Diesel Emissions and Lung Cancer:

Epidemiology and Quantitative Risk Assessment

A Special Report of the Institute's Diesel Epidemiology Expert Panel

Health Effects Institute

June 1999

HEI HEALTH EFFECTS INSTITUTE

The Health Effects Institute, established in 1980, is an independent and unbiased source of information on the health effects of motor vehicle emissions. HEI supports research on all major pollutants, including regulated pollutants (such as carbon monoxide, ozone, nitrogen dioxide, and particulate matter), and unregulated pollutants (such as diesel engine exhaust, methanol, and aldehydes). To date, HEI has supported more than 200 projects at institutions in North America and Europe and published over 100 Research Reports. Consistent with its mission to serve as an independent source of information on the health effects of motor vehicle pollutants, the Institute also engages in special review and evaluation activities.

Typically, HEI receives half its funds from the U.S. Environmental Protection Agency and half from 28 manufacturers and marketers of motor vehicles and engines in the United States. Occasionally, funds from other public or private organizations either support special projects or provide resources for a portion of an HEI study. Regardless of funding sources, HEI exercises complete autonomy in setting its research priorities and in reaching its conclusions. An independent Board of Directors governs HEI. The Institute's Research and Review Committees serve complementary scientific purposes and draw distinguished scientists as members. The results of HEI-funded research and evaluations have been used in public and private decision-making.

Statement from the HEI Board of Directors

Diesel engine technology has played a valuable role in the transportation industry worldwide since the Second World War. Diesel-powered engines are more efficient than gasoline-powered engines, and as a result they emit less carbon dioxide, which is a greenhouse gas. However, diesel engine emissions contain more oxides of nitrogen, which are ozone precursors, and particulate matter than gasoline engine emissions. The impact on human health of inhaling particulate matter, especially in the development of lung cancer, has been a matter of scientific concern and research.

For more than a decade, state, national, and international agencies have spent considerable effort to examine the resulting scientific literature to determine whether diesel engine emissions, and more specifically, diesel particulate matter, is carcinogenic in humans. The majority of the effort has been focused on lung cancer. Regulatory agencies that have a mandate to protect the public's health have examined the biologic, toxicologic, and epidemiologic data to identify the possible carcinogenic hazard from exposure to diesel exhaust and to develop quantitative risk estimates for diesel exhaust exposure and lung cancer. These risk estimates have been controversial.

The Health Effects Institute has supported research on the health effects of exposure to diesel emissions for many years. In the early 1980s, this research, via studies of mutagenicity, metabolism, and carcinogenesis, focused on the organic constituents of diesel exhaust particles. In the late 1980s, a major study compared the carcinogenicity of diesel particles and carbon black in rats, and found very similar results: The carbonaceous particles, not the organic chemicals adsorbed onto the particles, were responsible for the tumor response in rats, which was most likely due to particles overloading the lung clearance mechanism.

In 1995, HEI published a Special Report, *Diesel Exhaust: A Critical Analysis of Emissions, Exposure, and Health Effects*, in which the HEI Diesel Working Group raised questions about the use of animal data for quantitative risk assessment. With respect to epidemiologic studies, the Working Group noted those studies showed consistent, small increases in risk for exposed workers, but concluded that the absence of concurrent exposure measurements limited the utility of those studies for quantitative risk assessment.

As part of the Institute's continued interest in diesel engine emissions and health effects, and in an effort to understand how current research or possible new research could be useful for quantitative risk assessment, HEI initiated the Diesel Epidemiology Project in 1998. This multifaceted effort includes the work of the Diesel Epidemiology Expert Panel (contained in this report), a set of feasibility studies to inform the direction of new research, and the Diesel Workshop: Building a Research Strategy to Improve Risk Assessment, held in March of 1999.

We appointed seven scientific experts in biostatistics, epidemiology, exposure assessment, and exposure characterization as the Diesel Epidemiology Expert Panel to examine specifically the strengths and limitations of the published epidemiologic studies currently available for quantitative risk estimation, and to evaluate discrepancies in the exposure-response findings reported. We have reviewed the report, and the process the Panel used in preparing it, and believe it presents a systematic and fair evaluation of the available epidemiologic studies and the associated quantitative data. The Panel also has identified through its analysis a reasonable explanation for conflicting exposure-response findings previously reported.

We expect that the Panel's review, analyses, and recommendations will be used by HEI and others to address research priorities for further quantitative estimation of the relation between exposure to diesel emissions and lung cancer risk. We also note the need to consider the relation between diesel emissions and noncancer health effects. As the use of diesel fuels around the world expands to take advantage of fuel economy and reduced greenhouse gas emissions, the full range of health effects needs to be considered and weighed along with potential climate change benefits.

In addition to thanking the entire HEI Diesel Epidemiology Expert Panel, we would particularly like to thank Dr. John C. Bailar III, Chair of the Panel, and Dr. Diane J. Mundt, who served as HEI's scientific project manager.

Diesel Emissions and Lung Cancer: Epidemiology and Quantitative Risk Assessment

A Special Report of the Institute's Diesel Epidemiology Expert Panel

Statement from the HEI Board of Directors i

Contributors to the Project v

Executive Summary 1

Background 9

The HEI Diesel Epidemiology Project 13

Summary of Railroad Worker and Teamster Studies 17

Assessment of Railroad Worker and Teamster Studies 25

Findings and Recommendations for Future Research 39

Appendix A. Diesel Epidemiology Expert Panel Workshop 45

Appendix B. Framework for Systematically Evaluating Epidemiologic Studies for Quantitative Risk Assessment 47

Appendix C. The Panel's Exploration of Railroad Worker Data 55

Appendix D. Potential Impact of Control Selection Bias 61

Literature Cited 65

Abbreviations 69

Related HEI Publications: Diesel Engine Exhaust 71

Contributors to the Project

DIESEL EPIDEMIOLOGY EXPERT PANEL

John C. Bailar, III (Chair)
Professor, Department of Health Studies, The University of Chicago and the HEI Health Review Committee

Ethel Gilbert
Biostatistician, Radiation Epidemiology Branch, National Cancer Institute

David Hoel
Professor, Department of Biometry and Epidemiology, Medical University of South Carolina

Brian Leaderer
Professor and Head, Division of Environmental Health Sciences, Department of Epidemiology and Public Health, Yale University School of Medicine, and Fellow, John B. Pierce Laboratory

Charles Poole
Associate Professor, Department of Epidemiology, School of Public Health, University of North Carolina

Robert F. Sawyer
Professor of the Graduate School and Class of 1935 Professor of Energy (Emeritus), Department of Mechanical Engineering, University of California, Berkeley, and the HEI Health Research Committee

G. Marie Swanson
Professor of Family Practice and Director of the Cancer Center, Cancer Center and Department of Family Practice, Michigan State University

HEI PROJECT STAFF

Diane J. Mundt
Project Manager

Kathleen M. Nauss
Director for Scientific Review and Evaluation

Aaron J. Cohen
Senior Scientist

Daniel S. Greenbaum
President

Mary Stilwell
Senior Administrative Assistant

HEI PUBLICATION STAFF

Thomas Atwood
Manager of Publications and Information

Julia F. Campeti
Publications Assistant

L. Virgi Hepner
Senior Scientific Editor

Francine Marmenout
Senior Executive Assistant

Susan Shephard
Consulting Scientific Editor

Hope Steele
Acting Managing Editor

REVIEWERS

This report was submitted for outside peer review. The Health Effects Institute appreciates the thoughtful critiques provided by the reviewers listed below. The views expressed in this report are those of the Diesel Epidemiology Expert Panel and no endorsement by the external reviewers should be inferred.

The Panel extends its gratitude to Aaron Blair, William Bunn, Glen Cass, James Cogliano, Kenny Crump, Stanley Dawson, Eric Garshick, Dale Hattis, Daniel Krewski, Philip Lorang, Frank Mirer, Suresh Moolgavkar, Robert O'Keefe, Jonathan Samet, Michael Spallek, Kyle Steenland, Duncan Thomas, Gerald van Belle, Jane Warren, and Winthrop Watts.

Executive Summary

Diesel engines are an important part of the world's transportation and industrial infrastructure, especially in heavy-duty applications such as trucks, buses, construction and farm equipment, locomotives, and ships. Energy efficiency and durability account for the dominant use of diesel engines worldwide, and their use may expand in the future. In Europe, 20% to 50% of the new light-duty passenger fleet is powered by diesel engines. Although the percentage of diesel-powered light-duty vehicles is much lower in the United States, advanced technology diesel engines are being proposed as part of the nation's energy conservation and climate change strategies.

The economic advantages of diesel engines are clear; nevertheless, environmental concerns and related health issues must be addressed. Emissions from all types of engines are highly variable and complex mixtures. Diesel engines are more efficient than gasoline engines, and they emit less carbon dioxide (a greenhouse gas), carbon monoxide, and hydrocarbons. Therefore, diesel engines have some advantages over conventional gasoline engines in terms of global warming. However, they emit higher levels of oxides of nitrogen, which are ozone precursors, and particulate matter per vehicle mile traveled than do gasoline engines. The particulates are of special concern in possible health outcomes; they are small enough to be readily respirable, and they have many chemicals adsorbed to their surfaces, including known or suspected mutagens and carcinogens.

Cellular, animal, and human studies have investigated the association between exposure to diesel exhaust and adverse health effects, including cancer. Lung tumors have occurred in rats exposed to diesel exhaust, but the relevance of these lesions to human risk assessment has been questioned. Epidemiologic studies fairly consistently show an elevation in lung cancer rates among occupationally exposed individuals. In most studies, rates are 20% to 50% greater than those in unexposed individuals; however, these studies did not obtain quantitative measurements of exposure during the time period of the study.

Although epidemiologic data have been used generally to identify the hazards associated with exposure to diesel exhaust, questions remain as to whether the human data can be used to develop reliable estimates of the magnitude of any risk for lung cancer (that is, through quantitative risk assessment [QRA]), and whether new research efforts could provide any additional data needed. In response to such issues, the Health Effects Institute initiated the Diesel Epidemiology Project in 1998. The Project includes the evaluation by HEI's Diesel Epidemiology Expert Panel of occupational epidemiologic studies that have been used for QRA, and the development of new research initiatives to improve understanding about the health effects of diesel exhaust.

The Diesel Epidemiology Expert Panel was chaired by John C. Bailar, III, M.D., Ph.D., of The University of Chicago and the HEI Review Committee, and included six other scientists who have expertise in epidemiology, biostatistics, exposure characterization, and exposure assessment. It was charged to (1) review the epidemiologic data that form the basis of current QRAs for diesel exhaust, (2) identify data gaps and sources of uncertainty, (3) make recommendations about the usefulness of extending or conducting further analyses of existing data sets, and (4) make recommendations for the design of new studies that would provide a stronger basis for risk assessment.

Although lung cancer was the health outcome of interest to the Panel's charge, it was not charged to evaluate either the broad toxicologic or epidemiologic literature concerning exposure to diesel exhaust and lung cancer for hazard identification purposes, which has been done by others. State, national, and international agencies have all reviewed the broader animal and human evidence for carcinogenicity and, in either their draft or final reports, have all identified diesel exhaust as a probable human carcinogen or placed it in a comparable category (National Institute for Occupational Safety and Health 1988; International Agency for Research on Cancer 1989; World Health Organization 1996; National Toxicology Program 1998; Office of Environmental Health Hazard Assessment [California Environmental Protection Agency] 1998; U.S. Environmental Protection Agency 1998).

In response to the first charge, the Panel examined published epidemiologic studies of diesel exhaust emissions and lung cancer for possible use in support of QRA. Only two such studies reported any quantitative exposure data associated in some manner with the occupational epide-

miologic studies, and they were considered in the Panel's review.

The Panel recognized that no epidemiologic study can be perfect. Therefore, the Panel viewed its task as addressing the question: To what extent can limitations in the design and performance of a particular study affect its contribution to the body of epidemiologic knowledge under examination for QRA? The Panel also recognized that frequently it is very difficult to obtain retrospective data for estimating job-related work exposures, and that this process may require assumptions that cannot be validated. In the studies considered here, which form the core of the Panel's review, reasonable attempts were made to reconstruct past exposures to diesel engine emissions using approaches that were feasible when the studies were conducted. These data subsequently have been used, in some cases, for purposes that were not envisioned by the original investigators. The studies reviewed for this report include:

Railroad Worker Studies

- Case-control: Garshick et al. 1987
- Cohort: Garshick et al. 1988
- Industrial hygiene: Hammond 1988, and Woskie et al. 1988a,b
- Exposure-response analyses: Crump et al. 1991, Office of Environmental Health Hazard Assessment 1998, and Crump 1999

Teamster Studies

- Case-control: Steenland et al. 1990, 1992
- Industrial hygiene: Zaebst et al. 1991
- Exposure-response analysis: Steenland et al. 1998

The reports of these studies were supplemented by published articles and by presentations to the Panel by the principal investigators and others, including secondary analysts of the railroad worker data. The Panel did not consider other completed lung cancer and diesel epidemiologic studies because they included no directly associated quantitative exposure data.

