as categorical variables:

Lung cancer mortality = REC + smoking status (never; former; current; unknown) + smoking packs/day (never smokers; < 1 packs/day; 1 to < 2 packs/day; 2 or more packs/day; unknown) + location (ever-underground; surface-only) + duration (never smokers; smoked 2 to 39 years; smoked for 40 or more years; unknown duration) + location (ever-underground; surface-only) + covariates.

b. Using smoking duration and packs/day as continuous variables, along with smoking status as a categorical variable:

Lung cancer mortality = REC + smoking status (never; former; current; unknown) + packs/day (continuous) + duration (continuous) + location (ever-underground; surface-only) + covariates.

2.2.2.3 Additional Secondary Analytical Models Excluding Unknowns

a. All analyses in section 2.2.2.2 were repeated by excluding the unknowns from the categorical smoking variables: this was done by excluding subjects with missing information on these variables. When using the original investigator (Silverman et al. 2012) combination variable (smoking status, packs/day and location), the unknowns in both the smoking status and packs/day variables were excluded.

b. In our supplemental secondary analyses based on smoking status (categorical) and packs/day as a continuous variable (see section 2.2.2.2.3, b) when unknowns were excluded, the unknowns in both smoking status and in packs/day were excluded; therefore these results differ from the analyses that included unknowns that only exclude the subjects

with unknown packs/day. However, such a difference was not seen in analyses with smoking duration and packyears, as excluding unknowns for these variables effectively excluded unknowns in smoking status variables.

2.2.2.4 Trend Test *P* values for tests for trend in lung cancer risk were obtained by using categorical REC as a continuous variable in each of the models noted above (in sections 2.2.2.2 and 2.2.2.3). The median values of each of the four categories of REC were used to represent those categories in these models. The trend test was based on the Wald statistic, with significant results determined by $P < 0.05$.

2.2.3 Interactions Interactions between location of employment (ever-underground/surface-only) and smoking variables (duration, packs/day and packyears) were tested for all the models in section 2.2.2, with the exception of the original investigators' models and the models with both duration and packs/day (see section 2.2.2.2.5). A multiplicative interaction term between the smoking variable included in the model and location was added to the models without the interaction term, and the influence on the relationship between REC and lung cancer mortality was noted. The interaction was deemed significant if the *P* value based on the Wald statistic was less than 0.05. It should be noted that the purpose of this analysis was not primarily to evaluate the interaction between location and smoking, but rather to ensure that the possibility of interaction between these two variables had been considered when estimating the risk of lung cancer associated with REC.

3. RESULTS

Results of the primary analyses described above (where the various smoking variables were included as categorical variables) are summarized in Table D.2 for each of the four REC metrics. The original analyses from Silverman and colleagues (2012) are provided for comparison. Results of secondary analyses (i.e., with individual smoking variables expressed as continuous variables, included alone without smoking status, or excluding subjects with unknown smoking characteristics) are provided in Additional Materials 3 (available on the HEI Web site). The descriptive statistics could not be presented due to confidentiality requirements specified in the Data Use Agreements with the National Cancer Institute and the National Institutes of Occupational Health.

3.1 Descriptive Statistics on Smoking

Key findings from the descriptive analyses are as follows:

a. Compared with controls, a higher percentage of cases were smokers, and the percentage of current smokers in different categories of intensity, including those that smoked more than one pack per day, was higher among cases.
b. Both duration of smoking and packyears were higher among cases than controls.
c. A higher percentage of surface workers were nonsmokers compared with underground workers; however, the percentage of subjects who quit smoking was higher in underground miners.
d. As noted above, all responses for the cases were reported by proxies.
e. Among the control subjects, proxy respondents reported a higher percentage of regular smoking compared to self reports. However, self-reported data indicated a higher percentage of subjects who had quit smoking.
f. The proxy respondents for cases classified a higher percentage of cases to be current smokers than did proxy respondents for controls.

3.2 Logistic Regression Models

Conditional logistic regression models were used to estimate the risk of lung cancer associated with exposure to diesel exhaust, adjusted for smoking. The odds ratios and 95% confidence intervals for lung cancer mortality for the main (primary) models (that included smoking status and either duration, packs/day, or packyears as combination variables) and a model in which smoking status, duration and packs/day are included as separate categorical variables are tabulated in Table D.2. (For completeness they are also found in Additional Materials Tables 3.2, 3.3, 3.4, and 3.5, respectively, along with the interaction results). The results of the corresponding secondary analyses using duration only, packs/day only, or packyears only as smoking parameters are tabulated in Additional Materials Tables 3.6, 3.7, and 3.8, respectively. A summary of the results of the analysis of interactions between various smoking variables and location of employment are reported in Additional Materials Table 3.9.

Some key findings of these analyses are summarized below.

a. The combination variables involving smoking were used to avoid over-parameterization due to the high correlation between individual categorical smoking variables, including smoking status, duration, intensity (packs/day) and packyears.
b. The association between REC and lung cancer mortality were consistent with those of Silverman and colleagues (2012) when packs/day was used as a smoking

parameter; however, the results were less consistent when duration and packyears were used, possibly due to missing information on smoking duration for over 25% of the subjects. In their later analyses (Silverman et al. 2014), the footnote to Table 2 in that publication indicates that "For the 72 cases and 107 controls with missing age started smoking information, age started was assumed to be 17 years old, which was the average age started in the controls."

c. Interactions between the smoking and location variables were not significant when the complete data set, not excluding subjects with missing smoking data, was used for analysis, using the category 'unknown' for missing smoking data. All of the significant interactions observed were noted with continuous smoking variables, where subjects with unknown smoking information were excluded from data analysis (see Additional Materials Table 3.9).

d. Caution should be exercised in interpreting the interaction results based on combination variables, owing to the small number of subjects (< 5) in some cells.

e. No interaction was observed between location of employment and smoking when a measure of REC was excluded from the logistic regression models.

4. SUMMARY

Overall, the results of the present analyses of the relationship between REC and lung cancer mortality are consistent with those of the original investigators (Silverman et al. 2012) when smoking intensity (packs/day) was used to characterize tobacco smoking. Although the results of the analyses using smoking duration and packyears differed somewhat from those of the original investigators, it should be noted that more than a quarter of the data was represented as unknown. Because of this high degree of missingness, both the duration and packyears variables are thus less well suited to properly adjust for the possible confounding effects of smoking. No significant interactions were observed between smoking variables and location, with the exception of some significant interactions in a secondary analysis that excluded unknown subjects.

5. REFERENCES

Silverman DT, Lubin JH, Blair AE, Vermeulen R, Stewart PA, Schleiff PL, et al. 2014. Re: The diesel exhaust in miners study (DEMS): a nested case-control study of lung cancer and diesel exhaust. J Natl Cancer Inst 106:ii.

Silverman DT, Samanic CM, Lubin JH, Blair AE, Stewart PA, Vermeulen R, et al. 2012. The diesel exhaust in miners study: a nested case-control study of lung cancer and diesel exhaust. J Natl Cancer Inst 104:(11):855–868.

U.S. Department of Health and Human Services. 2014. The Health Consequences of Smoking — 50 Years of Progress: A Report of the Surgeon General. Atlanta, GA:U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. Available: *www.surgeongeneral.gov/library/reports/50-years-of-progress/index.html*.

Table D.1. Summary of Models Used in the Analysis With Various Smoking Variables[a]

			Covariates[b]				Packs/Day		Duration of Smoking		Packyears		Combination Variables[c]		
Models	REC[d]	Location[e]	Resp. Disease	High-Risk Job	OI Comb. Smoking[f]	Smoking Status	Categorical	Continuous	Categorical	Continuous	Categorical	Continuous	Status-Packs/Day	Status-Duration	Status-Packyears
OI Model	X		X	X	X										
Duration															
Status as combined variable	X	X	X	X										X	
Categorical	X	X	X	X		X			X						
Continuous	X	X	X	X		X				X					
Only categorical	X	X	X	X					X						
Only continuous	X	X	X	X						X					
Packs per day															
Status as combined variable	X	X	X	X									X		
Categorical	X	X	X	X		X	X								
Continuous	X	X	X	X		X		X							
Only categorical	X	X	X	X			X								
Only continuous	X	X	X	X				X							
Packs per year															
Status as combined variable	X	X	X	X											X
Categorical	X	X	X	X		X					X				
Continuous	X	X	X	X		X						X			
Only categorical	X	X	X	X							X				
Only continuous	X	X	X	X								X			
Duration and packs per day															
Categorical	X	X	X	X		X	X		X						
Continuous	X	X	X	X		X		X		X					

[a] All the models were conducted first by including unknowns in the analysis and then by excluding unknowns.

[b] Covariates are from the main Silverman et al. 2012 models. Resp. disease = history of respiratory disease 5 years or more before date of death/reference date. High risk job = history of employment in a high risk job for lung cancer for at least 10 years.

[c] These are all combination smoking variables, in categorical form.

[d] REC = respirable elemental carbon. Four categories of REC were evaluated: average and cumulative REC, unlagged and lagged 15 years from the date of death/reference date.

[e] Location = location worked = surface only or ever-underground.

[f] OI combination smoking variable = smoking status–packs/day–location.

OI = original investigator (Silverman et al. 2012)

Table D.2. Effect of Smoking Variable Choice on Relationship between REC and Lung Cancer[a]

	Silverman et al. 2012[b] OR (95% Cl)	Status–Duration OR (95% Cl)	Status–Duration Interaction OR (95% Cl)
Average REC			
Average REC intensity, quartiles, unlagged (μg/m³)			
0 to < 1	1.0 (referent)	1.0	1.0
1 to < 32	1.027 (0.503–2.094)	0.782 (0.388–1.574)	0.707 (0.344–1.452)
32 to < 98	1.881 (0.759–4.663)	1.34 (0.541–3.316)	1.250 (0.499–3.127)
≥ 98	2.398 (0.889–6.465)	1.379 (0.496–3.835)	1.336 (0.476–3.748)
P_{trend}[c]	**0.025**[d]	0.188	0.137
Quartiles, lagged 15 years (μg/m³)			
0 to < 1	1.0	1.0	1.0
1 to < 6	1.109 (0.593–2.073)	1.791 (0.934–3.432)	1.699 (0.875–3.298)
6 to < 57	1.899 (0.904–3.988)	2.336 (1.061–5.143)	2.194 (0.994–4.843)
≥ 57	2.280 (1.067–4.872)	2.861 (1.268–6.454)	2.771 (1.231–6.236)
P_{trend}[c]	0.062	0.079	0.069
Cumulative REC			
Cumulative REC, quartiles, unlagged (μg/m³-yr)			
0 to < 19	1.0	1.0	1.0
19 to < 246	0.871 (0.476–1.594)	0.751 (0.410–1.377)	0.724 (0.389–1.345)
246 to < 964	1.501 (0.671–3.356)	1.281 (0.566–2.899)	1.219 (0.535–2.773)
≥ 964	1.745 (0.767–3.967)	1.401 (0.603–3.257)	1.348 (0.578–3.146)
P_{trend}[c]	0.083	0.083	0.183
Quartiles, lagged 15 years (μg/m³-yr)			
0 to < 3	1.0	1.0	1.0
3 to < 72	0.740 (0.398–1.375)	0.824 (0.437–1.553)	0.783 (0.413–1.488)
72 to < 536	1.538 (0.740–3.195)	1.66 (0.770–3.576)	1.607 (0.750–3.441)
≥ 536	**2.831 (1.279–6.263)**	**2.933 (1.271–6.768)**	**2.83 (1.229–6.519)**
P_{trend}[c]	**0.001**	**0.002**	**0.002**

Table continues next page

[a] All analyses conducted with 864 subjects, including 'unknowns'. Analyses excluding unknowns, analyses with continuous variables, and analyses with individual smoking variables can be found online in Additional Materials 3.

[b] Results from Silverman et al. (2012) repeated on each page for comparison purposes.

[c] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[d] **Bolded** values are significant at $P < 0.05$.

Table D.2 (*continued*). Effect of Smoking Variable Choice on Relationship between REC and Lung Cancer[a]

	Silverman et al. 2012[b] OR (95% Cl)	Status–Packs OR (95% Cl)[b]	Status–Packs + Interaction OR (95% Cl)[b]
Average REC			
Average REC intensity, quartiles, unlagged (μg/m³)			
0 to < 1	1.0	1.0	1.0
1 to < 32	1.027 (0.503–2.094)	1.052 (0.530–2.090)	0.968 (0.472–1.987)
32 to < 98	1.881 (0.759–4.663)	1.815 (0.744–4.429)	1.804 (0.723–4.499)
≥ 98	2.398 (0.889–6.465)	2.306 (0.869–6.121)	2.248 (0.826–6.116)
P_{trend}[c]	**0.025**[d]	**0.036**	**0.028**
Quartiles, lagged 15 years (μg/m³)			
0 to < 1	1.0	1.0	1.0
1 to < 6	1.109 (0.593–2.073)	1.199 (0.653–2.203)	1.076 (0.573–2.021)
6 to < 57	1.899 (0.904–3.988)	1.978 (0.947–4.133)	1.948 (0.924–4.107)
≥ 57	2.280 (1.067–4.872)	2.444 (1.147–5.211)	2.328 (1.086–4.990)
P_{trend}[c]	0.062	**0.048**	**0.054**
Cumulative REC			
Cumulative REC, quartiles, unlagged (μg/m³-yr)			
0 to < 19	1.0	1.0	1.0
19 to < 246	0.871 (0.476–1.594)	0.881 (0.488–1.589)	0.856 (0.466–1.575)
246 to < 964	1.501 (0.671–3.356)	1.501 (0.676–3.333)	1.488 (0.662–3.344)
≥ 964	1.745 (0.767–3.967)	1.742 (0.770–3.941)	1.725 (0.754–3.948)
P_{trend}[c]	0.083	0.083	0.084
Quartiles, lagged 15 years, (μg/m³-yr)			
0 to < 3	1.0	1.0	1.0
3 to < 72	0.740 (0.398–1.375)	0.815 (0.445–1.492)	0.728 (0.391–1.357)
72 to < 536	1.538 (0.740–3.195)	1.619 (0.785–3.340)	1.56 (0.749–3.249)
≥ 536	**2.831 (1.279–6.263)**	**2.998 (1.353–6.641)**	**2.881 (1.296–6.404)**
P_{trend}[c]	**0.001**	**0.001**	**0.001**

Table continues next page

[a] All analyses conducted with 864 subjects, including 'unknowns'. Analyses excluding unknowns, analyses with continuous variables, and analyses with individual smoking variables can be found online in Additional Materials 3.