Certain strengths are evident in the studies reviewed by the Panel. The epidemiologic studies include large numbers of study subjects (55,407 subjects, and 1,694 lung cancers, for the railroad worker cohort study; 1,256 deaths from lung cancer for the railroad worker case-control study; and 996 deaths from lung cancer for the teamster case-control study), all of whom were employed in industries where many workers are exposed to diesel exhaust. Job categories with known exposure to asbestos were either excluded or controlled for in the analyses. Both of the case-control studies adjusted data analyses to control for cigarette smoking as a confounding variable. Overall, the results are generally consistent with findings of a weak association between lung cancer and exposure to diesel exhaust. However, published secondary analyses of exposure-response relations in the railroad worker cohort data produced conflicting results (Crump et al. 1991; Office of Environmental Health Hazard Assessment 1998).

Measurements from the industrial hygiene studies in general supported the job exposure categories used in the epidemiologic studies. The industrial hygiene studies measured different markers for diesel exhaust exposure—respirable-sized particles (RSP) for railroad workers and submicron-sized elemental carbon (EC_1) for teamsters. Although the RSP measures were adjusted for the environmental tobacco smoke component, EC_1 is more sensitive and specific to diesel exhaust than adjusted RSP.

In response to the second charge, the Panel developed a framework of general epidemiologic questions about study design, exposure assessment, outcome determination, and analysis. These are meant to help in systematically understanding and revealing the strengths and uncertainties of these studies. This framework was then used to evaluate the studies of railroad workers and teamsters. This process helped to address the third and fourth charges to the Panel, and to assist HEI in focusing its future research directions to inform apparent gaps for QRA.

The original findings of the cohort railroad worker study reported by Garshick and coworkers (1988) indicated a steadily increasing risk of lung cancer for exposed workers with increasing years of employment. This increase with duration of employment, however, was not supported in later, unpublished analyses (Garshick 1991). This increasing risk, plus the availability of some quantitative exposure data in railroad workers (Woskie et al. 1988a,b), prompted additional analyses to explore the exposure-response relation in these data (Crump et al. 1991; Office of Environmental Health Hazard Assessment 1994, 1998; Crump 1999). Crump and colleagues found a negative association between lung cancer risk and several measures of cumulative exposure; that is, risk decreased with increasing cumulative exposure. In contrast, the statistical models used by the Office of Environmental Health Hazard Assessment analysts, using the same data but different assumptions, showed a positive association in which risk increased with increasing cumulative exposure.

The Panel explored these apparent inconsistencies in the exposure-response relation to verify and obtain a better understanding of the previous analyses, and to help

clarify differences. These issues are central to whether the railroad worker data can be useful in a QRA for lung cancer.

The Panel's data exploration demonstrated that within the three broad railroad job categories of train workers (e.g., engineers, conductors), shop workers (e.g., electricians, machinists), and clerks and signalmen, the relative risk of lung cancer decreased with increasing duration of employment, and this decrease was statistically significant for the clerks/signalmen and train workers. Although the relative risk decreased with increasing duration of employment, overall risks for train workers, within each duration of employment group, were higher than those for clerks and signalmen, and shop workers had intermediate risks (Figure 1).

These findings are not consistent with a steadily increasing association between cumulative diesel exposure and lung cancer risk. Furthermore, if the difference in risk between train workers and clerks/signalmen was due primarily to differences in exposure to diesel emissions, one would expect the relative risk for train workers compared with that for clerks and signalmen to be reduced or even eliminated after adjusting for exposure. In fact, adjustment for exposure increased this relative risk. Such a systematic pattern of decreasing risk with increasing exposure suggests that some form of bias is present in the data, which makes it difficult to determine the true nature of an exposure-response relation. Bias can result from uncontrolled confounding by cigarette smoking or by other occupational exposure, differential misclassification of exposures by job category, longer survival of "healthier" workers, or differential ascertainment of lung cancer as a cause of death.

Initial findings from the teamster case-control study (Steenland et al. 1990) showed an increased risk of lung cancer with increasing years of employment. The investigators published an exposure-response analysis for the teamster study (Steenland et al. 1998) after the Panel's work started, thus the evaluation of this set of studies was necessarily less extensive.

Reconstructing past exposures for which actual data are limited or nonexistent requires several assumptions. The Panel had concerns about several of the assumptions used by Steenland and colleagues in the exposure-response analysis of the teamster data. These concerns include (1) the data on 1990 emissions used to estimate past exposures to diesel exhaust may underestimate average exposures over a range of work histories, given that more recent data show higher emissions for that time; (2) the date assumed for dieselization in the trucking industry, which, if too early, may overestimate exposures; (3) the degree to which vehicle miles traveled accurately reflects actual exposure to diesel exhaust for various job groups, which may affect exposure estimates in either direction; (4) the possible effects of using various scenarios of emission levels to account for long fleet turnover times in the trucking industry; and (5) the difficulty in distinguishing truck driver exposures from background levels, because measured estimates are close. Also, among the assumptions Steenland and colleagues used, nondiesel sources of elemental carbon in ambient air, especially from gasoline engine emissions, were not considered.

The Panel also was concerned about the controls used in the case-control study. Lung and bladder cancers and motor vehicle accidents were excluded as control causes of death, and controls were selected from other causes. If those causes of death were associated with exposure to diesel emissions, smoking, or both, the study findings could be biased.

Important work is currently under way to study the health effects of exposure to diesel exhaust in nonmetal miners in Germany (Säverin et al. 1998) and in the United States (National Cancer Institute–National Institute for Occupational Safety and Health 1997). The Panel did not review these studies because they are still in progress. However, the Panel heard presentations from these investigators at the HEI Diesel Workshop: Building a Research Strategy to Improve Risk Assessment (HEI 1999) at Stone Mountain, GA, March 7–9, 1999. In particular, the National Cancer Institute–National Institute for Occupational Safety and Health study is large and appears to be well designed and comprehensive. It includes a cohort and nested case-control component, as well as extensive current measurements of exposure to diesel exhaust, detailed reconstruction of historical exposure, and bio-

Figure 1. Panel's analysis depicting consistently elevated risk of lung cancer for train workers compared with clerks for each time period, but decreasing risk by job category over duration of employment. See Appendix C for details.

marker development. These studies in progress are likely to inform hazard identification, exposure estimation, and exposure-response analyses, all components of risk assessment.

The Panel recognizes that regulatory decisions need to be made in spite of the limitations and uncertainties of the few studies with quantitative data currently available. The findings described here and the systematic evaluation of these and other studies are designed to inform the ongoing process and provide a means to weigh a study's strengths and limitations.

FINDINGS

GENERAL

Enhanced exposure and epidemiologic data and analyses are needed for the purposes of QRA; these might come from further exploration of existing studies or from new studies.

RAILROAD WORKER STUDIES

At present, the railroad worker cohort study (Garshick et al. 1988), though part of a larger body of hazard identification studies, has very limited utility for QRA of lifetime lung cancer risk from exposure to ambient levels of diesel exhaust for the following reasons.

- The various exposure-response analyses are limited by the scope and quality of currently available exposure data. Quantitative exposure data were not obtained during the cohort study period. Also, there is a paucity of qualitative data on individual exposures before 1959, and on the variation in exposure by railroad site, by season, and over time. The potential impact of concurrent exposures (for example, to grease, dust, other fumes, asbestos, and active and passive cigarette smoke) were not examined in depth. The diesel exhaust exposure data are suitable for a crude categorical measure of exposure by job category; but other measures, including duration of employment in a job category exposed to diesel exhaust, intensity of exposure concentration ($\mu g/m^3$), and lifetime exposure ([$\mu g/m^3$]-years), are not adequate to support quantitative exposure-response analyses.
- The Panel's analysis of the exposure-response association in the railroad worker data showed that the evidence for a positive association of lung cancer with cumulative exposure to diesel exhaust depends entirely on differences in risks among job categories. Train workers (with higher exposures) have higher risks compared with clerks (with low or no exposure). However, within all job categories, the relation of lung cancer risk to duration of employment is negative.
- Factors that might explain a negative association between duration of employment and lung cancer in these data include bias introduced by systematic differences in exposure misclassification among and within job categories; differentially incomplete ascertainment of lung cancer deaths by job category; lack of information on other occupational exposures and air pollutants; the presence of a healthy worker survivor effect; confounding by cigarette smoking; and analysis of relative risks rather than absolute risks. Also, in a case-control study, if causes of death among controls were associated with exposure to diesel exhaust, smoking, or both, the results could be biased.

TEAMSTER STUDIES

The investigators' analysis of the teamster data reported an exposure-response relation (Steenland et al. 1998) that may be useful for QRA; this relation will be better understood with further exploration of uncertainties and assumptions, particularly those relating to the reconstruction of past exposures and the selection of controls. Exposures of teamsters are more similar to ambient exposures of the public than are exposures of railroad workers, and the diesel exhaust to which teamsters are exposed comes from a source that is likely to be relevant to regulatory issues.

The Panel reviewed the teamster study without the benefit of additional analyses and interpretations, and its comments are not as detailed as those about the railroad worker studies. Understanding the teamster study will evolve with time; however, some conclusions can be drawn now.

- The set of teamster studies may provide reasonable estimates of worker exposure to diesel exhaust, but significant further evaluation and development are needed. The marker for diesel exhaust that was selected for study by Steenland and associates, EC_1, is more sensitive and specific than RSP adjusted for environmental tobacco smoke, but has several limitations (e.g., the contribution of diesel emissions to ambient EC_1 concentrations has not been constant over time). The industrial hygiene study, which was conducted after the period when workers in the case-control study were exposed, identified a range of exposures for various job categories, but did not consider (1) site-to-site varia-

tions, (2) seasonal variations, (3) concurrent exposures to other agents, (4) historical ambient particle concentrations, or (5) intra- and interindividual variability. The estimation of historical exposures needs to incorporate recent data on diesel emissions from vehicles in use, reassessment of when dieselization occurred, alternatives to estimating exposure by vehicle miles traveled, and historical regional ambient pollution data.

- The exposure-response relation reported in the teamster study increases in a linear manner. However, more can be learned from other analysts examining these data using different approaches.
- Neither a roster of the study population nor an alternative method of selecting controls to represent it was available to the researchers. It cannot be established with certainty whether the causes of death used for controls adequately represent the joint distribution of exposure to diesel exhaust and smoking in the case-control study. If smoking, or diesel exhaust exposure as determined by job category, or both were associated with causes of death used for controls, results could be biased.

RECOMMENDATIONS

The Panel's recommendations reflect its general understanding, as expressed in its framework for evaluating studies, of what constitute adequate data for QRA. They also reflect the preceding evaluation of the studies of railroad workers and teamsters. The Panel is aware that research currently in progress will respond to some of these research needs; however, results are not yet available, and it is not yet clear whether all of the proposed needs will be met.

COMPLETED STUDIES

1. The Panel recommends against using the current railroad worker data as the basis for QRA in ambient settings.
2. Further scrutiny of the teamster data, including estimation of uncertainty in both the exposure estimates and selection of controls, is recommended in order to improve the use of these data in QRA. Strengths of the teamster study include the relevance of exposure levels to the general population and the use of an exposure marker for diesel engine emissions that was an improvement over RSP. The teamster study exposure-response analysis is relatively new, and its further review and analysis by both the original investigators and others should be accelerated. Alternative retrospective exposure models need to be developed that use the alternative assumptions described above and in more detail in the body of the text.

NEEDS FOR NEW TECHNIQUES AND DATA

3. Better measures of exposure to constituents of diesel emissions, with careful attention to selection of the sample studied, are needed. Of particular importance are the selection and validation of a chemical marker of exposure to the complex mix of diesel exhaust emissions. 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.
4. 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.
5. Although biomarker technology was not available when the studies reviewed were conducted, appropriate, validated, and specific biomarkers of diesel exposures, health outcomes, and susceptibility are needed.

DESIGN NEEDS FOR NEW STUDIES OF EXPOSURE-RESPONSE ANALYSES

6. Exposures should be adequately and accurately characterized with respect to magnitude, frequency, and duration, rather than solely by duration of employment. 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.
7. 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.
8. 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.

NEEDS FOR NEW STUDIES

A prospective epidemiologic study of the development of lung cancer in exposed and unexposed individuals could have many strengths. Information on confounders and exposures could be more complete than for a retrospective study, and many of the biases and uncertainties discussed in this report could be eliminated or reduced. These advantages, however, need to be weighed against the disadvantages, which include high costs and a long period of follow-up. Other study designs that include retrospective components are possible for a new epidemiologic study of lung cancer, but they are likely to include uncertainties and sources of bias that investigators will need to explore completely and acknowledge in their reporting.