[b] Results from Silverman et al. (2012) repeated on each page for comparison purposes.

[c] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[d] **Bolded** values are significant at $P < 0.05$.

Table D.2 (*continued*). Effect of Smoking Variable Choice on Relationship between REC and Lung Cancer[a]

	Silverman et al. 2012[b] OR (95% Cl)	Status–Packyears OR (95% Cl)	Status–Packyears + Interaction OR (95% Cl)
Average REC			
Average REC intensity, quartiles, unlagged (μg/m³)			
0 to < 1	1.0	1.0	1.0
1 to < 32	1.027 (0.503–2.094)	1.051 (0.530–2.083)	0.968 (0.482–1.944)
32 to < 98	1.881 (0.759–4.663)	1.713 (0.708–4.146)	1.658 (0.680–4.042)
≥ 98	2.398 (0.889–6.465)	2.151 (0.825–5.612)	1.993 (0.755–5.266)
P_{trend}[c]	**0.025**[d]	**0.052**	**0.057**
Quartiles, lagged 15 years (μg/m³)			
0 to < 1	1.0	1.0	1.0
1 to < 6	1.109 (0.593–2.073)	1.228 (0.668–2.257)	1.228 (0.662–2.277)
6 to < 57	1.899 (0.904–3.988)	1.766 (0.847–3.685)	1.907 (0.905–4.022)
≥ 57	2.280 (1.067–4.872)	2.145 (1.014–4.538)	2.299 (1.076–4.912)
P_{trend}[c]	0.062	0.094	0.073
Cumulative REC			
Cumulative REC, quartiles, unlagged (μg/m³-yr)			
0 to < 19	1.0	1.0	1.0
19 to < 246	0.871 (0.476–1.594)	0.91 (0.511–1.620)	0.825 (0.456–1.492)
246 to < 964	1.501 (0.671–3.356)	1.385 (0.635–3.020)	1.354 (0.614–2.986)
≥ 964	1.745 (0.767–3.967)	1.737 (0.775–3.895)	1.679 (0.742–3.799)
P_{trend}[c]	0.083	0.074	0.066
Quartiles, lagged 15 years (μg/m³-yr)			
0 to < 3	1.0	1.0	1.0
3 to < 72	0.740 (0.398–1.375)	0.763 (0.421–1.381)	0.721 (0.395–1.316)
72 to < 536	1.538 (0.740–3.195)	1.423 (0.703–2.880)	1.445 (0.709–2.943)
≥ 536	**2.831 (1.279–6.263)**	**2.499 (1.151–5.424)**	**2.658 (1.215–5.817)**
P_{trend}[c]	**0.001**	**0.002**	**0.001**

Table continues next page

[a] All analyses conducted with 864 subjects, including 'unknowns'. Analyses excluding unknowns, analyses with continuous variables, and analyses with individual smoking variables can be found online in Additional Materials 3.

[b] Results from Silverman et al. (2012) repeated on each page for comparison purposes.

[c] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[d] **Bolded** values are significant at $P < 0.05$.

Table D.2 (*continued*). Effect of Smoking Variable Choice on Relationship between REC and Lung Cancer[a]

	Silverman et al. 2012[b] OR (95% Cl)	Smoking Status, Packs/Day, Duration OR (95% Cl)
Average REC		
Average REC intensity, quartiles, unlagged (μg/m³)		
0 to < 1	1.0	1.0
1 to < 32	1.027 (0.503–2.094)	0.909 (0.452–1.829)
32 to < 98	1.881 (0.759–4.663)	1.668 (0.676–4.117)
≥ 98	2.398 (0.889–6.465)	1.959 (0.725–5.297)
P_{trend}[c]	**0.025**[d]	**0.048**
Quartiles, lagged 15 years (μg/m³)		
0 to < 1	1.0	1.0
1 to < 6	1.109 (0.593–2.073)	1.135 (0.616–2.089)
6 to < 57	1.899 (0.904–3.988)	1.872 (0.877–3.995)
≥ 57	2.280 (1.067–4.872)	2.277 (1.064–4.875)
P_{trend}[c]	0.062	0.063
Cumulative REC		
Cumulative REC, quartiles, unlagged (μg/m³-yr)		
0 to < 19	1.0	1.0
19 to < 246	0.871 (0.476–1.594)	0.931 (0.508–1.707)
246 to < 964	1.501 (0.671–3.356)	1.534 (0.676–3.482)
≥ 964	1.745 (0.767–3.967)	1.727 (0.755–3.952)
P_{trend}[c]		0.129
Quartiles, lagged 15 years (μg/m³-yr)		
0 to < 3	1.0	1.0
3 to < 72	0.740 (0.398–1.375)	0.793 (0.430–1.463)
72 to < 536	1.538 (0.740–3.195)	1.534 (0.735–3.204)
≥ 536	**2.831 (1.279–6.263)**	2.87 (1.285–6.412)
P_{trend}[c]	**0.001**	**0.001**

[a] All analyses conducted with 864 subjects, including 'unknowns'. Analyses excluding unknowns, analyses with continuous variables, and analyses with individual smoking variables can be found online in Additional Materials 3.

[b] Results from Silverman et al. (2012) repeated on each page for comparison purposes.

[c] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[d] **Bolded** values are significant at $P < 0.05$.

APPENDIX E. ALTERNATIVE RADON ANALYSES

CONTENTS

INTRODUCTION

Radon is a well-established risk factor for lung cancer (International Agency for Research on Cancer 2012), representing one of the most important causes of lung cancer after tobacco smoking (World Health Organization 2009). In the original Diesel Exhaust in Miners Study (DEMS*) case–control study (Silverman et al. 2012), the potential confounding effect of radon was not adjusted for in the relationship between diesel exhaust, measured as respirable elemental carbon (REC), and lung cancer mortality. The objective of the present analysis was to determine whether radon was a confounder for association between REC and lung cancer mortality in the study subjects. In addition, both a crude radon model (radon as the only covariate), and crude radon model with adjustment for various smoking variables, were used to evaluate the influence of radon on lung cancer risk in the DEMS case–control subjects. The methodology section describes the various analyses performed and the results that were tabulated. Some key results are described in the results section.

METHODOLOGY

Radon was available as a categorical (quartiles) and a continuous variable in the analytic data set provided by the National Cancer Institute. The confounding effect of radon was assessed by noting the difference in the odds ratios (OR) for the main effect of interest (i.e., the effect of REC on lung cancer mortality) in conditional logistic regression models with and with radon included as a covariate. All analyses were performed using SAS version 9.4 (SAS Institute. Cary, NC).

* A list of abbreviations and other terms appears at the end of the Special Report.

Variables Used in Radon Analysis

a. Radon exposure was measured in working level months (WLMs) and used in the models as either a continuous variable or as a categorical variable based on quartiles (unexposed; > 0 to < 0.64; 0.64 to < 1.87; 1.87 to < 2.98; ≥ 2.98 WLM).

b. Separate analyses were performed for various REC metrics, including average and cumulative REC, either unlagged or lagged 15 years from the death or reference date.

c. Duration of REC exposure, in years, was also used as a categorical variable in the analysis (unexposed; 0 to < 5; 5 to < 10; 10 to < 15; ≥ 15 years). The original investigators defined the unexposed category as "all subjects who worked surface jobs with either negligible or bystander exposure to REC, regardless of duration."

d. The common covariates used in all models were "history of respiratory disease 5 or more years before date of death/reference date" and "history of a high-risk job for lung cancer for at least 10 years." In what follows, these common covariates will be denoted simply as 'covariates'.

e. Smoking variables: As the main objective was to address the issue of confounding by radon in original analysis published by the original investigators, we used the same smoking status–location combination variable as was used by the original investigators (Silverman et al. 2012). We also conducted additional analyses using some of the smoking variables that were also used in the Panel's supplemental smoking analyses (Details of each of these smoking variables can be found in Appendix D):

- Smoking combination variable with location (Silverman et al. 2012). In the original analyses published by the study investigators, smoking was characterized by means of a combination variable that included smoking status, packs per day smoked, and location of employment (surface-only or ever-underground) as a single categorical variable.
- Smoking status and duration as a combined categorical variable (status–duration).
- Smoking status and packs per day as a combined categorical variable (status–packs/day).
- Smoking status and packyears as a combined categorical variable (status–packyears).

A categorical location variable (ever-underground or surface-only) was used to obtain certain descriptive statistics for radon by location.

Descriptive Analysis

Measures of radon exposure by case and control status and by location of employment (ever-underground; surface-only) were obtained for use in the analysis (Table E.1). In addition, associations between radon and REC were assessed using both Pearson and Spearman correlation coefficients Table E.2.

Analytical Models

Conditional logistic regression analysis was used to describe the relationship between diesel exhaust exposure and lung cancer mortality in the form of ORs. The confounding effect of radon was assessed by noting the difference in ORs of REC in models with and without radon: a change of more than 10% in the ORs for REC in models including radon compared with models excluding radon was used as a notional benchmark for confounding by radon. All models and variables included in the present analyses are summarized in a matrix provided in Table E.3.

Main Analyses to Determine the Confounding Effect of Radon To assess the potential confounding effect of radon, radon was added as an additional covariate in the original model used by (Silverman et al. 2012):

Model 1 (Silverman et al. 2012):

> *Lung cancer mortality = REC + smoking status–packs/day–location combination variable + covariates,*

where REC was described as average or cumulative REC, unlagged or lagged 15 years and the covariates were: 1) history of respiratory diseases 5 years or more before date of death or reference date; and 2) work in high risk jobs for cancer for at least ten years.

Model 1R_{cont}: Model 1 + radon as a continuous variable.

Model 1R_{cat}: Model 1 + radon as a categorical variable.

Because of the limited variability in radon levels measured in the mines, cumulative exposure to radon can be expected to be correlated with duration of REC exposure; to evaluate this possibility, both radon and duration of REC exposure were added as covariates to Model 1 above.

Model 2: Model 1 + Duration of REC as a categorical variable.

Model 2R_{cont}: Model 2 + Radon as a continuous variable.

Model 2R_{cat}: Model 2 + Radon as a categorical variable.

Results for these six sets of models are shown in Table E.4. The difference in ORs was obtained by comparing the ORs based on Model 1 with the ORs based on the other the models, for each of the four metrics for REC. Differences of more than 10% are underlined in the results tables. ORs that are significantly different from the null value of unity and significant trends ($P < 0.05$) are highlighted using a bold font. The test for trend was based on P values obtained by using categorical REC as a continuous variable in each of the models noted above. The median values within each category of each of the four metrics for REC were used to represent exposure for each subject in these models. The trend test was based on a Wald Chi-square statistic, using a significance level of $P < 0.05$.

For each of the above models that included radon as a covariate, the corresponding ORs for radon-related lung cancer were also estimated and are reported in Table E.5.

Additional Secondary Analyses Two additional sets of secondary analyses were performed. The first set was designed to evaluate directly the association between radon and lung cancer mortality in the DEMS case–control data set. These analyses were conducted using radon either as a continuous or categorical variable.

a. Crude analyses were performed including only radon, but not other covariates, in the models.
b. The crude radon model was adjusted one at a time for each of four different combination smoking variables (i.e., the original investigator's combination variable, status–packs/day, status–duration, and status–pack-years) to evaluate the potential confounding effects of smoking on the relationship between radon and lung cancer mortality.
c. Radon was added to the basic original investigators' models with the combination smoking-location variable and other covariates, after excluding REC from these models.

The results for these analyses can be found in Table E.6.

The second set of analyses assessed the confounding effect of radon and of REC on the relationship between duration of REC exposure and lung cancer. The overview of the models and their variables can be found in the matrix in Table E.7. In this analysis, the base model is essentially the same as Model 1 in the previous analyses (see Table E.3) but duration of exposure to REC, rather than average or cumulative REC, is used as the exposure variable. This model is then adjusted for radon as a continuous variable.

TABULATION OF RESULTS

Summary statistics describing radon exposures for cases and controls and by location (surface-only, ever-underground) are tabulated in Table E.1. Table E.2 summarizes Pearson and Spearman correlations between radon and different measures of REC. Analyses of the confounding effects of radon on the association between REC and lung cancer are reported in Table E.4; the corresponding ORs and 95% confidence intervals for radon in these models are presented in Table E.5.

The results of the first set of secondary analyses are summarized in Table E.6. Although the analyses with the categorical radon variable are subject to over-parameterization (i.e., some categories of the original investigators' smoking combination variable may be derived from a combination of remaining smoking categories and radon categories), SAS handles this problem by removing the over-parameterized category from the analysis. In fact, some over-parameterization did occur in the models using the 16 category combination smoking variables developed by the original investigators. (Because the primary focus of the supplemental analyses presented in this appendix are on the potential confounding effects of radon on the association between diesel exhaust and lung cancer mortality, categorical radon was entered first into models involving both categorical radon and smoking variables).

The results of secondary analyses using duration of REC as the primary REC exposure variable can be found in Table E.8.