9. The Panel recommends that a new, large, epidemiologic study of diesel exhaust emissions and lung cancer be considered after (1) currently ongoing or existing studies, including HEI's feasibility studies (to be completed in the spring of 2000), are evaluated, and (2) attempts to retrofit improved exposure assessments to existing epidemiologic studies are evaluated, including whether they can provide sufficiently accurate, complete, and relevant exposure data to support QRA.
10. Studies of lung cancer risk in general populations exposed to ambient diesel exhaust particulate matter will be difficult to conduct; however, such studies could usefully investigate other, noncancer health effects that occur in a shorter time after exposure.

Diesel Emissions and Lung Cancer: Epidemiology and Quantitative Risk Assessment

Background

Diesel engines are an important part of the world's transportation and industrial infrastructure, especially in heavy-duty machinery such as trucks, buses, construction and farm equipment, locomotives, and ships. The use of diesel engines may expand in the future because they are energy-efficient and durable.

In Europe, where fuel prices are three to four times higher than in the United States and concern for climate change is high, 20% to 50% of the new light-duty passenger fleet is powered by diesel engines. In 1996, 1.8% of the light-duty vehicles (passenger cars, light-duty trucks [up to 10,000 pounds gross vehicle weight], vans, and sport utility vehicles) sold in the U.S. were diesel-powered, almost all of which were light-duty trucks. Of all the light-duty trucks sold that year, about 3.7% were diesel-powered; whereas only about 0.1% of all passenger cars sold had diesel engines (Davis 1998). Advanced technology diesel engines are being proposed as key elements in the U.S. energy conservation strategy (e.g., for use in sport utility vehicles and in the Partnership for a New Generation of Vehicles' fuel-efficient car).

Although diesel engines have economic advantages, environmental issues must be considered. Emissions from all types of engines are highly variable and complex mixtures. Diesel engines are more efficient (per vehicle mile traveled) than gasoline engines and emit less carbon dioxide (a greenhouse gas), carbon monoxide, and hydrocarbons; however, they emit more oxides of nitrogen, which are ozone precursors, and particulate matter. It is difficult to estimate exposures to individual constituents of diesel engine emissions because the amount and concentration of each constituent depend on such factors as engine type, fuel, and operating conditions. Moreover, fuel reformulation and changes in engine technology have caused substantial changes in diesel emissions over time. In addition, it is challenging to distinguish diesel emission constituents in ambient air from other combustion products and cigarette smoke.

The particulates emitted in diesel exhaust are of special concern in possible health outcomes because (1) they are very small (less than 1 μm in size) and readily respirable, and (2) they have many chemicals adsorbed to their surfaces, including some known or suspected mutagens and carcinogens. Cellular, animal, and human studies have investigated the association between exposure to diesel exhaust and cancer and other diseases.

DIESEL EXHAUST AND LUNG CANCER

Scientists have conducted toxicologic and epidemiologic studies to examine the potential for diesel emissions to cause or contribute to the development of cancer and other diseases. Some studies have looked specifically at the particulate matter component of diesel emissions; others refer simply to "diesel exhaust." In this report, diesel particulate matter (DPM) is specified when applicable. Laboratory studies have established that lifelong exposures to high concentrations of DPM produce lung tumors (benign and malignant) in rats, equivocal results in mice, and no tumors in hamsters. The roles that high-dose exposure protocols and species-specific factors have in the induction of rat lung tumors by DPM have been investigated. (This material is reviewed in publications by HEI [1995], the World Health Organization [WHO] [1996], and the International Life Sciences Institute [ILSI] [1999]). Rats develop lung tumors (benign and malignant) when they are exposed to DPM at concentrations of 2,000 to 10,000 $\mu g/m^3$ for 35 hours or more each week over their lifetimes (HEI 1995). Prolonged exposure to high concentrations of a variety of other supposedly inert particulate materials also causes lung tumors in rats through a mechanism that involves impairment of lung clearance mechanisms (referred to as "lung overload response"). This impairment can lead to inflammation, cell proliferation, metaplasia, and ultimately the development of lung tumors (HEI 1995; ILSI 1999). The levels of DPM required to produce lung tumors in rats, however, are approximately three orders of magnitude higher than current estimates of average ambient (nonoccupational) concentrations of DPM. Because lung overload is not expected to occur in humans as a result of ambient or most occupational exposures to DPM, some organizations have suggested that the rat lung tumor response to high concentrations of particulate matter is not relevant for quantitative risk assessment (QRA) (HEI 1995; ILSI 1999).

More than 40 epidemiologic studies of workers have examined the association between exposure to diesel exhaust and the risk of lung cancer. Several review articles discuss this literature in depth (Cohen and Higgins 1995; WHO 1996; Boffetta 1997; Bhatia et al. 1998; Office of Health Hazard Assessment [OEHHA] 1998; U.S. Environmental Protection Agency [EPA] 1998). The epidemiologic studies generally show higher risks of lung cancer among persons occupationally exposed to diesel exhaust than among persons who have not been exposed, or who have been exposed to lower levels or for shorter periods of time. Occupational groups studied include railroad workers, truck drivers, bus garage workers, heavy equipment operators, dock workers, and underground miners. In these studies, the relative risk (RR), as a measure of association between exposure and lung cancer, generally has been between 1.2 and 1.5 (that is, an excess of 20% to 50% over the risk in unexposed persons); somewhat more variation in relative risk was reported among subgroups examined in individual studies. Some reviews critical of these data have cited study design flaws, including uncontrolled confounding and lack of exposure measures, leading to a lack of convincing evidence (Muscat and Wynder 1995; Stöber and Abel 1996; Morgan et al. 1997).

Two studies are under way to evaluate the association between exposure to diesel emissions and lung cancer among nonmetal miners in Germany (Säverin et al. 1998) and in the United States (National Cancer Institute–National Institute for Occupational Safety and Health [NCI-NIOSH] 1997). Some results of the former study are available now; those of the latter study are expected in about 2003.

RISK ASSESSMENTS OF DIESEL EMISSIONS

Several organizations have reviewed the relevant science, including the epidemiologic, toxicologic, and experimental studies of diesel engine exhaust, and have classified (or proposed to classify) the exhaust mixture, or the particulate component of the mixture, as a potential, probable, or definite human carcinogen (NIOSH 1988; International Agency for Research on Cancer [IARC] 1989; WHO 1996; OEHHA 1998; National Toxicology Program [NTP] 1998; U.S. EPA 1998). Each agency's current position on the carcinogenicity of diesel exhaust as cited in its draft or final report is as follows:

NIOSH (1988)

- Animal evidence "confirmatory" for carcinogenicity
- Human evidence "limited"
- Diesel exhaust classified as a "potential occupational carcinogen"
- No QRA

IARC (1989)

- Rat data "sufficient" for carcinogenicity
- Human epidemiologic data "limited"
- Diesel exhaust considered a "probable" human carcinogen
- No QRA

WHO (1996)

- Rat data support carcinogenicity
- Human epidemiologic data suggest "probably carcinogenic"
- Epidemiologic studies considered "inadequate for a quantitative estimate of human risk"
- Rat data used for QRA

California EPA (1998)

- Rat data "have demonstrated" carcinogenicity of diesel exhaust
- Causal association of diesel exhaust and lung cancer in epidemiologic studies is a "reasonable and likely explanation"
- Human epidemiologic data preferred for QRA because of uncertainties in rat data
- California Air Resources Board designated DPM a "toxic air contaminant"

NTP (1998 Draft)

- Committees have considered listing DPM as either "known to be a carcinogen" or "reasonably anticipated to be a carcinogen"
- Internal review complete; Directors' decision expected in 1999

U.S. EPA (1998 Draft)

- Rat experiments "adequate" for carcinogenicity
- Human epidemiologic studies "limited" evidence
- Diesel emissions considered "probable" human carcinogen
- Range of cancer risk estimates developed (on the basis of animal, epidemiologic, and comparative potency data)
- Revised risk assessment expected in 1999

In 1994, the U.S. EPA and the California EPA both released draft cancer risk assessments of diesel exhaust for public comment and review by scientific experts. Despite general agreement in their interpretation of the scientific literature, some important differences were apparent between the two reports. Most notably, the basis for the U.S. EPA's QRA was animal bioassay data because, in the view of the Agency's staff, the exposure data from human epidemiologic studies were too limited to support a QRA. In addition, the U.S. EPA staff, in conjunction with Dr. Kenny Crump, had tried to use data from a retrospective cohort study of U.S. railroad workers (Garshick et al. 1988) and an associated industrial hygiene survey (Hammond et al. 1988; Woskie et al. 1988a,b) to construct exposure-response estimates (Crump et al. 1991). Crump and colleagues did not find a positive exposure-response association for diesel exhaust and risk of lung cancer, however, and therefore concluded that it was not possible to use these data for a quantitative analysis.

In contrast, staff in the OEHHA of the California EPA, using the same railroad worker data, found an increasing exposure-response relationship (OEHHA 1994, 1998). Because of uncertainties in extrapolating from rat data to humans and the fact that some semiquantitative epidemiologic data were available, OEHHA determined that it was more appropriate to base risk assessment estimates for diesel exhaust on the epidemiologic data than on animal data.

The difference between these findings led HEI in 1996 to collaborate with the two agencies and others (NIOSH and WHO) to sponsor a scientific workshop, "Diesel Exhaust: Considerations in the Use of Epidemiologic Data for Quantitative Risk Assessments (QRA)," that focused on the strengths and limitations of the existing database. Participants discussed issues that underlie the differences in the exposure-response modeling of the railroad worker data. Although the findings were explored at length, workshop participants were unable to determine the reasons for the differences. The railroad worker data were the only epidemiologic data used in QRA until Steenland and colleagues (1998) published an exposure-response analysis of U.S. teamsters.

Questions still remain about how to develop a reliable QRA of diesel engine emissions and lung cancer. If the rat lung tumor data are not relevant for human cancer risk assessment, and if current epidemiologic studies do not provide the quantitative exposure measurements needed for exposure-response estimates, it is difficult to make informed decisions about possible health risks from exposure to diesel exhaust. Should existing epidemiologic studies be extended to include additional years of follow-up? Have diesel engines and fuels changed so much that studies of animals or humans exposed to diesel exhaust from old engines are no longer relevant? Government, industry, and the public have an interest in answers to these questions. HEI initiated its Diesel Epidemiology Project to help inform the decisions about appropriate ways to use the existing epidemiologic data and to suggest future research directions.

The HEI Diesel Epidemiology Project

In 1998, HEI initiated a multifaceted Diesel Epidemiology Project in response to the issues discussed in the Background chapter. This project includes the evaluation by HEI's Diesel Epidemiology Expert Panel of occupational epidemiologic studies that are currently being used for QRA, and the development of new research initiatives, including six feasibility studies to identify potential new cohorts to study or to improve exposure assessment estimates.

The Diesel Epidemiology Expert Panel (referred to here as "the Panel") was chaired by John C. Bailar III, M.D., Ph.D., of The University of Chicago and the HEI Health Review Committee, and included six other members (see Appendix A) with expertise in epidemiology, biostatistics, exposure characterization, and exposure assessment. Its charge was to (1) review the epidemiologic data that form the basis of current QRAs for diesel exhaust, (2) identify data gaps and sources of uncertainty, (3) make recommendations about the usefulness of extending or conducting further analyses of existing data sets, and (4) make recommendations for the design of new studies that would provide a stronger basis for risk assessment. The Panel was not charged to evaluate the broad epidemiologic literature concerning exposure to diesel exhaust and lung cancer for hazard identification purposes.

In response to the first charge, the Panel examined published epidemiologic studies for possible use in or contribution to QRA for diesel exhaust. However, quantitative exposure data were associated in some manner with only two epidemiologic studies, and those are considered in this report. Other diesel epidemiologic studies were not considered, because the lack of associated quantitative exposure data makes those studies less suitable for QRA. The studies forming the core of the Panel's review include:

Railroad Worker Studies

- Case-control: Garshick et al. 1987
- Cohort: Garshick et al. 1988
- Industrial hygiene: Hammond 1988, and Woskie et al. 1988a,b
- Exposure-response analyses: Crump et al. 1991, OEHHA 1998, and Crump 1999

Teamster Studies

- Case-control: Steenland et al. 1990, 1992
- Industrial hygiene: Zaebst et al. 1991
- Exposure-response analysis: Steenland et al. 1998

The Panel met on April 20–21, 1998, in Cambridge, MA, to hear presentations by the principal investigators of the epidemiologic studies listed above, and by the secondary analysts of the railroad worker cohort study. The Panel also began discussing how to approach its charge. An agenda for the workshop portion of the April 1998 meeting and a list of participants are in Appendix A. The Panel met in executive session on October 8–9, 1998; December 1, 1998; January 21, 1999; and by conference call on February 24 and April 9, 1999, to continue its work on this report.

At the first meeting of the Panel, inconsistent results of the exposure-response analysis of the railroad worker cohort data were presented by the secondary analysts. The Panel determined that it would need to explore the exposure-response relation in the original railroad worker data set in order to understand firsthand the reasons for the discrepancies. The Panel did not attempt to conduct its own exposure-response analysis for QRA. Eric Garshick, M.D., assisted the Panel by providing a copy of the original railroad worker data and documentation. Secondary analyst Stanley Dawson, Ph.D., of California EPA's OEHHA, provided documentation of his analysis of the original data set; likewise, Kenny Crump, Ph.D., provided the computer code listings he used for aggregating the original railroad worker data. The Panel is grateful to all these analysts for their cooperation and generosity in supporting this review.

THE PANEL'S APPROACH TO EVALUATING THE STUDIES

Epidemiologic studies of environmental pollutants and cancer are relevant to risk assessment, which includes some or all of the following steps: hazard identification (determination of whether or not an agent is causally linked to a health effect), exposure assessment (degree, timing, and level of exposure), dose-response assessment

Table 1. The Panel's Framework for Evaluating Epidemiologic Studies for Quantitative Risk Assessment

Element of Study	Questions for Evaluation
DESIGN	1. Was the study design efficient, and did it specifically consider power, potential types of bias, and latency?
	2. Were adequate quantitative exposure and covariate data planned and collected?
	3. Was the literature reviewed for relevant study methods and previous research findings?
	4. Were plans made for postpublication data scrutiny?
EXPOSURE ASSESSMENT	1. Were detailed lifetime histories of occupational exposures collected?
	2. Were known chemical and physical characteristics of the main exposure specified?
	3. Were potentially confounding exposures measured or estimated?
	4. Were magnitude, duration, and variability of exposure determined?
	5. Were industrial hygiene data and historical data on use and repair of machinery and equipment obtained?
	6. Were personal exposure measurements obtained, and were they representative of the population studied?
	7. Were uncertainties in exposure assessment quantified?
OUTCOMES	1. Were outcomes defined in specific and objective terms?
	2. Was the full range of outcomes included in cohort studies?
	3. Were participants actively followed to determine outcome status and the date outcome occurred?
ANALYSIS	1. Were analytic methods specified a priori?
	2. Was the appropriateness of the statistical approach demonstrated, and were potential biases explored?
	3. Were exposure-response relations statistically explored?
	4. Were uncertainties in risk estimates quantified, especially those resulting from exposure measurement error?

(determination of the relation between the magnitude of exposure and the health effect, including in subpopulations), and risk characterization (description of the nature and magnitude of risk, with uncertainties) (National Academy of Sciences 1983).

All of the characteristics that make a study relevant to hazard identification also help make a study relevant for QRA. But QRA requires more, including quantitative exposure data expressed in units that are comparable among the epidemiologic settings and the situations for which the risk estimates are desired. The quantitative data must be sufficient to construct an exposure-response curve on the basis of at least two levels of exposure. Occupational exposure levels often are much greater than ambient (nonoccupational) exposures. When occupational studies are used in a QRA that is to be extended to populations receiving ambient exposures, the analysts may need to extrapolate over a range from the high exposures observed to the lower exposures of interest; the closer the occupational exposures are to the ambient exposures of concern, the less extrapolation is required.

An important limitation of using most epidemiologic studies for QRA is the lack of adequate exposure data, especially historical information on exposure concentrations or

rates. In retrospective studies, exposure assessment is almost always less than ideal; the problems and uncertainties in using such data need to be clearly acknowledged.

The Panel recognized that few, if any, studies are designed specifically for use in QRA. It determined that the most objective way to identify data gaps and sources of uncertainty was to systematically review the epidemiologic studies of railroad workers and teamsters*. To guide this process, the Panel developed a framework of general epidemiologic principles, described in Table 1.

This framework presents a series of questions about elements of study design, exposure assessment, outcome determination, and analysis that the Panel determined to be important for this evaluation. An affirmative answer to all questions is clearly unattainable for any epidemiologic study. The answers, however, not only highlight the strengths and limitations of each study, but also indicate where to focus additional or new research efforts. A rationale for including each of the items in the framework is detailed in Appendix B.

A draft version of this entire report was peer-reviewed by 18 external reviewers, including Drs. Garshick, Crump, and Dawson. All of the reviewer comments and concerns were considered and addressed as appropriate by the Panel. Preliminary, draft findings were presented by Panel members at HEI's Diesel Workshop: Building a Research Strategy to Improve Risk Assessment (1999), held at Stone Mountain, GA, March 7–9, 1999.

* The reference to "railroad worker study or studies" in the report refers specifically to the studies by Garshick and colleagues (1987, 1988) and, if appropriate, to the industrial hygiene studies by Woskie and associates (1988a,b) and Hammond (1988). The reference to "teamster study or studies" in the report is specific to the case-control study by Steenland and coworkers (1990), subsequent reports based on these data (Steenland et al. 1992, 1998) and, if appropriate, to the industrial hygiene study by Zaebst and associates (1991).

Summary of Railroad Worker and Teamster Studies

Two sets of studies widely cited as "key" or "important" epidemiologic investigations of diesel exhaust and lung cancer have been used for QRA. One set of studies includes cohort (Garshick et al. 1988) and case-control (Garshick et al. 1987) studies of railroad workers, industrial hygiene measures in railroad work sites (Woskie et al. 1988a,b; Hammond et al. 1988), and secondary exposure-response analyses of these data (Crump et al. 1991; OEHHA 1994, 1998; Crump 1999). The second set includes a case-control study of teamsters (Steenland et al. 1990, 1992), industrial hygiene measures in the trucking industry (Zaebst et al. 1991), and a recent exposure-response analysis (Steenland et al. 1998). This chapter briefly summarizes each study's methods and results.

RAILROAD WORKER STUDIES

Garshick and colleagues conducted both cohort (Garshick et al. 1988) and case-control (Garshick et al. 1987) studies of lung cancer deaths among U.S. railroad workers registered with the Railroad Retirement Board (RRB). The authors' background materials report that the U.S. railroad industry began introducing large numbers of diesel engines in about 1949; dieselization of the industry was essentially completed within 10 years. Figure 2 shows the transition from the use of steam (mostly coal-fired) to diesel locomotives over time (Railroad Facts 1940–1970). In both studies, Garshick and colleagues defined four diesel exhaust exposure groups: shop workers; engineers and firemen; brakemen, conductors, and hostlers; and unexposed (clerks and signalmen).

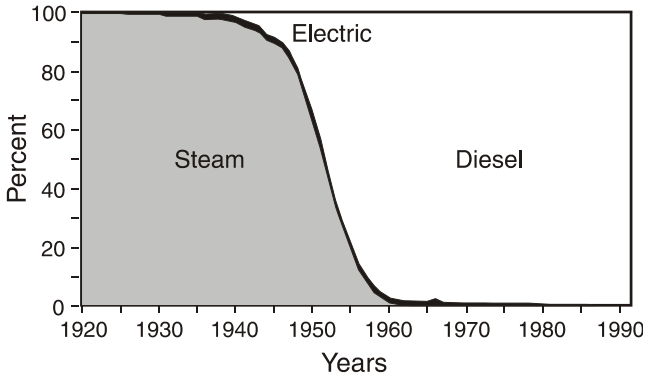

Figure 2. Percentage of trains powered by steam, electric, or diesel since 1920. (Data adapted from Railroad Facts 1940, 1944, 1946, 1948, 1950, 1953, 1954, 1957, 1964, 1967, 1970.)

COHORT STUDY

In the cohort study, Garshick and colleagues (1988) reported that they ascertained all deaths from 1959 to 1980 in a cohort of 55,407 white male railroad workers who had been actively employed and between 40 and 64 years of age in 1959, when nationwide conversion to diesel railway engines was nearly complete. Eligible subjects had begun work 10 to 20 years earlier (between 1939 and 1949), and in 1959, were employed in one of 39 jobs surveyed later in a companion industrial hygiene study (Woskie 1988a,b). Railroad Retirement Board records listed each person's job title, or titles, for each year starting in 1959. To reduce the potential for confounding by asbestos exposure, all workers whose jobs involved known exposure to asbestos (car repair and construction trade workers, and some trade workers in steam locomotive shops) were excluded from the cohort. Some shop workers and hostlers, not initially excluded, were also exposed to asbestos. Cigarette smoking information was not available for cohort members.

Fact of death was ascertained from the RRB through December 31, 1980. Among the 19,396 known deaths, 1,694 of the death certificates indicated lung cancer as a primary or contributing cause of death.

The investigators classified the workers' exposures according to the job held in 1959 or the cumulative years in an exposed job. Their analytic model including years of exposure to diesel exhaust (on the basis of duration of employment) and lung cancer showed an increase in risk of lung cancer with more years of exposure. The investigators also reported that 94% of the workers who were 40 to 44 years of age and working in a job exposed to diesel emissions in 1959 were still in an exposed job 20 years later. Unexposed workers also generally remained in their exposure category (97%). The relative risks for lung cancer and exposure to diesel exhaust on the basis of the job held in 1959 were inversely related to age in 1959; workers who were 40 to 44 years of age and working in a job category with exposure to diesel emissions in 1959

experienced an increase (RR = 1.5; 95% CI: 1.1, 1.9) in lung cancer mortality compared with those who were in that age category but held unexposed jobs in 1959. Excess relative risk of death from lung cancer declined as the worker's age in 1959 increased.

CASE-CONTROL STUDY

Garshick and colleagues (1987) also conducted a case-control study of RRB registrants who died between March 1, 1981, and February 28, 1982. Among 650,000 active and retired male railroad workers born in or after 1900 who had at least 10 years of railroad employment, 15,059 deaths were reported to the RRB. Cases consisted of all deaths for which primary lung cancer (International Classification of Disease, Version 8 [ICD 8] code 162) was indicated on the death certificate; this was the underlying cause of death in most cases. Investigators attempted to match each case with two deceased control subjects by age at death (within 2.5 years), and by date of death (within 31 days). Men who died from other cancers, suicides, accidents, or unknown causes were excluded as control subjects. The most common underlying causes of death among both older (age at death 65 years or older) and younger (age at death 64 years or younger) controls were diseases of the circulatory system (74% and 80%, respectively); deaths from nonmalignant respiratory disease also were included (15% of the older control subjects and 7% of younger controls). Overall, 1,256 lung cancer cases and 2,385 controls were considered in the analysis.

Exposure to diesel exhaust was assessed using (1) job histories beginning in 1959 for workers who retired after 1959, and (2) the last job worked before retirement for those who retired between 1955 and 1959. Each job was classified either as exposed or unexposed; cumulative exposure to diesel exhaust was summarized for each worker as diesel-years of exposure. Unlike the cohort study, the case-control study included persons who had worked in jobs other than the 39 jobs used to estimate diesel exhaust exposure levels in the industrial hygiene survey (Woskie et al. 1988a,b). These additional jobs were considered to be exposed or unexposed on the basis of (1) the similarity of job activities and work locations in question to jobs for which industrial hygiene samples had been taken, and (2) the extent of contact with operating diesel equipment that the job entailed.

Information on two potentially confounding variables (smoking and asbestos exposure) was collected. For smoking history, next of kin (usually a spouse) provided information on whether the subject had ever smoked; this information was obtained for 86% of cases and 82% of control subjects. If the age at which a worker began smoking was not available (which was true for less than 6% of both cases and controls), age 16 was assumed. Pack-years* of smoking history were missing for 22% of cases and 25% of controls. Asbestos exposure was ascertained from the work histories of those who retired after 1959 and from the last job held for those who retired between 1955 and 1959.

The investigators conducted separate analyses for younger workers (who died at or before age 64) and older workers (who died at age 65 or older), primarily because they reasoned that heavy cumulative exposure to diesel exhaust was more likely among the workers who died at a younger age. A second reason was attributed to a cited reference (Doll and Peto 1981) that indicated that the cause of death on a death certificate may be less accurate for workers who died at an older age. No excess risk of death from lung cancer in association with exposure to diesel exhaust was observed among the older workers. Among the younger workers, with diesel exposure modeled as a continuous variable, more than 20 years of exposure to diesel exhaust was associated with a crude RR = 1.4 (95% CI: 1.0, 1.8) for lung cancer mortality. Adjusting for asbestos and cigarette smoking had little effect on this estimate (RR = 1.4; 95% CI: 1.1, 1.9). Among younger workers, when years of diesel exposure were categorized (0–4, 5–19, 20+), an adjusted RR = 1.02 (95% CI: 0.72, 1.45) was found for workers with 5 to 19 years of diesel exposure, and an RR = 1.64 (95% CI: 1.18, 2.29) for those with 20 or more years of exposure, compared with the referent group with 0 to 4 years of diesel exposure.