RESULTS

Although a detailed discussion of the findings of the Panel's radon analysis will not be attempted here (see Chapter 4 for a more complete discussion), key results are described below:

a. Radon appears to be confounding the relationship between average and cumulative REC and lung cancer mortality to a certain extent, based on a change of more than 10% in the ORs for lung cancer associated with diesel exhaust following inclusion of radon in the main model (Table E.4).
b. In addition, a confounding effect of both radon and duration was observed for average REC (both unlagged and 15 years lagged REC), but not for cumulative REC lagged 15 years (Table E.4).
c. It should be noted that the confounding effect of radon persisted only for unlagged cumulative REC, but not average REC, either unlagged or lagged 15 years, or for cumulative REC lagged 15 years, after adjusting for the duration of REC (Table E.4).
d. Radon exposure (measured as cumulative radon exposure in WLM) is related to duration of underground mine work. At the HEI Diesel Epidemiology workshop in March 2014, Dr. Silverman suggested that 'double counting' of duration through the use of cumulative REC and cumulative radon measures could occur. It is difficult to disentangle the effects of radon and diesel exhaust on lung cancer risk, as adjusting for duration of REC exposure may have an effect similar to that obtained by adjusting for cumulative radon. The effect of this collinearity is evident in the analyses in Table E.8 where adjusting for radon in models including duration of REC exposure reduced the strength and significance of the trends in the ORs.
e. The ORs for cumulative REC (lagged 15 years) in the highest category remained significant even after adjusting for radon and/or duration, and also showed a significant positive trend. However, ORs for average REC (lagged 15 years) became nonsignificant after adjustment for radon and/or duration.

REFERENCES

International Agency for Research on Cancer. 2012. IARC: Diesel engine exhaust carcinogenic. Press Release 213. Lyon, France:IARC. Available: *www.iarc.fr/en/media-centre/pr/2012/pdfs/pr213_E.pdf*.

Silverman DT, Samanic CM, Lubin JH, Blair AE, Stewart PA, Vermeulen R, et al. 2012. The diesel exhaust in miners study: a nested case-control study of lung cancer and diesel exhaust. J Natl Cancer Inst 104:(11):855–868.

World Health Organization. 2009. WHO Handbook on Indoor Radon: A Public Health Perspective. WHO:Geneva. Available: *www.nrsb.org/pdf/WHO%20Radon%20Handbook.pdf*.

Table E.1. Cumulative Radon Summary Statistics by Case–Control Status

Status	Cases	Controls	Cases (*n*)		Control (*n*)	
			Surface Only[a]	Ever-Underground[b]	Surface Only	Ever-Underground
Number of subjects (*N*)	198	666	74	124	254	412
Radon (WLM)						
Mean	1.446	1.249		2.309		2.018
SD	1.893	1.751		1.932		1.845
Median	0.563	0.373		1.875		1.48
Minimum	0	0		0		0
Maximum	8.059	8.919		8.059		8.919
Radon quartiles (WLM)						
Not exposed	74	254				
Q1 0 to < 0.64	31	117				
Q2 0.64 to < 1.87	31	123				
Q3 1.87 to < 2.98	31	80				
Q4 ≥ 2.98	31	92				
Total exposed	124	412				

[a] Surface only = only worked in surface jobs.

[b] Ever-underground = ever worked underground for any time period.

Table E.2. Correlation of Radon with Average and Cumulative REC in the Case-Control Data[a,b]

REC	Correlation Coefficient (*P* Value)					
			Radon / Quartiles[c]			
	All Subjects	Exposed to Radon	Q1	Q2	Q3	Q4
Average REC						
Unlagged						
Pearson	0.36 (< 0.0001)	0.08 (0.08)	0.19 (0.02)	−0.01 (0.8808)	−0.16 (0.1022)	0.15 (0.1074)
Spearman	0.74 (< 0.0001)	0.16 (0.0003)	0.31 (0.0001)	0.0009 (0.9148)	−0.16 (0.0954)	0.17 (0.0607)
Lagged 15 years						
Pearson	0.35 (< 0.0001)	0.13 (0.003)	0.24 (0.0032)	0.04 (0.6213)	−0.13 (0.1677)	0.01 (0.8943)
Spearman	0.62 (< 0.0001)	0.27 (< 0.0001)	0.32 (< 0.0001)	0.09 (0.2569)	−0.1 (0.3075)	0.07 (0.4629)
Cumulative REC						
Unlagged						
Pearson	0.75 (< 0.0001)	0.67 (< 0.0001)	0.51 (< 0.0001)	0.39 (< 0.0001)	−0.13 (0.18101)	0.4 (< 0.0001)
Spearman	0.86 (< 0.0001)	0.7 (< 0.0001)	0.57 (< 0.0001)	0.34 (< 0.0001)	−0.16 (0.0847)	0.35 (< 0.0001)
Lagged 15 years						
Pearson	0.66 (< 0.0001)	0.58 (< 0.0001)	0.42 (< 0.0001)	0.16 (0.0482)	0.03 (0.7815)	0.29 (0.0014)
Spearman	0.68 (< 0.0001)	0.6 (< 0.0001)	0.41 (< 0.0001)	0.17 (0.0415)	0.03 (0.7705)	0.28 (0.0017)

[a] REC = respirable elemental carbon. Four categories of REC were evaluated: average and cumulative REC, unlagged and lagged 15 years from the date of death/reference date.

[b] Radon measured in working level months (WLM).

[c] Correlation coefficients between radon and REC within the specific quartiles of radon defined in Table E.1.

Table E.3. Summary of Models used in Analyses of Effect of Adjusting for Radon on REC–Lung Cancer Relationship[a]

Models vs. Variables	REC[c]	Covariates[b]		OI Smoking Combination Variable[d]	REC Duration[e]	Radon (WLM)		Combination Smoking Variables		
		Respiratory Disease	High-Risk Jobs			Continuous	Categorical	Status–Packs	Status–Duration	Status–Packyears
Main Models										
Model 1	X	X	X	X						
Model 1R_{cont}	X	X	X	X		X				
Model 1R_{cat}	X	X	X	X			X			
Model 2	X	X	X	X	X					
Model 2R_{cont}	X	X	X	X	X	X				
Model 2R_{cat}	X	X	X	X	X		X			
Secondary Analyses										
Radon as a continuous variable										
Model 1 without REC and radon		X	X	X		X				
Crude radon model						X				
Crude radon + OI smoking combination variable				X		X				
Crude radon + status–packs/day						X		X		
Crude radon + status–duration						X			X	
Crude radon + status–packyears						X				X
Radon as a categorical variable										
Model 1 without REC		X	X	X			X			
Crude radon model							X			
Crude radon + OI smoking combination variable				X			X			
Crude radon + status–packs/day							X	X		
Crude radon + status–duration							X		X	
Crude radon + status–packyears							X			X

[a] All smoking variables included unknowns in the analysis.

[b] Covariates from the main OI models. Resp. disease = history of respiratory disease 5 years or more before date of death/reference date. High risk job = history of employment in a high risk job for lung cancer for at least 10 years.

[c] REC = respirable elemental carbon. Four categories of REC were evaluated: average and cumulative REC, unlagged and lagged 15 years.

[d] OI combination smoking variable = smoking status–packs/day–location.

[e] Duration of REC = duration of exposure to REC. "Unexposed includes all subjects who worked surface jobs with either negligible or bystander exposure to REC, regardless of duration." (Silverman et al. 2012)

OI = original investigator (Silverman et al. 2012); 95% CI = 95% confidence interval; WLM = working level months.

Table E.4. Assessment of the Confounding Effect of Radon and/or Duration of REC Exposure on the Relationship between REC and Lung Cancer Mortality[a]

Exposure Metric	**Model 1** OR (95% CI)	**Model 1R$_{cont}$** OR (95% CI)	**Model 1R$_{cat}$** OR (95% CI)	% Difference in OR[b,c] 1 vs 1R$_{cont}$	% Difference in OR[b,c] 1 vs 1R$_{cat}$
Average REC Quartiles, Unlagged (µg/m³)					
0 to < 1	1 referent	1	1		
1 to < 32	1.027 (0.503–2.094)	1.028 (0.503–2.102)	1.026 (0.502–2.096)	0.10	−0.10
32 to < 98	1.881 (0.759–4.663)	1.578 (0.626–3.982)	1.722 (0.675–4.389)	−16.11	−8.45
≥ 98	2.398 (0.889–6.465)	1.959 (0.714–5.370)	2.206 (0.806–6.041)	−18.31	−8.01
P_{trend}[d]	**0.025**[e]	0.101	0.051		
−2logL	468.334	463.027	464.838		
AIC	512.334	509.027	514.838		
Average REC Quartiles, Lagged 15 Years (µg/m³)					
0 to < 1	1	1	1		
1 to < 6	1.109 (0.593–2.073)	1.11 (0.591–2.082)	1.092 (0.581–2.051)	0.09	−1.53
6 to < 57	1.899 (0.904–3.988)	1.582 (0.734–3.410)	1.626 (0.746–3.543)	−16.69	−14.38
≥ 57	**2.28** (1.067–4.872)	1.834 (0.831–4.047)	1.948 (0.886–4.285)	−19.56	−14.56
P_{trend}	0.062	0.207	0.142		
−2logL	469.61	463.971	466.616		
AIC	513.61	509.971	516.616		
Cumulative Quartiles, Unlagged (µg/m³-yr)					
0 to < 19	1	1	1		
19 to < 246	0.871 (0.476–1.594)	0.954 (0.517–1.758)	0.921 (0.500–1.696)	9.53	5.74
246 to < 964	1.501 (0.671–3.356)	1.438 (0.638–3.242)	1.607 (0.683–3.782)	−4.20	7.06
≥ 964	1.745 (0.767 to 3.967)	1.162 (0.472–2.856)	1.52 (0.592–3.904)	−33.41	−12.89
P_{trend}	0.083	0.961	0.566		
−2logL	469.978	464.732	467.113		
AIC	513.978	510.732	517.113		
Cumulative Quartiles, Lagged 15 Years (µg/m³-yr)					
0 to < 3	1	1	1		
3 to < 72	0.74 (0.398–1.375)	0.754 (0.405–1.405)	0.722 (0.387–1.349)	1.89	−2.43
72 to < 536	1.538 (0.740–3.195)	1.446 (0.684–3.055)	1.531 (0.724–3.238)	−5.98	−0.46
≥ 536	**2.831** (1.279–6.263)	2.263 (0.938–5.457)	**2.746** (1.125–6.699)	−20.06	−3.00
P_{trend}	**0.001**	**0.031**	**0.009**		
−2logL	461.065	459.59	460.012		
AIC	505.065	505.59	510.012		

Table continues next page

[a] See Table E.3 for model and variables.

[b] % Difference in ORs take the general form = $([OR_{model\ 1R} - OR_{model\ 1}]/OR_{model\ 1}) \times 100$.

[c] Underlined values show a greater than 10% difference between the ORs, adjusted for radon and the ORs in the model, unadjusted for radon.

[d] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[e] **Bolded** values are significant at $P < 0.05$.

95% CI = 95% confidence interval; OI = Original Investigators, Silverman and colleagues (2012); AIC = Akaike information criterion; −2logL = minus 2 log likelihood; REC = respirable elemental carbon; WLM = working level months, a measure of cumulative exposure to radon.

Table E.4 (*continued*). Assessment of Confounding Effect of Duration of REC exposure and/or Radon on the Relationship between REC and Lung Cancer Mortality[a]

Exposure Metric	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 2R$_{cont}$ OR (95% CI)	Model 2R$_{cat}$ OR (95% CI)	% Difference in OR[b,c] 1 vs 2	1 vs 2R$_{cont}$	1 vs 2R$_{cat}$
Average REC Quartiles, Unlagged (µg/m³)							
0 to < 1	1	1	1	1			
1 to <32	1.027 (0.503–2.094)	0.831 (0.301–2.299)	0.878 (0.317–2.437)	0.878 (0.316–2.439)	−19.08	−14.51	−14.51
32 to < 98	1.881 (0.759–4.663)	1.629 (0.516–5.150)	1.589 (0.501–5.035)	1.608 (0.506–5.104)	−13.40	−15.52	−14.51
≥ 98	2.398 (0.889–6.465)	1.999 (0.621–6.437)	1.893 (0.586–6.119)	2.031 (0.624–6.603)	−16.64	−21.06	−15.30
P_{trend}	**0.025**	**0.033**	0.071	**0.041**			
−2logL	468.334	459.453	457.405	457.292			
AIC	512.334	511.453	511.405	515.292			
Average REC Quartiles, Lagged 15 Years (µg/m³)							
0 to < 1	1	1	1				
1 to < 6	1.109 (0.593–2.073)	1.112 (0.579–2.136)	1.133 (0.589–2.179)	1.126 (0.584–2.172)	0.27	2.16	1.53
6 to < 57	1.899 (0.904–3.988)	1.779 (0.804–3.934)	1.679 (0.754–3.742)	1.664 (0.745–3.713)	−6.32	−11.59	−12.37
≥ 57	**2.28** (1.067–4.872)	2.076 (0.937–4.600)	1.891 (0.843–4.246)	1.969 (0.881–4.398)	−8.95	−17.06	−13.64
P_{trend}	0.062	0.123	0.213	0.153			
−2logL	469.61	462.365	459.575	459.938			
AIC	513.61	514.365	513.575	517.938			
Cumulative Quartiles, Unlagged (µg/m³-yr)							
0 to < 19	1	1	1	1			
19 to < 246	0.871 (0.476–1.594)	0.942 (0.503–1.763)	1.062 (0.560–2.016)	1.032 (0.542–1.964)	8.15	21.93	18.48
246 to < 964	1.501 (0.671–3.356)	1.956 (0.811–4.718)	2.073 (0.857–5.014)	1.962 (0.803–4.794)	30.31	38.11	30.71
≥ 964	1.745 (0.767–3.967)	1.88 (0.707–4.997)	1.672 (0.623–4.488)	1.754 (0.650–4.728)	7.74	−4.18	0.52
P_{trend}	0.083	0.336	0.765	0.59			
−2logL	469.978	461.328	458.25	459.57			
AIC	513.978	513.328	512.25	517.57			
Cumulative Quartiles, Lagged 15 years (µg/m³-yr)							
0 to < 3	1	1	1	1			
3 to < 72	0.74 (0.398–1.375)	0.671 (0.356–1.266)	0.696 (0.367–1.318)	0.687 (0.363–1.300)	−9.32	−5.95	−7.16
72 to < 536	1.538 (0.740–3.195)	1.542 (0.719–3.308)	1.54 (0.714–3.320)	1.509 (0.702–3.244)	0.26	0.13	−1.89
≥ 536	**2.831** (1.279–6.263)	**2.828** (1.146–6.977)	**2.599** (1.032–6.544)	**2.69** (1.064–6.799)	−0.11	−8.19	−4.98
P_{trend}	**0.001**	**0.004**	**0.015**	**0.01**			
−2logL	461.065	453.906	452.957	453.115			
AIC	505.065	505.906	506.957	511.115			

Table continues next page

[a] See Table E.3 for model and variables.