INDUSTRIAL HYGIENE STUDIES OF RAILROAD WORKERS

Woskie and colleagues (1988a,b) conducted an industrial hygiene survey of U.S. railroad workers in four small northern railroads in the early 1980s to estimate occupational exposures to diesel exhaust. (This study was conducted during a time after the period when the workers in the epidemiologic studies would have been exposed.) They first identified 39 job titles (from among the more than 150 U.S. Interstate Commerce Commission railroad job titles) that encompassed large numbers of workers and that were thought to indicate either minimal or substantial exposure to diesel exhaust. These job titles were then collapsed into 13 job groups and, for some analyses, into 5 career exposure groups (clerks; signal maintainers; engineers and firers; brakers, conductors, and hostlers; and shop workers).

* Pack-years is the number of packs of cigarettes smoked per day times the number of years smoked.

The investigators developed three markers of diesel exhaust exposure: (1) the concentration of respirable-sized particles (RSP); (2) the adjusted respirable particle (ARP) concentration, which removed the particle contribution of environmental tobacco smoke (ETS) from the RSP; and (3) the adjusted extractable mass (AEM) (Hammond et al. 1988).

The RSP was the simplest marker to estimate in over 550 air samples from workers' breathing zones in 13 job groups at the four railroads, as well as in 23 air samples from fixed samplers in various railroad locations. Respirable matter (median aerodynamic diameter ≤ 3.5 μm) was collected on filters for analysis. Investigators also developed a method to collect and analyze nicotine from the same samples as a marker for ETS. The estimated ETS concentration was subtracted from the estimated RSP to obtain the second diesel marker, ARP. Dichloromethane was used to extract the organic chemical components of diesel exhaust, including the mutagens and carcinogens, from the RSP. The fraction of extractable material in the RSP was adjusted for the fraction of extractable ETS for the third index of exposure, AEM. Crump and colleagues (1991) used another index, total extractable material (TEX), which is the concentration of extractable RSP without the ETS fraction removed. None of these measures accounts for other respirable matter in the RSP, such as sand, dirt, or fibers.

The investigators used these data from four railroad yards to estimate a national career mean exposure for ARP (Woskie et al. 1988a,b). They used a linear statistical model to adjust for climate differences and variability of exposures among railroads across the country. This model included weighting factors to estimate the fraction of the year a railroad was in a "cold" (below 10°C) or "warm" (above 10°C) climate (Woskie et al. 1988b).

EXPOSURE-RESPONSE ANALYSES OF RAILROAD WORKERS

The railroad worker cohort study suggested that lung cancer risk increased with increasing cumulative years of exposure (Figure 3). Some quantitative exposure data were available for the industry as well. Because both types of information were available, analysts have used the data to develop exposure-response estimates for diesel emissions and lung cancer.

Crump and colleagues (1991) were the first to develop quantitative estimates of lung cancer risk associated with exposure to diesel exhaust by combining data from the railroad worker cohort study (Garshick et al. 1988) with exposure estimates from the industrial hygiene studies (Hammond et al. 1988; Woskie et al. 1988a,b). Crump and colleagues assigned average exposures as defined in the industrial hygiene studies to members of the railroad worker cohort, on the basis of yearly job codes beginning in 1959. They conducted analyses using RSP, ARP, AEM, and TEX estimates.

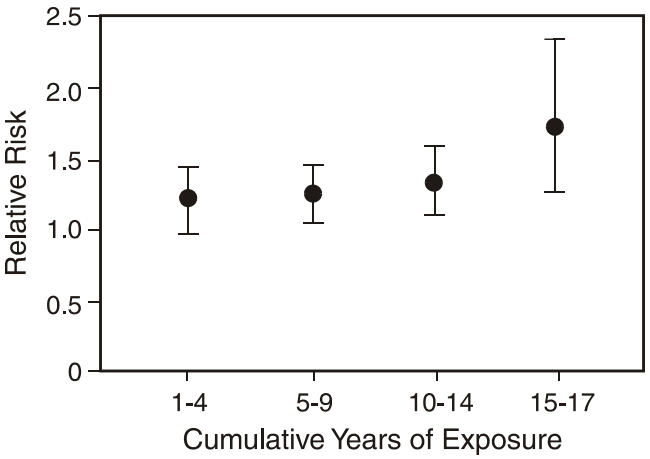

Figure 3. Relative risk of lung cancer by years of exposure to diesel exhaust in the railroad industry through 1980. Relative risk and 95% confidence intervals are shown. (Adapted from Figure 1 in Garshick et al. 1988.)

Crump and colleagues (1991) constructed several different exposure metrics that combined measures of particulate levels with information on regional climates for the U.S.; they used these metrics, plus age, calendar year, and five job categories (clerks; signal maintainers; engineers and firers; brakers, conductors, and hostlers; and shop workers) to conduct more than 50 analyses of the relation between exposure to diesel exhaust and death from lung cancer. All but two analytic models showed that subjects with the highest estimated cumulative exposures had the lowest risk of death from lung cancer (Figure 4).

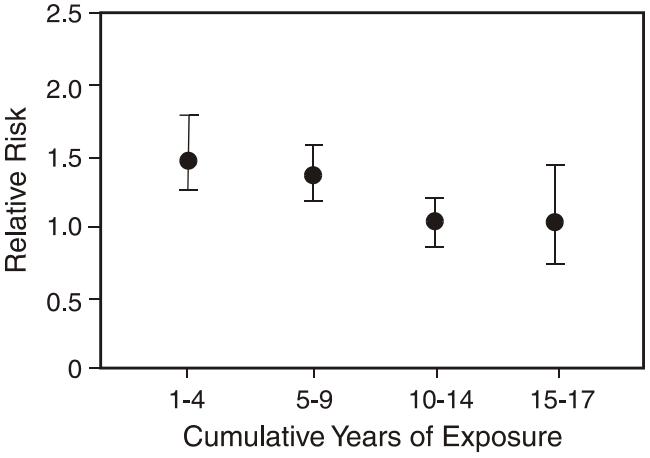

Figure 4. Relative risk of lung cancer by years of exposure to diesel exhaust in the railroad industry, using data from the cohort study. (Adapted from data presented in Crump et al. 1991.)

Crump and associates also discovered a limitation of the original cohort data. Whereas Garshick and colleagues (1988) had compared lung cancer mortality rates between exposed and unexposed railroad workers within the cohort, Crump and colleagues also compared the overall age- and year-specific death rates of the cohort to those rates in the U.S. population (Crump et al. 1991). This comparison suggested that follow-up (that is, determination of vital status) was incomplete between 1976 and 1980 as ascertained by the RRB. Garshick confirmed this, as noted in a letter to the U.S. EPA (Garshick 1991). This underascertainment of deaths from the RRB did not affect the major findings of Garshick and colleagues (1988) because Crump's analyses, limited to the years 1959 through 1976, revealed similar overall excess lung cancer mortality risk (Crump et al. 1991; Garshick 1991). In his letter, Garshick (1991) also presented a new analysis of lung cancer risk by years of diesel exposure, in which the data were modeled to allow the effect of age to vary in the cohort in a time-dependent manner between 1959 and 1980. Although the relative risks for all four exposure groupings were elevated, the estimated effect did not increase with increasing duration of exposure (Figure 5), in contrast to results Garshick had reported previously.

As part of the California EPA's risk assessment of diesel particulate matter, OEHHA analyzed exposure-response relations in the railroad worker data. The five job categories described were combined to form three exposure groups: exposed (engineers and firers; and brakers, conductors, and hostlers; collectively referred to as "train workers"), unexposed (clerks and signalmen), and uncertain (shop workers). Shop workers were excluded from some analyses because their exposure was assessed as uncertain (OEHHA 1998).

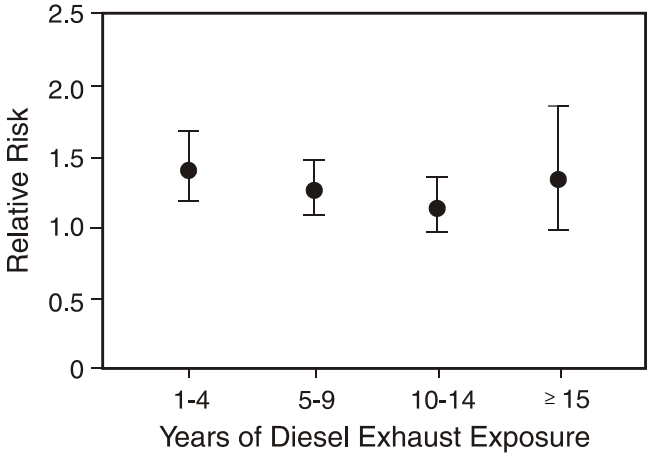

Figure 5. Relative risk of lung cancer by years of employment in the railroad industry through 1976. (Adapted from Table 4 in Garshick 1991.)

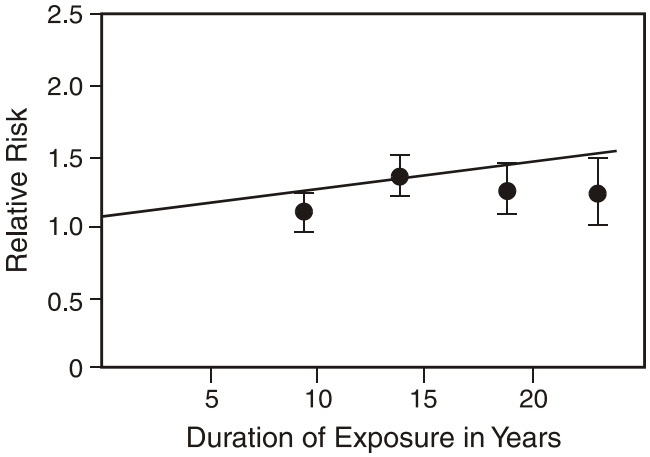

Figure 6. Relative risk of lung cancer by years of exposure to diesel exhaust. (Adapted from Figure 7-2 in OEHHA 1998.)

The analysis by OEHHA (1994, 1998) found a steadily increasing risk of lung cancer with increasing duration of exposure (Figure 6). This result conflicts with the findings of Crump and colleagues (Figure 4) and the revised Garshick analysis (Figure 5). Although the linear regression coefficient is positive, the relation shown in Figure 6 does not appear to be monotonic. On the basis of the linear increase, the California EPA estimated a range of lifetime unit risk (95% upper limit) of lung cancer from exposure to DPM to be from 1.3×10^{-4} (lifetime-μg/m^3)$^{-1}$ to 2.4×10^{-3} (lifetime-μg/m^3)$^{-1}$. The estimated risks are based on several assumptions, including (1) a linear increase in DPM exposure concentrations from zero in 1945 to a peak in 1959 (described below as the "roof" pattern of exposure) that was 1 to 10 times the concentration measured in 1980; (2) a linear decline to the 1980 value after the initial increase and peak in 1959; (3) exclusion of shop workers from some analyses because their degree of exposure was uncertain; (4) various statistical methods to control for age and calendar year; and (5) subtraction of "background" exposure levels measured for clerks and signal maintainers from the exposure levels measured for the train workers.

Different assumptions were made by OEHHA and Crump to reconstruct diesel exposures before 1959, because no actual data were available. Figure 7 shows a schematic representation of how pre-1959 exposure was represented in analytic models by Garshick and colleagues (1988), Crump and coworkers (1991), and OEHHA (1998). The cohort analysis (Garshick et al. 1988) assumed that exposure began in 1959, and that the 1959 exposure level remained constant until 1980 (which has been referred to as the "block" exposure pattern). Crump used the "ramp" exposure pattern, with a linear increase in exposure from the

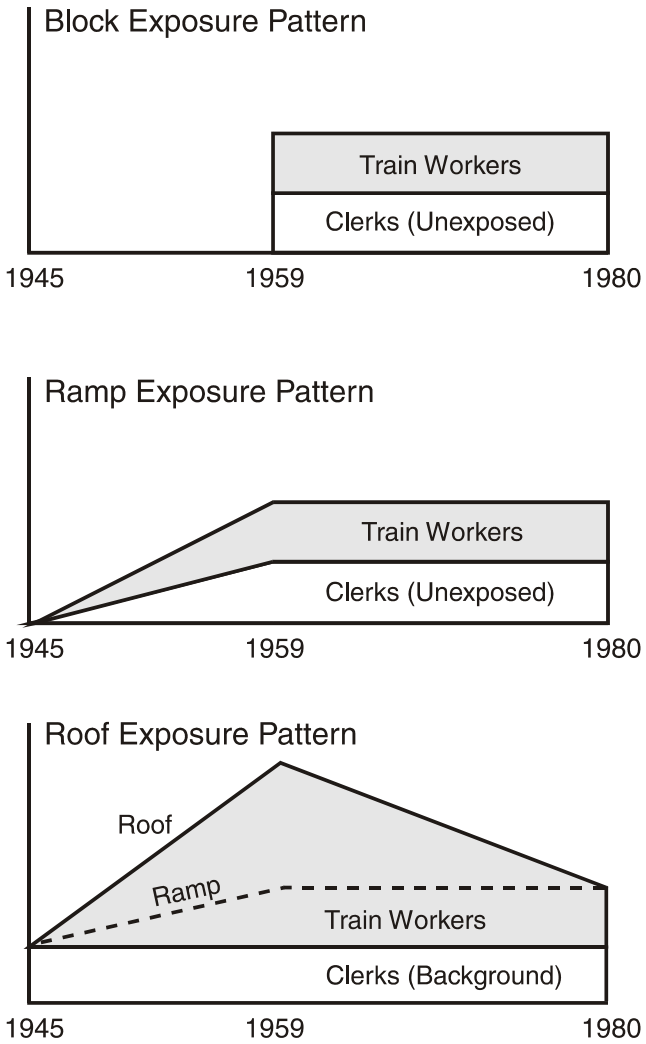

Figure 7. Different exposure patterns from assumptions made about pre-1959 diesel exposures. Garshick and colleagues (1988) made no assumptions about exposure before 1959, which has been referred to as a block pattern. Crump and coworkers (1991) assumed a linear increase in exposure from 1945 to 1959 and constant exposure thereafter (the ramp pattern). The roof pattern (OEHHA 1998) is based on the assumption that engines were "smokier" in the past, increased to a peak in 1959, and decreased to measured levels in 1980. (Adapted from Figure 7-4 in OEHHA 1998.)

time diesel-powered locomotives were introduced in 1945 to a peak of diesel engine use in 1959 and constant exposure levels from 1959 until 1980. This model assumes that exposures for both train workers and clerks followed this pattern, though at different magnitudes of exposure. The "roof" exposure pattern assumes a constant background level of exposure for the clerks before and after 1959. The additional exposure for the train workers was assumed to increase to a maximum in 1959 of three times the exposure level of the ramp model, followed by a decreasing exposure to the levels measured in 1980. Analysts at OEHHA used the roof model of diesel exposure because they thought it more nearly approximated anecdotal reports that engines were "smokier" in the past, and decreased to levels measured in 1980 (OEHHA 1998).

TEAMSTER STUDIES

CASE-CONTROL STUDY

Steenland and colleagues (1990) conducted a case-control study of lung cancer mortality in the Central States Teamsters Union. Cases and controls were selected from among 10,699 male union members who had filed for pension benefits and who died in calendar years 1982 and 1983. Cases (n = 996) constituted all deaths in which lung cancer was reported as an underlying or contributing cause of death (ICD code 162 or 163). Control subjects (n = 1,085) consisted of every sixth death in the file of decedents, excluding deaths from lung or bladder cancer and motor vehicle accidents. As with the railroad worker case-control study, deaths from nonmalignant respiratory disease were included among the control subjects.

Exposure to diesel exhaust was ascertained in two ways. Interviews with the next of kin were conducted for 82% of cases and 80% of control subjects to obtain a lifetime work history. Study subjects were classified according to the job category in which they had worked the longest: diesel truck driver, gasoline truck driver, driver of both types of trucks, truck mechanic, or dock worker (a person who worked on truck loading docks or in warehouses). Subjects who had never worked in any of the above categories were defined as unexposed to diesel exhaust.

The second source of exposure information consisted of Teamsters Union pension applications that had been completed by the study subjects; these listed each occupation, employer, and the dates of employment. As with the data provided by next of kin, study subjects were categorized according to the job they held the longest: long-haul driver, short-haul or city driver, truck mechanic, or dock worker; others were classified as unexposed. Most subjects had worked in only one job category. The Teamsters Union data did not provide information on whether trucks were gasoline or diesel. The concordance between exposure classifications based on Teamsters Union records and on the next-of-kin interviews was generally high, but it varied among job categories. Over 90% of the men identified by their next of kin as diesel truck drivers were listed as long-haul drivers in the Teamsters Union records, and 82% of workers identified as mechanics by their next of kin were listed as such in the same records. Information

on the amount and duration of cigarette smoking was obtained from the next of kin, as was information on asbestos exposure and diet.

When subjects who had ever been employed in any of the index job categories were compared with those who had not, the relative risk of lung cancer was elevated (but not statistically significant) in all major occupational categories except dock workers. Relative risk estimates for lung cancer appeared to increase with duration of employment after 1959 for both long-haul and short-haul truck drivers.

A similar pattern was observed in analyses that were based on length of employment as a driver of diesel trucks, using job histories provided by the next of kin. Employment as a diesel truck driver for 35 years or longer was associated with an 89% increase in lung cancer mortality (RR = 1.89; 95% CI: 1.0, 3.4). However, no relation between duration of employment and excess lung cancer mortality was evident for mechanics working 35 years or more (RR = 1.09; 95% CI: 0.44, 2.7).

INDUSTRIAL HYGIENE STUDY OF TRUCKING INDUSTRY WORKERS

In conjunction with the case-control study of teamsters, Zaebst and colleagues (1991) surveyed exposures to DPM in the four job categories identified by the Teamster Union records: long-haul or road drivers, short-haul or local drivers, mechanics, and dock workers. This survey was conducted in a time period after workers in the epidemiologic study would have been exposed. Samples were measured among workers in these four groups at six "breakbulk" terminals (hubs where inbound long-distance loads are broken into smaller loads for delivery). Additional data were obtained for dock workers at another terminal and for mechanics at a small truck repair shop. In addition, area samples were collected near a highway and in a residential area to estimate background concentrations. Levels of submicrometer-sized elemental carbon (EC_1) particles were used as the principal marker of exposure to whole diesel exhaust. Table 2 presents the EC_1 data from the industrial hygiene study and the odds ratios from the epidemiologic study.

Steenland and colleagues (1992) interpreted the EC_1 measurements as being generally consistent with the epidemiologic results. The industrial hygiene survey (Zaebst et al. 1991) indicated that measured levels of EC_1 for drivers did not differ substantially from highway background levels at the time of the survey.

Different numbers of dock workers were reported by Zaebst and colleagues (1991) and Steenland and colleagues (1998). In 1991, the industrial hygiene study reported that dock workers used equipment with engines powered by diesel (n = 54), gasoline (n = 9), or propane (n = 12). However, primarily propane-powered engines were used by dock workers until diesel-powered engines were introduced in the early 1980s, which is an insufficient latent period for development of lung cancer in 1982 and 1983, when subjects were identified for the epidemiologic

Table 2. Sample Means of EC_1 and Estimates of an Association Between Exposure to EC_1 and Lung Cancer Risk by Job Category in the Trucking Industry[a]

Job Category or Location[b]	n	EC_1 Geometric Mean ± SD[b] (μg/m^3)	EC_1 Arithmetic Mean ± SE[b] (μg/m^3)	Odds Ratio (95% CI)[c]
Dock workers	12	1.3 ± 2.0	1.6 ± 0.4	0.93 (0.55, 1.55)[d]
Mechanics	80	12.1 ± 3.7	26.6 ± 4.1	1.69 (0.92, 3.09)
Short-haul drivers	56	4.0 ± 2.0	5.4 ± 0.9	1.31 (0.81, 2.11)
Long-haul drivers	72	3.8 ± 2.3	5.1 ± 0.4	1.27 (0.83, 1.93)
Roadside area samples	21	2.5 ± 2.4	3.4 ± 0.5	NA
Off-roadway area samples	23	1.1 ± 2.0	1.4 ± 0.2[e]	NA

[a] Table compiled from Steenland and associates (Table 2 in 1990, Table 1 in 1998) and Zaebst and coworkers (Table 3 in 1991).

[b] From Zaebst and coworkers (1991).

[c] From Steenland and associates (1990). Reference job category for odds ratios is workers who had never worked in the other job categories listed nor in any other diesel-exposed job. NA = Not applicable.

[d] Odds ratio from Steenland and associates (1998) for dock workers using propane forklifts only.

[e] Value taken from Zaebst and coworkers (1991); same value appears as geometric mean for off-roadway area samples in Steenland and associates (1998).

study. Subsequent reports utilizing these industrial hygiene data include exposure levels measured only among those 12 dock workers using propane-powered equipment (Steenland et al. 1992, 1998).

EXPOSURE-RESPONSE ANALYSES

Steenland and colleagues (1998) reported results of an exposure-response analysis for EC_1 exposure as a marker for diesel exhaust and lung cancer in the teamster study. In the original case-control study, Steenland and colleagues (1990) had estimated exposures for subjects from work history. Industrial hygiene measures (Zaebst et al. 1991) were combined with past estimates of worker exposure on the basis of the investigators' evaluation of changes over time in both diesel engine emissions and patterns of use in the transportation industry. The estimated exposures from work histories and the industrial hygiene measures were combined with the following assumptions to develop the exposure-response relation.

- Ambient diesel exposure for workers in the trucking industry increased in proportion to the use of diesel engines.
- Past exposures to diesel emissions were estimated using heavy-duty trucks as a marker of diesel engine use and data on vehicle miles traveled (VMT) by heavy-duty trucks from 1949 to 1990.
- Worker exposures were assumed proportional to diesel engine emissions, as estimated from past engine emission levels and changes in emissions over time using existing data on grams emitted per mile traveled.
- The mid-point value of emissions decreased from 4.5 grams per mile in the 1970s to 0.4 grams per mile in the 1990s for new heavy-duty trucks.
- For the subjects in the epidemiologic study, past exposure to EC_1, as a marker for diesel exposure, was estimated by assuming that (1) the average 1990 level could be assigned to all workers in a specific job category; and (2) levels before 1990 were proportional to VMT by heavy-duty trucks and the estimated emission levels of diesel engines.
- Long-haul drivers received some exposure from their own trucks, increasing their estimated exposure by 50% (based on Ziskind et al. 1978).
- Ambient air background levels (1 $\mu g/m^3$/year) were added to the cumulative worker exposure.

Cumulative exposure estimates were calculated for workers by each year of work history and job category. The largest job category was long-haul drivers, and estimates of lung cancer risk were calculated for this group. All analyses were controlled for age, race, smoking, diet, and reported asbestos exposure. Results indicated that a lifetime excess risk of lung cancer death from exposure to 5 $\mu g/m^3$ EC_1 (through age 75) for a male truck driver was 1.6% (95% CI: 0.4, 3.1); that is, an excess of 1.6 deaths from lung cancer for each 100 men. (The excess risk at 1 $\mu g/m^3$ EC_1 is 3×10^{-3}.) The assumptions in the estimate were that emissions in 1970 were 4.5 grams per mile, the worker had 45 years of exposure (from age 20 to 65) at 5 $\mu g/m^3$ EC_1, and there was no lag time when the cumulative exposure was calculated. (Investigators reported that similar findings resulted when a lag was included.) Varying the assumption of level of exposure in 1970 resulted in an estimated range of lifetime risk from 1.4% (95% CI: 0.3, 2.7) to 2.3% (95% CI: 0.5, 4.6). (This range is higher than the unit risk estimate calculated by OEHHA from the railroad worker cohort data.)

Assessment of Railroad Worker and Teamster Studies

When the Panel started to evaluate the epidemiologic studies of diesel-exposed workers, one challenge was understanding why the exposure-response association from various analyses of the railroad worker cohort data produced apparently conflicting results. Another was whether and how to address the exposure-response analysis of the teamster data (Steenland et al. 1998), which was published after the Panel's work was under way, and had not been subjected to the same level of scrutiny as the railroad worker studies. The Panel concluded that both sets of studies should be systematically reviewed to weigh their strengths and limitations for QRA.

In general, the results of both sets of studies are consistent with findings of a weak association between death from lung cancer and occupational exposure to diesel exhaust. Although the secondary exposure-response analyses of the railroad worker cohort data are conflicting, the overall risk of lung cancer was elevated among diesel-exposed workers.

In addition to the availability of some associated exposure data, several strengths are evident in these railroad worker and teamster studies. Both included large numbers of subjects (55,407 employees with 1,694 lung cancers comprised the railroad worker cohort, and 1,256 lung cancers were the basis for the case-control study; the teamster case-control study was based on 996 lung cancers) from industries in which some occupations entailed exposure to diesel exhaust. Job categories with known exposure to asbestos were either excluded or controlled for in the analyses. Both of the case-control studies adjusted for cigarette smoking as a confounding variable in analyses of the relation between diesel exposure and lung cancer mortality.

The Panel addressed the questions posed in its framework (see Table 1 in the Background chapter and Appendix B) for evaluating the railroad worker and teamster studies. Because the questions represent an ideal study, the Panel did not expect a strict "yes or no" response to each question. Instead, the systematic evaluation of these studies within this framework was intended to inform the process for others who may want to understand or use these data, and to provide a means to weigh study strengths and limitations. This systematic approach also was intended to sharpen the focus on where limitations from previous work might help define research needs.

RAILROAD WORKER STUDIES

DESIGN

1. Was the study design efficient, and did it specifically consider power, potential types of bias, and latency?

A cohort study (Garshick et al. 1988) and a case-control study (Garshick et al. 1987) reported findings on railroad workers. The authors did not report how power calculations were conducted or whether factors other than overall association between diesel exposure and lung cancer were considered. For weak associations, even large studies may not be powerful enough to detect associations, particularly in subanalyses.