[b] % Difference in ORs take the general form = $([OR_{model\ 1R} - OR_{model\ 1}]/OR_{model\ 1}) \times 100$.

[c] Underlined values show a greater than 10% difference between the ORs, adjusted for radon and the ORs in the model, unadjusted for radon.

[d] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[e] **Bolded** values are significant at $P < 0.05$.

95% CI = 95% confidence interval; OI = Original Investigators, Silverman and colleagues (2012); AIC = Akaike information criterion; −2logL = minus 2 log likelihood; REC = respirable elemental carbon; WLM = working level months, a measure of cumulative exposure to radon.

Table E.5. Radon and Lung Cancer Odds Ratios Corresponding to the Radon-adjusted Models Depicted in Table E.4[a]

Exposure Metric	Model 1 + Radon OR (95% CI)	Model 2 + Radon OR (95% CI)
Radon Quartiles + Average REC, Unlagged		
Radon as quartiles (WLM)		
Not exposed	1 referent	1
> 0 to < 0.64	1.93 (0.530–7.026)	1.217 (0.296–5.013)
0.64 to < 1.87	1.797 (0.469–6.891)	1.920 (0.486–7.582)
1.87 to < 2.98	2.654 (0.695–10.126)	2.726 (0.651–11.413)
≥ 2.98	3.33 (0.877–12.649)	2.466 (0.562–10.819)
P_{trend}[b]	< 0.086	< 0.400
Radon as continuous (WLM)	**1.181 (1.026–1.360)**	1.162 (0.946–1.428)
P value	**0.0205**	0.1533
Radon Quartiles + Average REC, Lagged 15 years		
Radon as quartiles (WLM)		
Not exposed	1	1
> 0 to < 0.64	1.984 (0.581–6.783)	1.443 (0.358–5.810)
0.64 to < 1.87	2.074 (0.594–7.236)	2.491 (0.652–9.516)
1.87 to < 2.98	2.766 (0.777–9.853)	3.255 (0.773–13.700)
≥ 2.98	3.558 (0.982–12.892)	3.071 (0.694–13.584)
P_{trend}[b]	< 0.090	< 0.377
Radon as continuous (WLM)	**1.188 (1.032–1.368)**	1.190 (0.970–1.461)
P value	**0.017**	0.095
Radon Quartiles + Cumulative REC, Unlagged		
Radon as quartiles (WLM)		
Not exposed	1	1
> 0 to < 0.64	2.28 (0.667–7.795)	1.373 (0.333–5.668)
0.64 to < 1.87	1.97 (0.525–7.385)	2.023 (0.496–8.260)
1.87 to < 2.98	2.786 (0.749–10.362)	2.819 (0.651–12.213)
≥ 2.98	3.792 (0.981–14.650)	2.914 (0.635–13.370)
P_{trend}[b]	< 0.146	< 0.326
Radon as continuous (WLM)	**1.224 (1.030–1.454)**	1.216 (0.977–1.515)
P value	**0.0216**	0.0803
Radon Quartiles + Cumulative REC, Lagged 15 yrs		
Radon as quartiles (WLM)		
Not exposed	1	1
> 0 to < 0.64	2.037 (0.602–6.890)	1.410 (0.356–5.594)
0.64 to < 1.87	1.586 (0.447–5.623)	1.828 (0.476–7.021)
1.87 to < 2.98	1.915 (0.531–6.900)	2.260 (0.537–9.515)
≥ 2.98	2.202 (0.592–8.182)	1.902 (0.426–8.490)
P_{trend}[b]	< 0.660	< 0.823
Radon as continuous (WLM)	1.105 (0.941–1.297)	1.112 (0.898–1.377)
P value	0.223	0.3305

[a] See Table E.3 for model and variables.

[b] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[c] **Bolded** values are significant at $P < 0.05$.

95% CI = 95% confidence interval; OR = odds ratio; REC = respirable elemental carbon; WLM = working level months, a measure of cumulative exposure to radon.

Table E.6. Six Secondary Analyses: Radon Odds Ratios Corresponding to Crude Radon Models, Adjusted for Smoking[a]

Exposure Metric	Number of Cases/ Controls	Model 1 Without REC + Radon[b]	Crude Radon Model	Crude Radon Model + Combination Smoking Variable: OI Smoking Combination Variable[b,c]	Crude Radon Model + Combination Smoking Variable: Smoking Status–Packs/Day	Crude Radon Model + Combination Smoking Variable: Smoking Status – Duration	Crude Radon Model + Combination Smoking Variable: Smoking Status–Packyears
Radon quartiles (WLM)							
Not exposed	74/254	1 referent	1	1	1	1	1
> 0 to < 0.64	31/117	0.500 (0.129–1.940)	0.819 (0.487–1.376)	0.639 (0.174–2.348)	0.731 (0.428–1.250)	0.783 (0.448–1.370)	0.734 (0.427–1.262)
0.64 to < 1.87	31/123	0.540 (0.137–2.132)	0.847 (0.520–1.379)	0.682 (0.184–2.524)	0.847 (0.503–1.429)	0.765 (0.446–1.313)	0.820 (0.490–1.373)
1.87 to < 2.98	31/80	0.780 (0.201–3.024)	1.342 (0.808–2.230)	0.952 (0.257–3.533)	1.104 (0.642–1.896)	1.211 (0.688–2.131)	1.084 (0.631–1.861)
≥ 2.98	31/92	1.052 (0.265–4.171)	1.230 (0.742–2.040)	1.124 (0.299–4.233)	1.327 (0.761–2.314)	1.103 (0.612–1.988)	1.179 (0.670–2.075)
P_{trend}[d]	198/666	**0.024**[e]	0.198	0.07	0.198	0.489	0.386
Radon continuous (WLM)	198/666	**1.224 (1.070–1.401)**	1.078 (0.986–1.178)	**1.181 (1.039–1.344)**	1.101 (0.996–1.217)	1.067 (0.961–1.186)	1.073 (0.970–1.187)
P value		**0.003**	0.098	**0.011**	0.061	0.223	0.169

[a] See Table E.3 for definition of models and variables.

[b] OI combination smoking variable = smoking status–packs/day–location.

[c] This variable was rendered to zero for some categories (i.e., no estimate was possible) when the smoking combination variable and radon quartiles were in linear combination (i.e., over parameterized).

[d] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[e] **Bolded** values are significant at $P < 0.05$.

95% CI = 95% confidence interval; OI = Original Investigators, Silverman and colleagues (2012); REC = respirable elemental carbon; WLM = working level months, a measure of cumulative exposure to radon.

Table E.7. Overview of Models and Covariates Used to Evaluate the Effect of Radon on the Relationship Between Duration of REC Exposure and Lung Cancer Mortality

Models vs Variables	Duration of REC[b]	REC Metric[c]	Covariates[a]		OI Smoking Combination Variable	Radon Continuous (WLM)
			Respiratory Disease	High Risk Jobs		
Model 1D[d]	**X**		**X**	**X**	**X**	
Model $1DR_{cont}$[e]	**X**		**X**	**X**	**X**	**X**

[a] Covariates from the main Silverman et al. 2012 models. Respiratory disease = history of respiratory disease 5 years or more before date of death/reference date. High risk job = history of employment in a high risk job for lung cancer for at least 10 years. OI combination smoking variable = smoking status–packs/day–location.

[b] Duration of REC = duration of exposure to REC. "Unexposed includes all subjects who worked surface jobs with either negligible or bystander exposure to REC, regardless of duration" (Silverman et al. 2012).

[c] REC = respirable elemental carbon; average and cumulative REC, unlagged and lagged 15 years from the date of death/reference date.

[d] Model 1D (Model 1-Duration) is the Model 1 (OI Main model in Table E.3) but with Duration of REC exposure only as a measure of exposure to REC instead of average and cumulative REC as in the main model.

[e] Model $1DR_{cont}$ = Model 1D adjusted for continuous radon.

OI = Original Investigators, Silverman and colleagues (2012); WLM = working level months.

Table E.8. Assessment of Confounding Effect of Radon on the Relationship Between Duration of REC Exposure and Lung Cancer Mortality[a]

Exposure Metric / Quintiles	Model 1D OR (95% CI)	Model $1DR_{cont}$ OR (95% CI)
Duration of REC Exposure (yr)[b]		
Unexposed	1	1
0 to < 5	1.16 (0.53–2.55)	1.38 (0.61–3.09)
5 to < 10	0.88 (0.38–2.03)	0.90 (0.39–2.09)
10 to < 15	0.93 (0.39–2.21)	0.82 (0.34–1.98)
≥ 15	2.09 (0.89–4.90)	1.32 (0.50–3.51)
P_{trend}[c]	**0.043**[d]	0.950
−2logL	466.103	462.205
AIC	512.103	510.205

[a] See Table E.7 for definition of models and variables.

[b] "Unexposed includes all subjects who worked surface jobs with either negligible or bystander exposure to REC, regardless of duration" (Silverman et al. 2012).

[c] *P* values for tests of trend conducted with 2-sided Wald test. For categorical analyses, all subjects were assigned the median for each quartile.

[d] **Bolded** values are significant at $P < 0.05$.

95% CI = 95% confidence interval; OR = odds ratio; REC = respirable elemental carbon; AIC = Akaike information criterion; − 2logL = minus 2 log likelihood.

APPENDIX F. TECHNICAL BACKGROUND AND FURTHER EVALUATION OF THE DEMS RETROSPECTIVE EXPOSURE MODEL

CONTENT

INTRODUCTION

Over the course of its deliberations the Panel received a great deal of input from the engine manufacturing and engineering communities regarding the exposure assessment used in the Diesel Exhaust in Miners Study (DEMS*). The focus of this input, which came in the form of both published papers and presentations, was on the validity of the underlying assumptions used to estimate historical exposures. These concerns were considered by the Panel in its critical evaluation of the DEMS exposure assessment approach as part of its assessment of the utility of DEMS for quantitative risk assessment.

Chapter 4 of the report reflects the Panel's overall conclusion that the DEMS retrospective exposure assessment was a logical, thorough, well-designed effort that ultimately yielded estimates of historical respirable elemental carbon (REC) levels that were consistent with the limited measurement data available to the original investigators and with levels that have been observed in other mines, albeit not without uncertainties that would need to be considered in using the results of the analysis for risk assessment purposes.

This appendix, authored primarily by Dr. David Foster with input from other members of the Panel, provides detailed technical background to the Panel's efforts to understand, and where possible, reconcile the differing views on the fundamental assumptions of the DEMS retrospective exposure model and their implications for uncertainty in the REC exposure estimates. Specifically, the appendix explores 1) the apparent assumption of a 'universal' relationship among REC emissions (a surrogate for particulate matter [PM]), carbon monoxide (CO) emissions, and horsepower (HP) across all engines and mines; and 2) the potential impact on the interrelationships between these pollutants, and thus on estimated exposures, from changes in engine technology over time (Borak et al. 2011; Crump and Van Landingham 2012; Hesterberg et al. 2012a; McClellan et al. 2012).

CHALLENGES FOR ESTIMATION OF HP–CO–REC RELATIONSHIPS

The basic concepts behind the retrospective exposure model are logical as a first approximation — that as more diesel engine HP is used in a mine, more diesel exhaust is emitted, and as ventilation is increased, the concentration of diesel exhaust in the mine air is reduced. These concepts are illustrated by a series of figures shown by Dr. Vermeulen in his presentation at the March 2014 HEI Workshop, reproduced in a more generalized form below as Figure F.1.

The logic is also supported by basic box, or compartmental, modeling principles in which a steady-state contaminant concentration (e.g., in g/m^3) is directly proportional to the emissions rate (e.g., in g/min) divided by a ventilation rate (e.g., ft^3/min [CFM]). The importance of the variable AdjHP/CFM (adjusted horsepower/CFM), a surrogate for the emissions rate divided by the ventilation rate, in the DEMS exposure model can be seen by comparing the graph of this variable for the mine operator in Mine E (Figure F.2, left panel) to the predicted REC exposures from the DEMS model for the period of the study (Figure F.2, right panel, from Vermeulen et al. 2010a).

While the temporal patterns are consistent, translating from the surrogate emissions term (AdjHP/CFM) to an estimate for personal exposure to REC for each year requires additional calculations. It is the assumptions underlying this step in the model that have been most controversial, specifically, the appropriateness of the assumption of consistent relationships between HP (AdjHP/CFM) and CO concentrations and between CO and REC concentrations over time (Borak et al. 2011; Crump and Van Landingham 2012; Hesterberg et al. 2012a; McClellan et al. 2012). The controversy lies in what these assumptions imply or assume about the underlying characteristics of engine emissions. To better understand the basis for these criticisms, it is helpful to understand some basic diesel engine combustion fundamentals.

* A list of abbreviations and other terms appears at the end of the Special Report.

Diesel Engine Combustion Fundamentals

The HP of an engine is usually presented in the form of an engine map, which is constructed by measuring the engine work output at different conditions and displaying it on a plot of work output (load) versus speed. The precision of the HP measurement and its presentation in the engine map is not sufficiently resolved to distinguish the extremely small variations in HP that would occur with routinely measurable variations in the CO emitted from the engine. Because of the much larger amounts of carbon dioxide (CO_2) relative to CO in emissions, and the impact of additional factors which affect the HP, there can be large variations in CO that could only be detected by very high resolution measurements of the CO_2, and such resolution is not present in the HP measurements used for rating the engine. In short, within the precision of the HP data presented in an engine map, and the fact that the HP is influenced by factors in addition to the completeness of the burn, such as the phasing and duration of the combustion relative to the piston position, there can be large variations in the emission of CO which will not be distinguishable though assessment of changes in HP.