A cohort study design, as used by Garshick and colleagues (1988), is appropriate for enumerating the workers to be followed for development of disease. A case-control study design requires additional assumptions for use in QRA; when the case-control study is not nested in a cohort design (as it was not in Garshick et al. 1987), still more assumptions are required, which can limit the conclusions to be drawn. For example, it is possible that the deaths in the case-control study did not include all lung cancer deaths that occurred in the population eligible for RRB death benefits, because both cases and controls were deaths identified between March 1, 1981, and February 28, 1982, which was after the period when follow-up was considered complete in the cohort study as reported by Crump and colleagues (1991).

A case-control study may be less suitable than a cohort study for QRA; if the control causes of death used are related to smoking or to other risk factors for lung cancer, including the exposure of interest, this may lead to biased results. Control deaths (Garshick et al. 1987) were primarily from cardiovascular disease (CVD), which is also associated with cigarette smoking. If risk of CVD differed among job categories (that is, by exposure), and within category of smoking status, a selection bias may have been introduced. The authors did not discuss such a possibility

in their published report. A case-control study nested in the cohort study would be a preferable study design because it avoids such bias.

In the cohort study, subjects were followed for up to 21 years. This may have limited the power of the study by providing too short a latency period for the manifestation of lung cancer, given that the latency period from exposure to a substance until the development of solid tumors is often 20 to 40 years (although some can appear in 10 years or less).

2. Were adequate quantitative exposure and covariate data planned and collected?

Both railroad worker studies were retrospective studies conducted on the basis of death certificates, which obviously precluded interviewing subjects to obtain exposure data. However, an industrial hygiene survey of the railroad industry was conducted to measure RSP adjusted to remove the ETS fraction, as a marker for DPM (Woskie et al. 1988a). This survey was conducted at a nonrandom sample of four small railroads, and measurements were in a time period after the cohort members were exposed. The diesel-exposed or unexposed job categories in the epidemiologic studies, which had been designated on the basis of job title and duties, were confirmed with the industrial hygiene samples of railroad jobs (Woskie et al. 1988a,b). (For a more detailed discussion of exposure issues, see the Exposure Assessment section.)

3. Was the literature reviewed for relevant study methods and previous research findings?

The literature review in the main report of the cohort study did not include references related to QRA. Such references would not be expected, however, because a QRA was not originally intended.

4. Were plans made for postpublication data scrutiny?

The Panel did not know whether the researchers originally intended to share their data. The investigators did, however, generously share data with the U.S. EPA (and, through the EPA, with Dr. Crump). In turn, Dr. Crump shared the data with Dr. Dawson in the OEHHA (California EPA). Most recently, all of these individuals have shared data with the HEI Panel. This was extremely valuable in understanding the complexities inherent in this data set, and the Panel greatly appreciated all of their cooperation.

EXPOSURE ASSESSMENT

1. Were detailed lifetime histories of occupational exposures collected?

All members of the cohort were active railroad workers in 1959 and had at least 10 years' history of employment; investigators reported that no information on work histories was available for the years before 1959 (Garshick et al. 1988). This lack of data on exposures before 1959 adds to the uncertainty in quantitative risk estimates developed from studying this cohort. The researchers chose 1959 as the starting date because approximately 95% of railroad locomotives were diesel-powered (as opposed to steam-powered) by that time, according to a U.S. Department of Labor report (1972). Duration of employment since 1959 was used to represent the duration of exposure to diesel exhaust after 1959. Although the investigators had neither a detailed occupational history nor a history of exposure, they verified that more than 90% of workers who were between 40 and 44 years of age in 1959 remained, over their careers, in job categories with the same exposure classification. Garshick and colleagues (1988) accounted for exposure history in "diesel-years," that is, the total number of years in a diesel-exposed job category from 1959 until death or retirement.

2. Were known chemical and physical characteristics of the main exposure specified?

In the original cohort study, the main exposure of the railroad workers was considered to be "diesel exhaust," which is a complex mix of many compounds in both particle and gas phases. Little historical information was available on the chemical and physical nature of the diesel emissions from locomotives, or on how the emissions might have varied by railroad or job category over the time period covered in the cohort study. The industrial hygiene study, conducted between 1981 and 1983, focused on RSP as a marker for exposure to diesel exhaust, but diesel exhaust is only one of many sources of ambient particulate matter. The contribution of ETS, a major source of indoor particles, to the total particulate samples was adjusted for by a correction factor derived from measured particle-phase nicotine (Hammond et al. 1988). Levels of NO_2, a constituent of diesel exhaust, were examined at the four studied railroads across seasons, but this measurement was not useful for distinguishing among the job categories (Woskie et al. 1989).

3. Were potentially confounding exposures measured or estimated?

In the railroad worker studies, sources of exposure to respirable particulate matter in addition to diesel exhaust were likely to include outdoor nondiesel pollution, personal activities (e.g., smoking), indoor residential exposures (e.g., ETS or residential coal heating), and other occupational exposures (e.g., asbestos). These exposures, if related to both diesel exposure and lung cancer, could be considered possible confounders. That all of these exposures were confounding variables is unlikely.

Asbestos, however, is a potentially confounding exposure among railroad workers. The railroad worker cohort study excluded workers in job categories known to be "asbestos-exposed," and analyses were conducted both with and without the inclusion of job categories in which asbestos exposure was suspected. The case-control study controlled for asbestos exposure in the statistical analysis (Garshick et al. 1987). The possibility of confounding by other exposures before 1959 could not be completely evaluated.

Retrospective mortality studies have limited ability to obtain good data on risk factors such as smoking or lifestyle variables, which often must be obtained from next of kin. The case-control study included smoking data from next of kin; the investigators reported that when results were adjusted for smoking, they did not differ substantially from unadjusted results. Thus, it is likely that results were not seriously confounded by smoking. However, because smoking is strongly associated with lung cancer in most studies, and because the reported association between diesel emissions and lung cancer is weak, smoking needs to be controlled precisely in analyses of the effects of exposure to diesel exhaust. Even if smoking is not shown to be a confounder in a particular study, when a weak association between a risk factor and lung cancer is under investigation, the validity of any findings is likely to be questioned if the analysis did not control for smoking. In addition, smoking may modify the risk of exposure to diesel exhaust, and this modification can be investigated only if smoking data are available.

In the industrial hygiene study by Hammond and colleagues (1988), the ARP removed the ETS contribution. The adjusted marker (ARP) is an improvement over a crude classification of each job as either exposed or unexposed to diesel exhaust with no consideration of either smoking or ETS. When the ETS estimate was subtracted from the RSP concentration, ranking of the diesel-exposed job categories changed from that based on RSP alone.

Woskie and colleagues (1988b) developed a model for estimating mean levels of DPM exposure in national career groups. Although this model included data on ARP concentrations, railroad job category, and climate, it did not consider the impact of outdoor nondiesel particulate matter on estimates of exposure to DPM; instead, one estimate of the contribution of outdoor particulate levels was applied to all workers regardless of location or time. This procedure is questionable, however, because outdoor particulate mass concentrations show considerable geographic variability across the U.S., and over time; substantial reductions have been noted in measured levels. For example, the trend in particle emissions data indicates a threefold reduction in emissions between 1940 and 1982 (National Air Pollutant Emissions Trends 1997), a time that overlaps the period during which the railroad worker cohort was exposed. Considering background outdoor particulate levels over space and time, and by railroad, might substantially alter estimates of total particle exposures, and hence alter estimates of the health effects of diesel exhaust.

4. Were magnitude, duration, and variability of exposure determined?

Exposure data were collected by Woskie and colleagues (1988a,b) several years after the workers in the epidemiologic studies were exposed. The 13 job codes from the industrial hygiene studies were combined to develop the 5 career groups as shown in Table 3. (Arithmetic means are presented rather than geometric means, because comparable data were available in this form for each of the exposure estimates.) The authors discussed some concerns they had about the exposure data used in the model to estimate the national career group mean exposures (Table 3, last column). Particle mass measurements varied considerably among the job groups for both corrected (ARP) and uncorrected (RSP) concentrations. The uncertainty associated with the final grouping of occupations into these job categories should be considered when the exposures for each group are discussed.

Table 3 indicates that several job exposure groupings are possible, with all such groupings having an undefined but large uncertainty. For example, clerks (who are considered "unexposed") appear to have RSP exposures (RSP = 125) similar to those (RSP = 126) of freight conductors (who are considered "exposed"). Engineers and firers, and brakers and conductors are combined as "exposed"; however, their RSP estimates range from 75 (passenger engineers and firers) to 231 (hostlers). When ARP values are considered, clerks (ARP = 42) and signal maintainers (ARP = 58) appear to have more similarly low exposures than when RSP is used. However, passenger engineers and firers (ARP = 51), yard engineers and firers (ARP = 69),

Table 3. Railroad Worker Exposure Assessment Data[a]

Career Group	Job Group	Number of Workers Monitored[b]	Number of Samples Used for ETS Estimates[c]	Particle Concentration as Arithmetic Mean				Modeled National Career Group[h] Mean ± SD (μg/m^3)
				RSP[d] (± SD)	ETS[e]	ARP[f] (± SD)	AEM[g]	
Clerks	Clerks	59	36	125 ± 75	88	42 ± 36	7	33 ± 1
Signal maintainers	Signal maintainers	13	14[i]	69 ± 39	10	58 ± 33	23	58 ± 4
Engineers and firers	Freight	55	37	115 ± 67	18	94 ± 55	30	71 ± 3
	Yard	50	20	108 ± 109	44	69 ± 70	24	
	Passenger	23	21	75 ± 52	23	51 ± 35	16	
Brakers and conductors	Freight conductors	62	48	126 ± 65	52	69 ± 52	30	89 ± 3
	Freight brakers	21	16	145 ± 80	36	102 ± 62	8	
	Passenger	35	33	111 ± 62	6	104 ± 58	27	
	Yard	32	7	180 ± 117	75	114 ± 76	17	
	Hostlers	8	8	231 ± 134	7	224 ± 130	33	
Shop workers	Electricians	42	16 (Summer)	256 ± 332	37	192 ± 248	37	141 ± 8
	Machinists	110	32 (Summer)	191 ± 146	29	147 ± 120	55	
	Supervisors and other shop workers	24	12	244 ± 141	30	155 ± 83	43	

[a] Adapted from Hammond and colleagues (Table 2 in 1988) and from Woskie and coworkers (Tables 2, 4, and 5 in 1988a; Table 1 in 1988b).

[b] Number of workers monitored over one work shift with a personal monitor for particle mass (Woskie et al. 1988a).

[c] From Hammond and associates (1988).

[d] RSP = concentration (μg/m^3) of respirable particle mass (≤ 3.5 μm) measured by personal monitors for four railroads (Woskie et al. 1988a).

[e] ETS = respirable particle mass concentration (μg/m^3) associated with ETS (Hammond et al. 1988), determined from the analysis of composite personal monitoring particle mass samples by job group and railroad for three railroads.

[f] ARP (in μg/m^3) = (RSP – ETS)/volume of air sampled (Woskie et al. 1988a).

[g] AEM (in μg/m^3) = [(μg RSP × fraction extractable) – (μg ETS × fraction ETS extractable)]/volume of air sampled (Hammond et al. 1988).

[h] Modeled national career group exposure determined from model that incorporates ARP concentrations, national railroad worker job data, and climate data in a linear model (Woskie et al. 1988b).

[i] Numbers as reported.

and freight conductors (ARP = 69) also appear to have low-level exposures, although for analysis they are considered "exposed" along with those who are listed with much higher ARP levels. Such misclassification as to which jobs are considered "exposed" or "unexposed" could bias study results.

The variability in job exposure estimates may be the result of several assumptions that were made in the industrial hygiene study. First, the selection of a 10°C cut point to represent climate effects on DPM exposures is arbitrary and not well supported. Second, repeated personal monitoring on a subpopulation of workers in different job groups would have been preferable to a single measurement for estimating measurement error and day-to-day variability. Variability associated with these sources may add considerable uncertainty to estimated job group exposures. Concentrations of RSP in various job groups might have changed substantially over time in ways that could be related to newer engine technology, changes in ventila-

tion, differences in diesel maintenance practices, or other variables not examined. The model of career exposure means (last column of Table 3) would have been more informative for use in QRA if the model had considered various scenarios of time-varying concentrations.

5. Were industrial hygiene data and historical data on use and repair of machinery and equipment obtained?

Although Woskie and colleagues (1988b) obtained some information on national dieselization of the railroad industry, actual exposure measurements were made at four small railroads that were unlikely to represent the national diversity of equipment and exposures. Detailed historical exposure conditions in the railroad shops were evaluated in terms of NO_2 levels, historical records of use and changes in ventilation systems, and a comparison of records of historical locomotive use with locomotives in use at the time of the industrial hygiene study. Detailed historical information on the purchase and maintenance of diesel locomotives, fuel use, and repair shop design, helpful to the historical reconstruction of exposures, was not available. Such additional information might better define historical exposure trends by railroad.