Figure F.3 gives some indication of this challenge when looking at a group of engines by plotting Ln(CO) against Ln(HP) data from a set of detailed chassis dynamometer tests for more than 250 vehicles spanning model years 1981 to 1997 (Yanowitz et al. 2000). Even in natural log transformed space there is large scatter in the data and the correlation between CO and HP yielded an $R^2 = 0.01$.

In addition to the HP–CO relationship, concerns have also been expressed regarding the assumption of a consistent CO–REC relationship across engines over time (e.g., $\beta = 1$). The assumption of a 1:1 proportionality in CO:REC concentrations over time in the DEMS model was based, not on the figure above, but on the figure by Yanowitz and colleagues (2000) showing similar decline in PM and CO emissions for engines arrayed by model year.

At the level of individual engine combustion, several factors directly affect the emission levels of CO and PM at a given HP, principally the completeness of combustion, the completeness of the air-fuel mixing processes, the duration of the combustion, and the phasing of the combustion relative to the piston position within the expansion stroke. The fuel-air mixing process is essential for complete diesel engine combustion and different mixing histories can drive potential relationships between CO and PM in the opposite direction. Because of different engine geometries, engine intake configurations, injection equipment, etc., the relative quantities of over-mixed and/or under-mixed regions within the cylinder during combustion will vary from engine to engine and among operating conditions for a given engine. Depending on the details of combustion processes, two engines of the same HP could have drastically different emission profiles, and very different ratios of CO to HP and PM to CO. Changes in fuels, engine design, and aftertreatment technologies that affect these variables are also likely to affect the CO–PM relationship for a given engine.

Variability in HP–CO–REC Emissions Relationships

These factors all suggest that, at a given point in time, there could be significant variability in CO and REC emissions characteristics and their relationships with HP. This argument is reinforced by the next set of Figures F.4 through F.6 as one progresses from studying the performance of a single engine to the more complex multiengine, multiduty cycle operating conditions. In each of these figures, PM is assumed to be a proxy for REC.

The first figure, Figure F.4, shows that for a single 2004 engine, operating on specified duty cycles, there is a strong relationship between CO and PM emissions (Xu et al. 2005).

The second figure, Figure F.5, illustrates how the CO and PM emissions are still positively related, but the relationship is more variable and likely weaker (no statistical analysis was provided) when different engines are operated on a specified duty cycle (Clark et al. 1999).

Finally, in an experiment where different engines and vehicles were tested over a range of driving cycles, Xu and colleagues (2005) found that there was no correlation between PM and CO emissions (We note, however, the scaling of the y-axis in Figure F.6 unfortunately makes it difficult to see what the relationships are at lower levels and no statistical analyses of the data were provided).

What these figures suggest is that it is not surprising that Vermeulen and colleagues (2010b), under real-world conditions, also observed substantial scatter or variability in the relationships between the concentrations of CO and REC measured in the mines during the DEMS survey in 1998–2001 (Figure F.7). Their data in part reflect variability in the underlying relationships between CO and REC emitted from the individual engines operated in each mine for the period of the DEMS survey data. The presence of modest correlations between CO and REC concentrations in the mines reflects the common influence of ventilation (CFM), source density or proximity, and other factors operating on those same emissions.

In a sense, the group of engines in use in a mine during any given year or time period could be considered a single aggregate engine and one could infer that the HP produced by those engines during the year was indicative of a specified duty cycle. Time periods over which the HP in a mine was constant from year to year could then be interpreted as periods for which the work pattern, or duty cycle, for that mine also stayed constant. In such a case, one could expect to see a

correlation between the CO and PM emissions for those periods as long as the "aggregate engine" did not change drastically during that time period. As the mix of engines changes relatively slowly from year to year in a given mine, this assumption seems plausible for some time intervals.

Figure F.8 graphs the adjusted horsepower and ventilation (kCFM) used for each year in each mine over the full period of the study (1947–2001) (These and related data were also available in the National Institute for Occupational Safety and Health (NIOSH) data made available on the National Cancer Institute Web site, *http://dceg.cancer.gov/research/what-we-study/environment/diesel-exhaust-miners-study-dems*). For the time period from 1998–2001, AdjHP was approximately constant for most mines in comparison with the changes in AdjHP that had occurred over the entire period of the study. From reviewing the NIOSH data, the average age of the engines did not change appreciably over the time period from 1998–2001. Consequently, the assumption of a specific aggregate engine operating on a single duty cycle for the period 1998–2001 seemed reasonable. Coupled with the influence of ventilation acting on a common source of emissions, a correlation between CO and PM concentrations could likely be observed for this time period for each of the mines. This would be somewhat analogous to analyzing the relationship between CO and PM emissions from the data shown above in Figure F.5. Given that each mine would have a different mine-specific aggregate engine and a different duty cycle, as signified by a different AdjHP for this time period, one could expect a different correlation between CO and PM concentrations for each mine. This is what the DEMS data show (see Figure F.7) where one sees relatively strong correlations in some mines (Facility E), and weaker correlations in others (Facility I).

The impact of engine operating conditions on variability in emissions is reflected in recent Mine Safety and Health Administration (MSHA) ventilation guidelines (Tomko 2013). MSHA classifies engines via a particulate index (PI) which designates the amount of ventilation necessary to dilute an engine's exhaust PM to 1000 $\mu g/m^3$ (although it was not clear at which engine year these classifications began). Particulate indices for MSHA-approved engines are listed on their internet website. Once mine operators know the engine's PI they can determine the necessary ventilation via multiples of the PI. For example, 2 × PI will dilute the PM in the exhaust stream to 500 $\mu g/m^3$, 5 × PI would dilute the PM in engine's exhaust stream to 200 $\mu g/m^3$, and so on. Their guidelines give an example in which two different engines, both producing 150 HP, required ventilation flow rates that were different by a factor of 9 to achieve the same PM concentration in the exhaust.

IMPACTS OF TEMPORAL TRENDS IN ENGINE TECHNOLOGY ON EMISSIONS AND ON EXPOSURES ESTIMATED IN THE MINES

In addition to the uncertainties in the HP–CO–REC emissions relationships discussed above, a further concern raised in critiques of the DEMS study, discussed in Chapter 4, is how appropriate it is to use the empirical regression relationships observed in 1998–2001 in the estimation of the relative trends in CO and REC concentrations back to the start of mine dieselization. For this to be the case, the underlying engines and operating conditions (e.g., the aggregate engine working at a mine specific duty cycle) would have to hold for the entire period of the study. The variations in the AdjHP with time for each mine shown in Figure F.8, coupled with NIOSH data on the average engine model year over time, show that there were large changes in the number and age of engines in the mines throughout the period of the study. The underlying engines and operating conditions that gave rise to the regression relationships estimated from the 1998–2001 data would not necessarily be the same over time.

The DEMS investigators recognized that these changes in engines, aftertreatment systems, and fuels were important. They included an explicit indicator variable for post-1990 engines in their CO regression model to evaluate the potential impact of newer engine technologies. Their analyses indicated that the presence of post-1990 engines reduced the estimated CO concentrations. As discussed in more detail in the text, the DEMS investigators also attempted to address uncertainty in this relationship by conducting limited sensitivity analyses using alternative proportionality constants ($\beta = 1$ in their main model versus $\beta = 0.58$ from their regression analysis) in *RELtrend* (Equation 2, Chapter 4). These efforts continued to rely on the underlying HP–CO–REC model, with its attendant uncertainties, discussed above.

The following sections provide support for the discussion of this issue in Chapter 4 beginning with 1) scientific background on the impacts of changes in engine, fuel, and exhaust aftertreatment technology on emissions from diesel engine over time; followed by 2) an evaluation of changes in emissions implicit in the DEMS model and an illustrative analysis to examine the uncertainties in REC concentrations in the mines associated with different assumptions about the rate at which fuel and technological innovations were reflected in the mines.

Impacts of Changes in Engine, Fuel, and Exhaust Aftertreatment Technology on Emissions

Innovations in engine technology have been motivated by the desire to improve the air utilization within the combustion chamber, to obtain more power from the same size engine, and to reduce fuel consumption. For example, early

diesel engines were naturally aspirated with low pressure pump-line injectors. In the late 1940's and early 1950's, turbocharged engines started being introduced which were in turn followed by charge-air cooling, and by air system and combustion chamber design changes. In the late 1970's the low-pressure pump-line injection systems started being replaced with high pressure unit injectors, which also brought electronic control of the injection system and the advent of multiple injections per combustion event. Ultimately by the late 1990's high pressure electronically controlled flexible injection systems were in wide spread application. These innovations resulted in better air utilization, a significant reduction in PM emissions and altered relationships between CO and HP and CO and PM in those emissions.

In addition to engine technology, emission aftertreatment systems were being added to the engines. Particularly relevant to the studies in question would be the addition of oxidation catalysts to the exhaust system, which began in the mid to late 1970's. Oxidation catalysts promote the conversion of CO and gas phase hydrocarbons into CO_2 and water, as well as promoting the oxidation of hydrocarbons adsorbed onto the surface of the PM into CO_2 and water. Their use in the mine would have been motivated by the desire to reduce CO concentrations in the exhaust. Consequently, oxidation catalysts would reduce the CO concentration in the exhaust without changing the REC concentration. Finally, changes in the sulfur content of fuel contributed to reductions in PM in later years. Fuel sulfur contributes to PM formation (as sulfates), which acts as nucleation sites for the initial formation of PM, including elemental carbon. It also makes possible the use of different aftertreatment systems.

Many of these changes in engine technology and on emissions have been influenced by regulations. Figure F.9 demonstrates the decreasing trend in PM emissions from onroad diesel engines over time, superimposed by the limits imposed by heavy-duty diesel PM standards over time. These data were obtained by running production, onroad, heavy-duty engines through a transient emission test and measuring the exhaust leaving the engine. They indicate that the major changes in emissions from these engines began in the late 1980's and show a roughly 8-fold decrease on average in particulate emissions per brake-HP-hour between 1975 and 1995.

Impact of Changing Technology and Emissions Trends on Uncertainty in Estimated REC Exposures in the Mines

Without a much more detailed comparison of the engines used in the mines with the timing of these changes in technology described above and inclusion of the history of the fuel composition, it is difficult to predict the quantitative impact of these concerns on historical estimates of REC over the course of the study. Such a comparison might be done given the data made available by the DEMS investigators on the engines used in the mines and if additional emissions data for these engines could be obtained.

As an alternative, we explore this issue with an illustrative comparison of trends in PM emissions from heavy-duty diesel engines with trends in REC emissions implicit in the DEMS model. An advantage of this approach is that it does not rely on the HP–CO–REC assumptions embedded in the DEMS model. Again using Mine E and exposure to the mine operator as an example, the first step was to estimate the implicit trend in PM emission rates in the DEMS model.

We have seen in Figure F.2 for Mine E that REC exposure is proportional to the determinant AdjHP/CFM. If one were to convert the right hand graph shown in Figure F.2 from AdjHP/CFM to an exposure of μg/m^3 it would represent the exposure of the Mine E operator assuming that the emissions of PM per unit HP from the engines were constant at the 1998 engine level over the entire period of the study. That is, it would remove the CO/REC proportionality from the model and treat the PM emission rate from the engine as being equal to the 1998 aggregate engine throughout the period of study.

To make this conversion one would need to convert AdjHP/CFM curves, such as that shown for Mine E in Figure F.2, into concentrations. Then one could compare the results to the exposures shown on the left. The conversion can be accomplished by anchoring the value of the AdjHP/CFM curve on the right hand graph to the 1998 REC value from the DEMS survey shown on the left (139.7 μg/m^3), and thus determining a conversion factor between AdjHP/CFM and exposure in μg/m^3. We refer to this as a "simple model" and the equation for the mine operator in Mine E is given here, but could be generalized to any mine, job, and reference year:

$$REC_i = REC_{1998} \times \frac{AdjHP_i / CFM_i}{AdjHP_{1998} / CFM_{1998}},$$

where REC is the concentration in any year i, REC_{1998} is the REC level in 1998, $AdjHP_i$ and CFM_i are the total AdjHP and CFM for that year. The two graphs can then be displayed in the same exposure units, (μg/m^3), experienced by the mine operator job category in Mine E. This is the same approach taken by Crump and colleagues (2015) in their REC 6 model introduced in Chapter 4.

Figure F.10 compares the estimated Mine E operator personal REC exposures from the simple or REC 6 exposure

model to those from the actual DEMS model. The figure shows that the simple model predicts a history of lower exposures than those in the DEMS model, since the simple model essentially assumes that the implicit 'emissions rate' per AdjHP-hr in engines present during the 1998 survey is the same across the full period of study, although it is almost certainly higher in earlier years as might be inferred from Figure F.9.

The ratio of the exposure predicted with the DEMS model to the exposure predicted by the simple model for each year indicates the relative change in "emission rates" for the mine engines per unit of HP before ventilation that is implicitly contained within the DEMS model. This is of interest because it would form a basis of comparison to the specific REC emissions of the engines that were actually used in the mine. Although this information is not within the data sets provided for the engines used in the mines, in principal it could be obtained through cooperative agreements with the engine manufacturers. This would facilitate an alternative approach to assess the fidelity and potential uncertainty of the current model.

When conducting this analysis in Mine E as an illustrative example (Figure F.11), the 1976 engine is predicted to have an HP-specific PM emission that is about 60 percent (1.6 times) greater than that of a comparable HP 1998 engine. This suggests that engine REC emissions per unit of engine HP-hour were higher in earlier years in the DEMS model relative to 1998; this is consistent with the downward adjustment for 1990 and older engines in the model.

How likely is it that the DEMS model underestimated historical REC concentrations in the mines? The 1.6-fold difference in REC emissions that is implicit in the current model is substantially lower than the roughly 8-fold difference in PM emissions per brake-HP-hour suggested in Figure F.9. At face value, this finding might suggest that the DEMS model underestimated historical concentrations of REC in the mines by a similar factor.

However, the U.S. Environmental Protection Agency (U.S. EPA) regulations governing nonroad engines (e.g., construction and other equipment) and the Mine Safety and Health Administration's regulations governing levels of diesel particulate matter (DPM) in metal and nonmetal mines followed quite different schedules. U.S. EPA regulations governing emissions from nonroad engines were first proposed in 2004 to be phased in according to level of HP beginning in 2008, long after the DEMS measurement survey was conducted. Although MSHA proposed a rule governing exposure to DPM in 2001, the rule was subject to multiple challenges and the rule setting a limit of 160 $\mu g/m^3$ total carbon was made final in 2008 (Pomroy and Saseen 2008). As part of the DPM rule, MSHA required that new engines introduced to the mines meet MSHA guidelines or to meet or exceed the U.S. EPA requirements (at the time, Tier 1 and Tier 2 standards had been published).

The DEMS aggregate data on the HP, ventilation levels and average model year of the engines in the mines for a given calendar year show that adoption of newer engines was not immediate in any of the mines (Table F.1). Using 1998 as a reference year, Table F.1 shows a difference of 5 to 10 years between the calendar year and the average model year of the equipment in the mines. Similar differences were evident throughout the years covered by the study.

The simple model developed above, combined with the U.S. EPA (2002) data above for onroad engines, provide a basis for exploring this question. Assuming that PM is a proxy for REC, the regression equation in Figure F.9 provides a means for estimating the change in emissions rate (in g/brakeHP-hr), which can be used to scale changes in REC as measured in 1998 (note that we use 1995 as that is the latest date for which emissions data was available in the U.S. EPA report [2002]). Given the emissions profile provided by Cummins in Figure F.9, we assume the emissions prior to 1975, the earliest date in the U.S. EPA figure, to be constant at 1975 rates (although an alternative assumption could have constant emissions dating back from the late 1980's). We conducted analyses assuming the average model year of engines in the mines directly corresponded to model year in the U.S. EPA (2002) data along with their associated technology and PM emissions.

The results of the analysis using the emissions trends from the U.S. EPA (2002) data are shown graphically in the top line of Figure F.12 and are compared with the DEMS estimates and with the simple (REC 6) model for the same Mine E, mine operator job. They suggest that historical estimates of worker exposure to REC in the mines could have been underestimated by as much as a factor of 2 (taking 1976 as the comparison year). Crump and colleagues (in press) have conducted a similar analysis with slightly different assumptions but overall similar results.

As indicated at the outset of this section, this analysis is considered illustrative of an approach that might be taken to examine the uncertainties associated with fundamental assumptions about emissions trends implicit in the DEMS model. It is an approach that does not rely on assumptions about the relationships between CO and REC over time. We have conducted this analysis using Mine E data as an example and have assumed that the average age of the engines in Mine E corresponds directly to model year and associated engine technology and PM emissions in the data reported in U.S. EPA's (2002) analysis. Mine-specific data exist for individual mines that could be used to explore this question further. Furthermore we realize that the fidelity of

the exposure model depends on the validity of all its assumptions. In Chapter 4 we considered uncertainties in the proportionality of the HP–CO–REC concentration relationships as well. A more complete uncertainty analysis than this illustrative example could explore more systematically and completely the impact of all the key assumptions.

REFERENCES

Borak J, Bunn WB, Chase GR, Hall TA, Head HJ, Hesterberg TW, et al. 2011. Comments on the diesel exhaust in miners study. Ann Occup Hyg 55:339–342.

Clark NN, Jarrett RP, Atkinson CM. 1999. Field measurements of particulate matter emissions, carbon monoxide, and exhaust opacity from heavy-duty diesel vehicles. J Air Waste Manage Assoc 49:76–84.

Crump K, Van Landingham C. 2012. Evaluation of an exposure assessment used in epidemiological studies of diesel exhaust and lung cancer in underground mines. Crit Rev Toxicol 42:599–612.

Crump KS, Van Landingham C, Moolgavkar SH, McClellan R. 2015. Reanalysis of the DEMS nested case-control study of lung cancer and diesel exhaust: suitability for quantitative risk analysis. Risk Anal 35(4):676–700. doi: 10.1111/risa.12371.

Crump KS, Van Landingham C, McClellan R. In Press. Influence of alternative exposure estimates in DEMS miners study: diesel exhaust and lung cancer. Risk Analysis.

Hesterberg TW, Long CM, Bunn WB, Lapin CA, McClellan RO, Valberg PA. 2012a. Health effects research and regulation of diesel exhaust: An historical overview focused on lung cancer risk. Inhal Toxicol 24 Suppl 1:1–45.

McClellan RO, Hesterberg TW, Wall JC. 2012. Evaluation of carcinogenic hazard of diesel engine exhaust needs to consider revolutionary changes in diesel technology. Regul Toxicol Pharmacol 63:225–258.

Pomroy WH, Saseen GP. 2008. MSHA's DPM rule for metal and nonmetal mines. In: 12th U.S./North American Mine Ventilation Symposium 2008 (Wallace KG ed). Clovis, CA: Mine Ventilation Services, 455–462. Available: *www.mvsengineering.com/index.php/products/97*.

Tomko DM. 2013. Chief, Environmental Assessment & Contaminate Control Branch, Mine Safety & Health Administration, Pittsburgh Safety & Health Technology Center, Technical Support / Dust Division / Field GroupDiesel "*Particulate Matter Control Strategies,*" PowerPoint presentation sent to the panel via e-mail 9/19/2013 (slides 6 and 7). Available: *www.msha.gov/01-995/Powerpoints/Diesel%20Particulate%20Matter%20(Louisville).pdf)*.

U.S. Environmental Protection Agency. 2002. Health assessment document for diesel engine exhaust. EPA/600/8-90/057F. Washington, DC:National Center for Environmental Assessment.

Vermeulen R. 2014. Adapted from "Estimating Historical Exposures to Diesel Exhaust in Underground Non-metal Mining Facilities" presented by Professor Roel Vermeulen at the HEI "Diesel Exhaust, Lung Cancer and Quantitative Risk Assessment", March 6, 2014, Boston MA.

Vermeulen R, Coble JB, Lubin JH, Portengen L, Blair A, Attfield MD, et al. 2010a. The Diesel Exhaust in Miners Study: IV. Estimating historical exposures to diesel exhaust in underground non-metal mining facilities. Ann Occup Hyg 54:774–788.

Vermeulen R, Coble JB, Yereb D, Lubin JH, Blair A, Portengen L, et al. 2010b. The Diesel Exhaust in Miners Study: III. Interrelations between respirable elemental carbon and gaseous and particulate components of diesel exhaust derived from area sampling in underground nonmetal mining facilities. Ann Occup Hyg 54:762–773.

Xu S, Clark NN, Gautam M, Wayne WS. 2005. Comparison of heavy-duty truck diesel particulate matter measurement: TEOM and traditional filter. SAE Technical Paper 2005-01-2153. doi:10.4271/2005-01-2153.

Yanowitz J, McCormick RL, Graboski MS. 2000. In-use emissions from heavy-duty diesel vehicles. Environ Sci Technol 34:729–740.

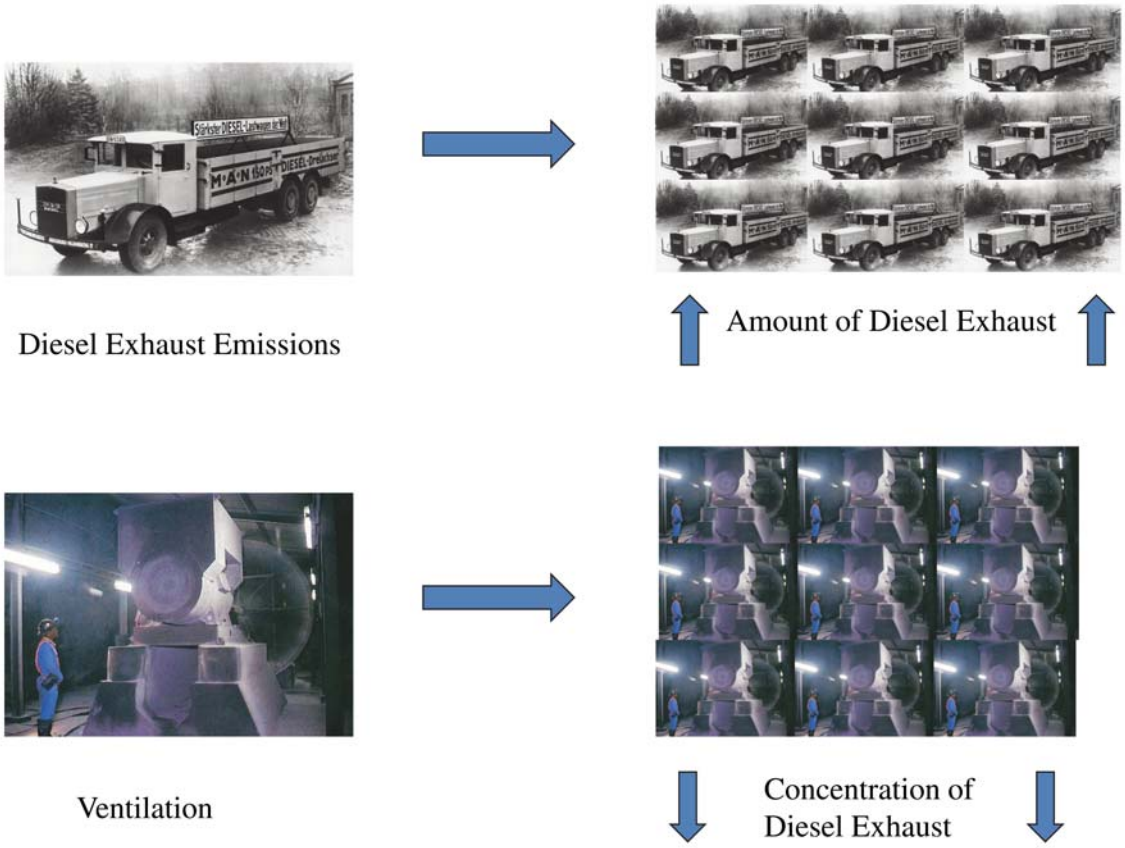

Figure F.1. General overview of the logic of the DEMS retrospective exposure model. (Modified from a presentation by Roel Vermeulen at the March 2014 HEI Workshop on Diesel Epidemiology.)

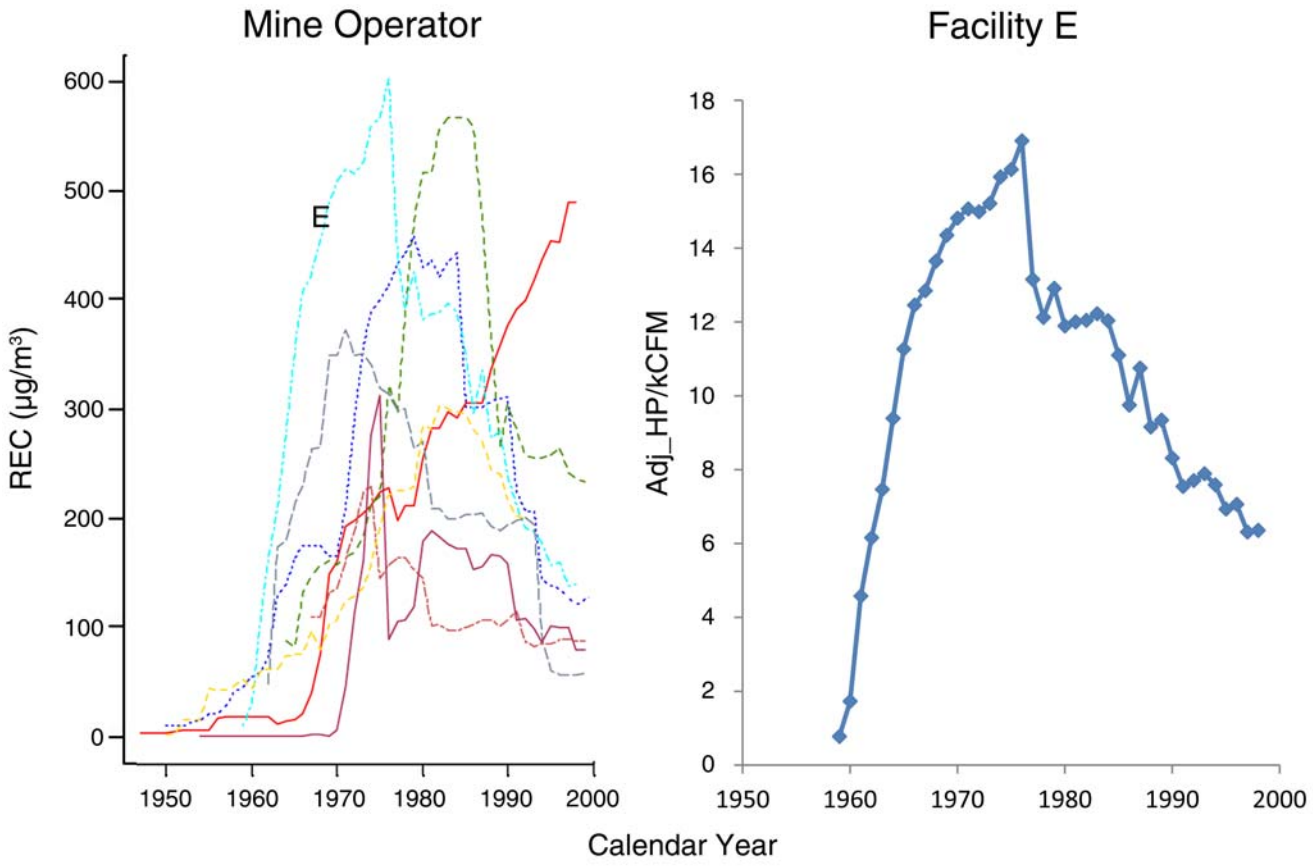

Figure F.2. REC historical predications (µg/m³) for the mine operator, based on the primary facility-specific CO models, by mining facility (left) (source: Vermeulen et al. 2010a) compared to plot of AdjHP/CFM for Mine E (right).

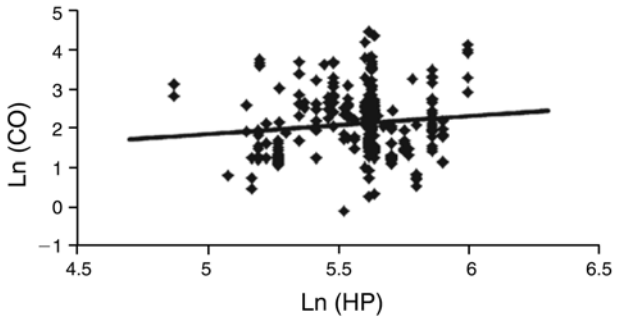

Figure F.3. Scatterplot of LnCO) and Ln(HP) emissions (g/mile) from chassis dynamometer tests of heavy-duty diesel vehicles. Based on data from more than 250 different vehicles, model years 1981–1997, reported in 20 different studies in a review by Yanowitz et al. 2000. (Adapted from Crump and Van Landingham 2012, Figure 2.)

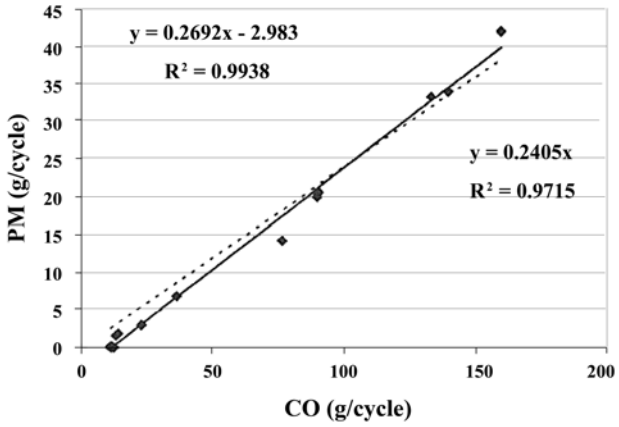

Figure F.4. Relationship between CO and PM for a 2004 Freightliner truck powered by a Detroit Diesel DDC Series 60 engine. The engine was operated within Idle, UDDS, Creep, Transient, Cruise, and HHDDT_S cycles at 30,000 lb, 56,000 lb, and 66,000 lb test weight. (Reprinted, with permission, from Xu et al. 2005 [SAE paper 2005-01-2153, available at *www.sae.org*].)

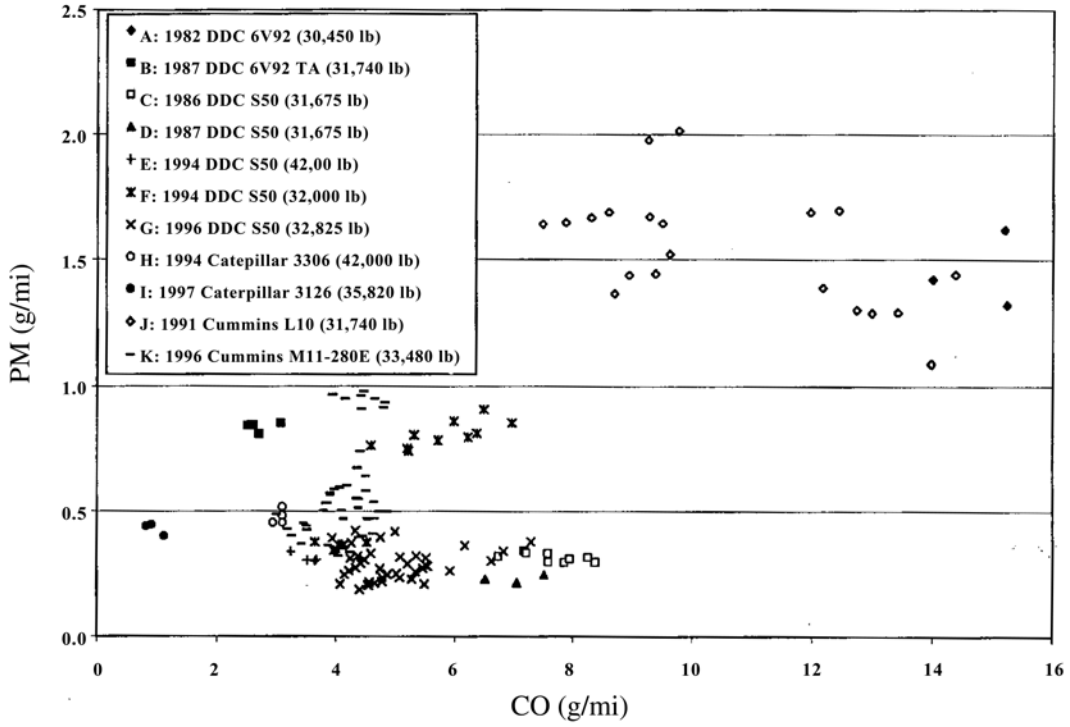

Figure F.5. Relationship between cycle emissions of CO and PM from different transit buses driven through the CBD cycle. (Reprinted from Clark et al. 1999 by permission of the Air & Waste Management Association.)

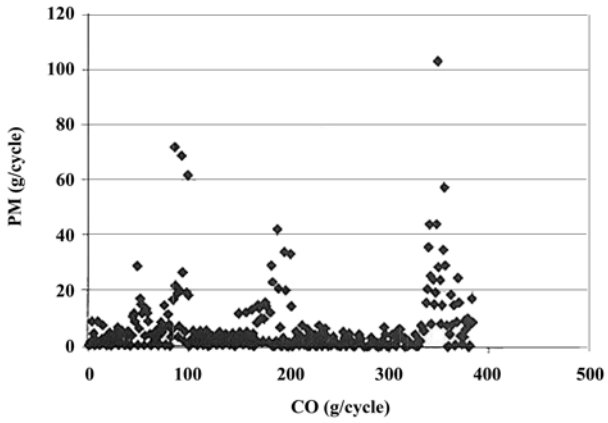

Figure F.6. Plot of CO and PM emissions per cycle for different engines, operated over different duty cycles. (Reprinted, with permission, from Xu et al. 2005 [SAE paper 2005-01-2153, available at *www.sae.org*]. Copyright © 2005 SAE International. Further use or distribution is not permitted without permission from SAE.)

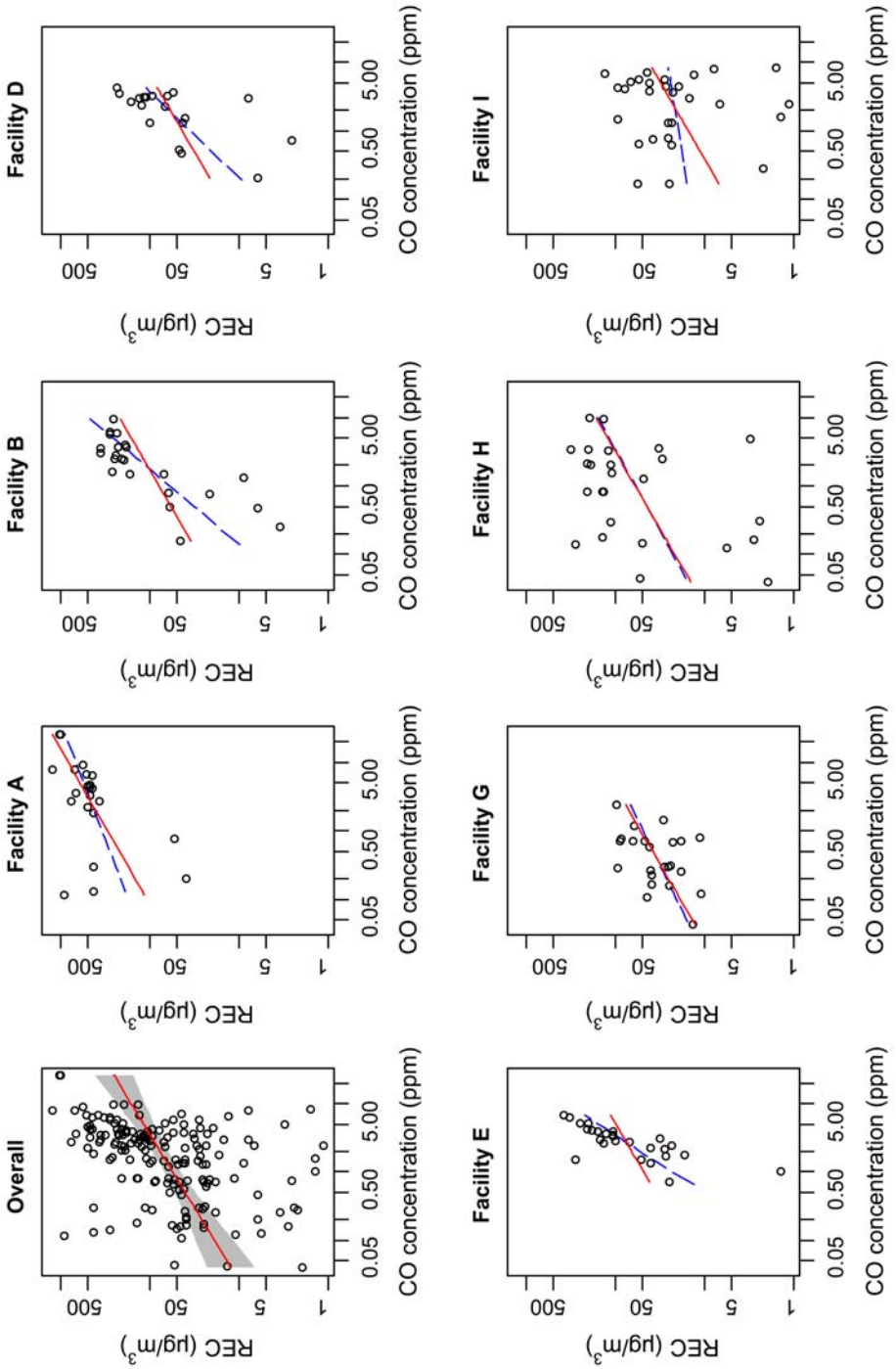

Figure F.7. Regression analyses between natural log-transformed (Ln) CO and REC measurements for all mining facilities combined (Overall) and by facility (A-I). The solid line represents a fixed-effects model allowing for facility-specific intercepts and a common slope, and the dashed line represents one mixed-effects model with fixed facility-specific intercepts and facility-specific random slopes. The shaded area in the overall plot represents the 95% CI. (Reproduced from Vermeulen et al. 2010b, Figure 2.)

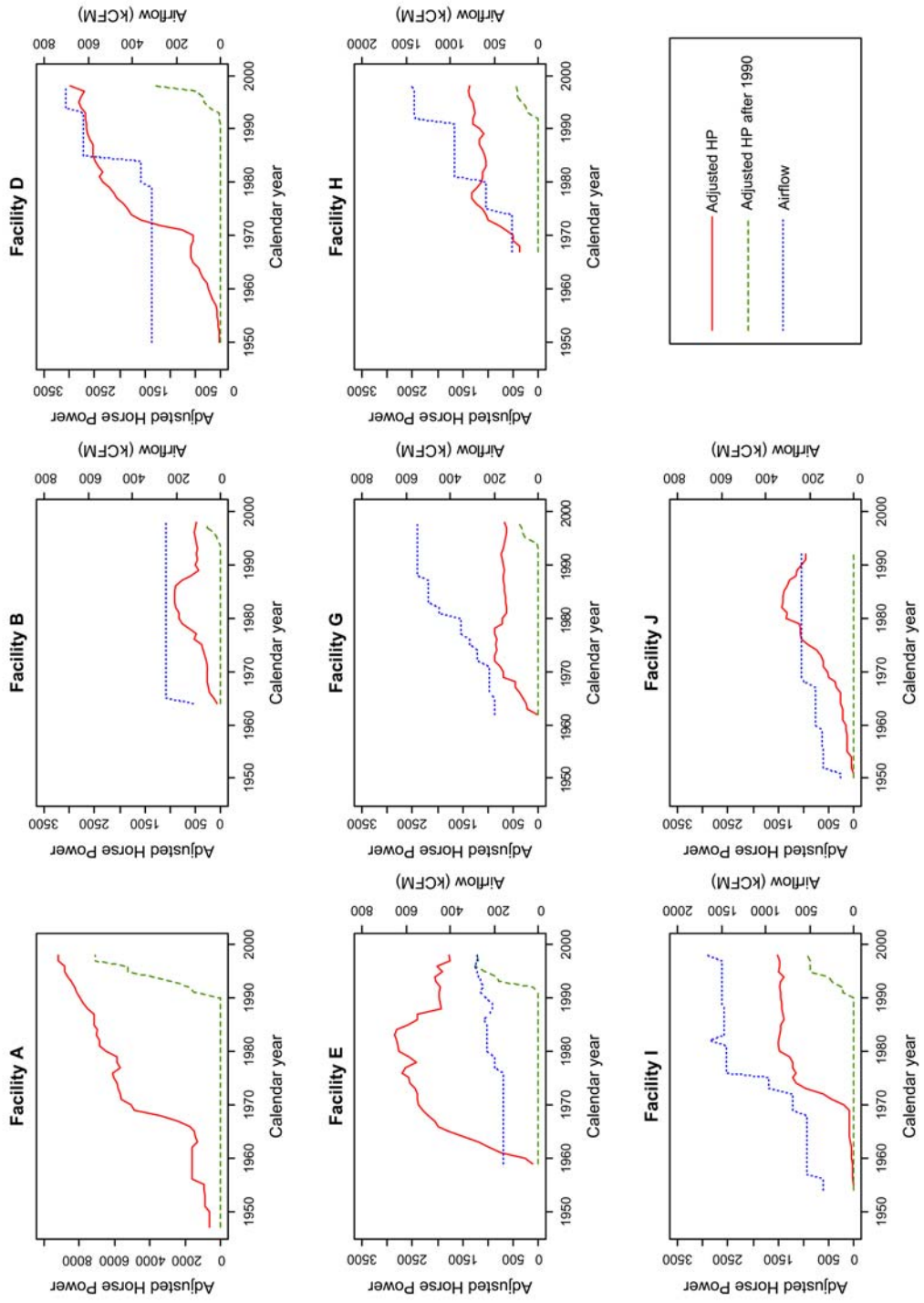

Figure F.8. AdjHP, AdjHP after 1990, and total airflow exhaust rates for each mine. (Reproduced from Vermeulen et al. 2010a, Figure 1).

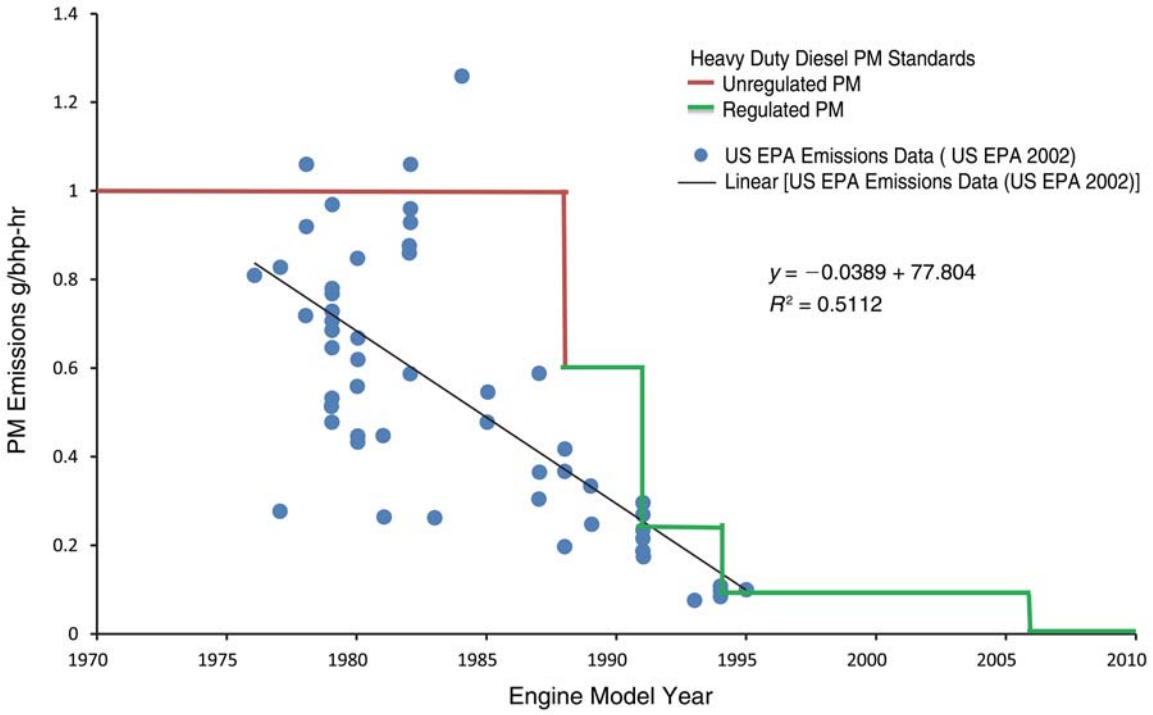

Figure F.9. PM emissions from onroad diesel engine certification data in grams per brake horsepower-hour (g/bhp-hr) as a function of model year compared with the timeline for heavy-duty diesel emissions standards. The black line is the regression line (slope: −0.0389 g/bhp-hr per year). (Reproduced from U.S. EPA 2002, Figure 2-20.) The regulatory timeline for emissions is from data provided by Cummins (personal communication).

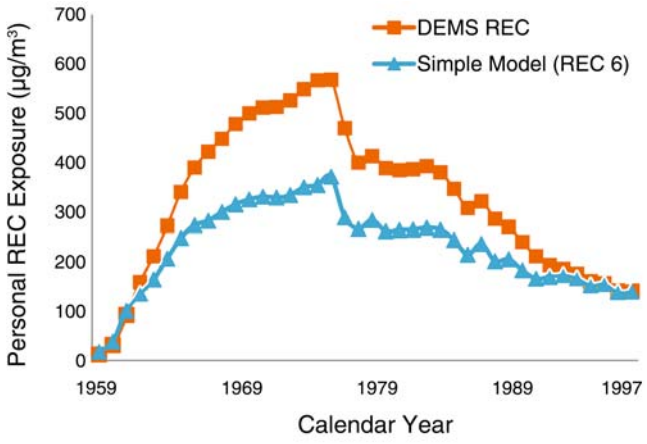

Figure F.10. Comparison of personal REC exposures to the Mine E operator using predictions from the DEMS model (top line) and from the simple model (or REC 6) (bottom line).

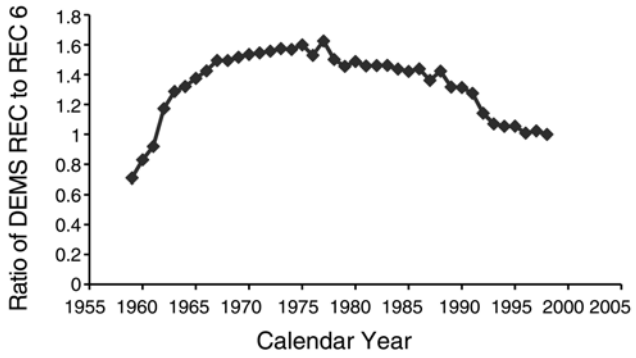

Figure F.11. Ratio of DEMS REC to simple model (REC 6) estimates from the Mine E operator: A comparison of relative emissions rates.

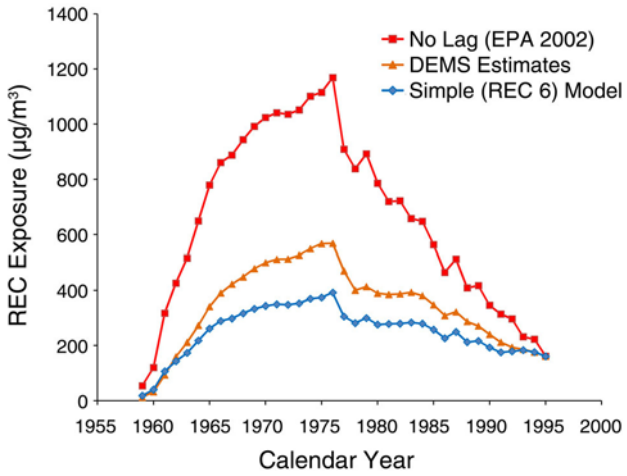

Figure F.12. Illustrative analysis incorporating relative trends in emissions from onroad engines into simple model (REC 6) exposure estimates for the Mine E operator.

Table F.1 Average Model Year of Equipment in the Mines in 1998

Mine	Average Model Year in 1998[a]
A	1993
B	1992
D	1987
E	1992
G	1993
H	1991
I	1989
J	1983[b]

[a] Source: NCI/NIOSH exposure data (*http://dceg.cancer.gov/research/what-we-study/environment/diesel-exhaust-miners-study-dems*).

[b] The last date of operation for Mine J was 1993. 1983 is the average model year of the equipment at that time.

Additional Materials Available on the Web

Additional Materials 1 through 3 contain supplemental material not included in the printed report. They are available on the HEI Web site at *http://pubs.healtheffects.org*.

Additional Materials 1. Graphs of Cox Regression Analyses with Penalized Splines from Garshick et al. 2012.

Additional Materials 2. Analytical Data Sets for the Cohort (Attfield et al. 2012) and Case–Control (Silverman et al. 2012) Studies, and HEI Diesel Epidemiology Panel Replication of Selected Analyses in the DEMS Case–Control Study.

Additional Materials 3. HEI Diesel Epidemiology Panel's Additional Analyses of Adjustment for Smoking.

Abbreviations and Other Terms

AIC	Akaike Information Criteria
ACES	Advanced Collaborative Emissions Study
AdjHP	adjusted horsepower
ATS	American Thoracic Society
BoM	Bureau of Mines
CFM	cubic feet per minute
CI	confidence interval
CO	carbon monoxide
CO_2	carbon dioxide
CPH	Cox proportional hazards
DEMS	Diesel Exhaust in Miners Study
DPM	diesel particulate matter
EC	elemental carbon
EMA	Engine Manufacturers Association
HP	horsepower
HR	hazard ratio
IARC	International Agency for Research on Cancer
ICCT	International Council on Clean Transportation
IIASA	International Institute for Applied Systems Analysis
IRIS	Integrated Risk Information System
LC	lung cancer
LOD	limits of detection
MATES	Multiple Air Toxics Exposure Study
MESA	Mine Safety Enforcement Administration
MIDAS	mine information data system
MSHA	Mine Safety and Health Administration
NCI	National Cancer Institute
NDI	National Death Index
NIOSH	National Institute for Occupational Safety and Health
NO	nitric oxide
NO_2	nitrogen dioxide
NO_x	oxides of nitrogen
NRC	National Research Council
OR	odds ratio
PAH	polycyclic aromatic hydrocarbon
pCi	picocuries
PI	particulate index
PM	particulate matter
$PM_{1.0}$	particulate matter ≤ 1 µm in aerodynamic diameter
$PM_{2.5}$	particulate matter ≤ 2.5 µm in aerodynamic diameter
REC	respirable elemental carbon
RELtrend	factor adjusting for changes in annual CO
SE	standard error
SEC	submicron elemental carbon
SMR	standardized mortality ratio
STROBE	strengthening the reporting of observational studies in epidemiology
SCAQMD	South Coast Air Quality Management District
TEC	total elemental carbon
TSCE	three-stage clonal expansion
U.S. EPA	U.S. Environmental Protection Agency
WHO	World Health Organization
WL	working level
WLM	working level month

HEI BOARD, COMMITTEES, and STAFF

Board of Directors

Richard F. Celeste, Chair *President Emeritus, Colorado College*

Sherwood Boehlert *Of Counsel, Accord Group; Former Chair, U.S. House of Representatives Science Committee*

Enriqueta Bond *President Emerita, Burroughs Wellcome Fund*

Purnell W. Choppin *President Emeritus, Howard Hughes Medical Institute*

Michael T. Clegg *Professor of Biological Sciences, University of California–Irvine*

Jared L. Cohon *President Emeritus and Professor, Civil and Environmental Engineering and Engineering and Public Policy, Carnegie Mellon University*

Stephen Corman *President, Corman Enterprises*

Linda Rosenstock *Dean Emerita and Professor of Health Policy and Management, Environmental Health Sciences and Medicine, University of California–Los Angeles*

Henry Schacht *Managing Director, Warburg Pincus; Former Chairman and Chief Executive Officer, Lucent Technologies*

Warren M. Washington *Senior Scientist, National Center for Atmospheric Research; Former Chair, National Science Board*

Health Research Committee

David L. Eaton, Chair *Dean and Vice Provost of the Graduate School, University of Washington–Seattle*

Jeffrey R. Brook *Senior Research Scientist, Air Quality Research Division, Environment Canada, and Assistant Professor, University of Toronto, Canada*

David Christiani *Elkan Blout Professor of Environmental Genetics, Harvard T.H. Chan School of Public Health*

Francesca Dominici *Professor of Biostatistics and Senior Associate Dean for Research, Harvard T.H. Chan School of Public Health*

David E. Foster *Phil and Jean Myers Professor Emeritus, Department of Mechanical Engineering, Engine Research Center, University of Wisconsin–Madison*

Uwe Heinrich *Professor, Hannover Medical School; Executive Director, Fraunhofer Institute for Toxicology and Experimental Medicine, Hanover, Germany*

Barbara Hoffmann *Professor of Environmental Epidemiology and Head of Environmental Epidemiology of Aging, IUF-Leibniz Research Institute for Environmental Medicine, and Professor, Medical Faculty, Heinrich Heine University of Düsseldorf, Germany*

Allen L. Robinson *Raymond J. Lane Distinguished Professor and Head, Department of Mechanical Engineering, and Professor, Department of Engineering and Public Policy, Carnegie Mellon University*

Ivan Rusyn *Professor, Department of Veterinary Integrative Biosciences, Texas A&M University*

Richard L. Smith *Director, Statistical and Applied Mathematical Sciences Institute, University of North Carolina–Chapel Hill*

HEI BOARD, COMMITTEES, and STAFF

Health Review Committee

James A. Merchant, Chair *Professor and Founding Dean, College of Public Health, University of Iowa*

Michael Brauer *Professor, School of Environmental Health, University of British Columbia, Canada*

Bert Brunekreef *Professor of Environmental Epidemiology, Institute of Risk Assessment Sciences, University of Utrecht, the Netherlands*

Mark W. Frampton *Professor of Medicine and Environmental Medicine, University of Rochester Medical Center*

Roger D. Peng *Associate Professor, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health*

Lianne Sheppard *Professor of Biostatistics, School of Public Health, University of Washington–Seattle*

Officers and Staff

Daniel S. Greenbaum *President*

Robert M. O'Keefe *Vice President*

Rashid Shaikh *Director of Science*

Jacqueline C. Rutledge *Director of Finance and Administration*

April Rieger *Corporate Secretary*

Zachary Abbott *Research Assistant*

Sarah Benckart *Science Administration Assistant*

Hanna Boogaard *Staff Scientist*

Aaron J. Cohen *Principal Scientist*

Maria G. Costantini *Principal Scientist*

Philip J. DeMarco *Compliance Manager*

Hope Green *Publications Associate*

Anny Luu *Executive Assistant*

Hilary Selby Polk *Managing Editor*

Evan Rosenberg *Staff Accountant*

Robert A. Shavers *Operations Manager*

Geoffrey H. Sunshine *Senior Scientist*

Annemoon M.M. van Erp *Managing Scientist*

Donna J. Vorhees *Senior Scientist*

Katherine Walker *Senior Scientist*

HEALTH
EFFECTS
INSTITUTE

101 Federal Street, Suite 500
Boston, MA 02110, USA
+1-617-488-2300
www.healtheffects.org

SPECIAL
REPORT
19

November 2015