6. Were personal exposure measurements obtained, and were they representative of the population studied?

The industrial hygiene exposure assessment provided information that identified categories of railroad workers who might be at risk of exposure to diesel particulate emissions from locomotives. As noted, however, the results in Table 3 indicate that other groupings of diesel-exposed jobs are possible, which introduces additional uncertainty in the exposure data for use in QRA.

The exposure assessment conducted by Woskie and colleagues (1988a,b) probably does not represent personal exposures of railroad workers to diesel exhaust in either space or time, and raises several issues. First, the four small railroads sampled by Woskie and colleagues are in the northern U.S. and are not necessarily representative of the national average railroad exposure, during either the period of sampling (1981–1983) or the period of exposure covered by the epidemiologic studies (1959–1980). Second, a convenience sample of 530 workers from 39 job codes was selected for personal monitoring of particle mass during a single work shift, which may not accurately represent exposures of all workers in those job codes. Third, corrections to particle mass exposures were determined from composites of personal samples within job group at each railroad (Hammond et al. 1988). Data on ETS were available from only three of the four railroads. Although the composite samples were necessary to have a sufficient mass of particulate matter for marker analysis, the few resulting observations did not allow assessment of the variability of the contribution of ETS and inorganic mass (e.g., sand, dirt, fibers) to the measured personal diesel exposures either by job group or by railroad. A statistically drawn random (e.g., stratified random) sample from railroads across the country would have provided greater confidence in the representativeness of the data.

7. Were uncertainties in exposure assessment quantified?

When Woskie and colleagues (1988a) developed their exposure intensity estimates for various job categories, it was not with the intention that the estimates would be used for QRA. Woskie and colleagues provided estimates of error resulting from sampling variation but did not quantify uncertainties from all sources, as would be desirable for results intended for use in QRA. Sources of uncertainty include (1) extrapolation of data from four railroads to the entire U.S., (2) use of results from 1980 to estimate earlier exposures, (3) lack of exposure histories prior to 1959, (4) failure to account for seasonal variations in exposures, (5) problems in selecting the appropriate job category groupings, and (6) use of respirable particulate matter as a surrogate measure of DPM. Analysts who have made use of the Woskie and colleagues data (Crump et al. 1991; OEHHA 1998) also have not quantified uncertainties from all sources, although these analysts discussed problems with the exposure measurements.

OUTCOMES

1. Were outcomes defined in specific and objective terms?

Death certified as primary lung cancer was defined as the outcome for the railroad worker case-control study (Garshick et al. 1987). In the cohort study (Garshick et al. 1988), the outcome included lung cancers mentioned either as the underlying cause of death or elsewhere on the death certificates. Lung cancers identified on death certificates can include false positive identification of disease, especially if a cancer is metastatic rather than primary in the lung, in addition to false negatives from certification of deaths to some other disease when lung cancer is the cause. That is, death certificates are likely to overestimate the number of lung cancer deaths by including metastatic sites. The best identification of lung cancer for incident cases would be by pathologic examination of a tissue specimen; however, this process is more

expensive than a death certificate review. If some cancers metastatic to the lung were differentially misclassified as primary cancers, the impact on risk estimates used in QRA could be in either direction.

2. Was the full range of outcomes included in cohort studies?

Garshick and colleagues (1988) reported only lung cancer in their cohort study, whereas Crump (1999) has looked at other mortality outcomes in association with exposure to diesel exhaust, including heart disease and stroke. These findings have shown decreased risk for overall mortality, as well as for cause-specific mortality, with longer duration of employment.

3. Were participants actively followed to determine outcome status and the date outcome occurred?

Crump and colleagues (1991) compared all-cause mortality in the railroad cohort with U.S. mortality rates, and showed that follow-up was incomplete after 1976. They found that although age-specific death rates in the cohort (overall job categories) remained fairly constant through 1976 (consistent with the U.S. pattern in this period), after 1976 death rates for the cohort dropped (Crump et al. 1991). Garshick reexamined the data and verified that mortality follow-up after 1976, as determined by the RRB, was incomplete (1991). In the cohort study, mortality was determined using the records of the RRB benefits plan. This appears to have been an incomplete method of mortality follow-up, and whether the incomplete follow-up is nondifferential by job category is unknown.

ANALYSIS

1. Were analytic methods specified a priori?

The original investigators were clear that they would use job category and employment duration to estimate exposures to diesel exhaust among job categories without attempting to make quantitative estimates. Although the Panel did not have the original protocol for this study, no evidence would indicate that the approach used was influenced by the hypothesized relation between potential exposure measures and risk of death from lung cancer. Crump and colleagues (1991) used quantitative data to distinguish among job categories and considered many different measures of exposure, all of which led to a similar conclusion; that is, the risk of lung cancer decreased with longer duration of employment. The OEHHA (1998) based most of its analyses of the cohort data on a metric of duration of employment and on the difference between train workers and clerks; however, it considered a variety of assumptions regarding increasing diesel exposure until 1959 and its gradual decrease afterward.

2. Was the appropriateness of the statistical approach demonstrated, and were potential biases explored?

The published models for the cohort and case-control studies did not report any validation procedures. Of possible concern in these models is that age in 1959 was treated as a continuous variable*, and no term was included for interaction with calendar year, which would have adjusted for attained age† and calendar year. Other analysts of the railroad worker data have used different approaches for adjusting for age and calendar year.

The OEHHA did not account for variations in risk by job category, and none of the analysts fully investigated the possibility that time patterns of lung cancer development might differ for various subgroups. The investigators did not publish a description of how *p* values and confidence intervals were calculated.

3. Were exposure-response relations statistically explored?

The conflicting results obtained by secondary analyses of the railroad worker data are central to the issue of whether these data should be used in a QRA of diesel exhaust. Although the railroad worker data have been used for QRA, Garshick (1998) has not supported this use.

The model used for the original railroad data analysis (Garshick et al. 1988) was limited, and neither goodness of fit nor attempts to quantify the risk as a function of exposure were explored. The analysis included a comparison of estimates by age in 1959, and interpreted the larger risks for the younger age groups as reflecting larger cumulative exposures for the younger subjects rather than as effect modification.

Results in the railroad worker study have been analyzed with either no lag time to allow for latency or induction of cancer, or with a five-year lag. The five-year lag was used by secondary analysts (Crump et al. 1991; OEHHA 1998; Crump 1999).

In the original analysis of the cohort railroad data (Garshick et al. 1988), exploration of alternative models was limited to an analysis considering a lag period and an analysis that excluded shop workers and hostlers who

* When age is treated as a continuous variable, the actual age of the subject is included in the analysis. If age is considered as a categorical variable, ages of the subjects are divided into either 5- or 10-year groups, for example, and a variable representing the age group is included in the analysis rather than the actual age.

† Attained age is the age at risk for a particular calendar period and would include everyone in the cohort who is at risk.

may have been exposed to asbestos. The OEHHA explored ramp and roof patterns of pre-1959 exposures, as well as several approaches for adjusting for age and calendar year period. Crump's original work (Crump et al. 1991) explored different models for adjusting for the effects of age, calendar year, and exposure groups.

In the analysis of the railroad worker cohort study (Garshick et al. 1988), with all job categories combined, relative risks appeared to increase with years of exposure to diesel exhaust (see Figure 3 in the chapter Summary of Railroad Worker and Teamster Studies); exposure in the year of death and the preceding four years was disregarded. This finding was important because it supported an exposure-response effect and stimulated interest in using these data for QRA.

The Panel has conducted its own limited analysis of the railroad worker data. The objectives of these analyses were specifically to assist the Panel in its task of verifying and better understanding previous analyses, and to clarify reasons for differences between the results obtained by Crump (see Figure 4; a negative exposure-response relation) and by OEHHA (see Figure 6; a positive exposure-response relation). The Panel's analyses were limited to these objectives only, and were not intended to be a complete exploration and evaluation of the railroad worker data or to provide a model for QRA. A detailed explanation of the Panel's methods of analysis, assumptions, and results is presented in Appendix C. The Panel also recognizes that its own analyses are subject to the same outside scrutiny as others' have received.

The Panel's data exploration indicates that overall, lung cancer risks for train workers, within each duration of employment group, were higher than those for clerks and signalmen; shop workers had intermediate risks (Figure 8). However, within the exposed groups (train workers and shop workers), lung cancer risk decreased with increasing duration of employment. Simple measures of exposure, determined by multiplying duration of employment by exposure intensities assumed to be constant for the clerks/signalmen and train worker groups, were also analyzed. With no adjustment for job category, a positive slope was obtained, which was statistically significant when zero intensity was assumed for the clerks and signalmen group. To clarify the extent to which this positive response was due to the difference in baseline risks for train workers compared with clerks and signalmen, the variable *GRP* to measure this difference was included in the model. When this was done, the direction of association with exposure became negative; this result reflects entirely the decreasing association with increasing employment duration for train workers, and demonstrates that the positive

Figure 8. Panel's analysis depicting consistently elevated risk of lung cancer for train workers compared with clerks for each time period, but decreasing risk by job category over duration of employment. See Appendix C for details.

exposure-response relation obtained, without adjusting for *GRP*, was entirely due to the baseline differences between the two job categories.

These patterns are not consistent with a monotonically increasing association between cumulative exposure to diesel exhaust and lung cancer risk. If risk increased consistently with increasing exposure, a positive trend with duration of employment would be expected for the exposed groups (including train workers), even if exposure magnitudes were incorrect (see Appendix C).

The Panel's analytic model that was similar to the model that served as the basis for the California OEHHA risk assessment did not fit the data; a strong improvement was seen in the fit of the model with the addition of a variable reflecting the difference between train workers and clerks/signalmen. Although some definitions of exposure used by the Panel and by the OEHHA were different, it seems unlikely that these differences affected the Panel's qualitative conclusions.

The Panel's conclusions regarding the railroad worker data analyses did not depend either on method of adjusting for age and calendar time or on the assumption about exposure patterns before 1959.

Crump and colleagues (1991) and Crump (1999) have reported various analyses of the railroad worker data, and have investigated duration of exposure since 1959 as well as more complex measures of exposure than those cited here. However, all results are likely to reflect the negative association of risk with increased duration of employment within job category groups. Crump (1999) has noted that the OEHHA positive slope is driven by the difference in risk between train workers and clerks, and the Panel's analyses confirm this. By showing results of analyses with

and without adjustment for job category (intended to indicate exposure intensity), the Panel has attempted to clarify the roles of job category and duration of employment in the differences in results obtained by the OEHHA (1998) and by Crump and colleagues (1991, 1999). An expanded discussion of the Panel's analyses is in Appendix C.

4. Were uncertainties in risk estimates quantified, especially those resulting from exposure measurement error?

All analysts of the railroad worker data have evaluated statistical uncertainties, although the assumptions underlying the normal approximations were not validated. Uncertainties in the risk estimates resulting from exposure measurement error were not explored by any of the analysts. However, this was not relevant for the cohort study (Garshick et al. 1988), because that study did not use quantitative measures of exposure. Crump and colleagues (1991) did not specifically address the issue of measurement error, but presented results on the basis of several definitions of cumulative exposure and examined the effect of excluding shop workers from the analyses. They also commented on the uncertainties in exposure estimates and the possibility that these uncertainties could mask or distort diesel-related effects. The OEHHA (1998) listed uncertainties in the mathematical aspects of modeling, and used several simple models to explore some of the variability due to uncertainty (see Appendix F of its report).

DISCUSSION OF USING RAILROAD WORKER DATA FOR QUANTITATIVE RISK ASSESSMENT

The original railroad worker studies (Garshick et al. 1987, 1988) reported an elevated overall risk of lung cancer with increasing years of exposure to diesel exhaust, when all job categories were combined; however, when the Panel analyzed the cohort data by duration of employment for each job category, a negative association was seen in each group. That is, among train workers, and to a lesser extent among shop workers, the relative risk of lung cancer does not appear to increase with longer duration of employment (Appendix C).

A negative exposure-response relation might be present in these data for several possible reasons: several types of bias could affect the data, alone or in combination, in such a way as to mask a true positive association. For example, results could be affected by unmeasured confounding variables, such as cigarette smoking, previous occupational exposures, or other sources of pollution, that might be associated with diesel exposure as well as lung cancer. As noted previously, smoking, which is strongly associated with lung cancer risk, could not be assessed directly in the cohort study (Garshick et al. 1988), and estimates of lung cancer risk from the case-control study might have been biased.

Exposure misclassification is another possible source of bias in these data, especially because a dichotomous assignment of diesel exposure ("yes" or "no" by job categories) is a crude method for determining exposure. Also, the year when workers were first exposed to diesel is unknown. If dieselization was 50% complete in the early 1950s, some workers in a job category could actually have more "exposure duration" than others in the same category, thus diluting a possible association among those exposed the longest. Dosemeci and Stewart (1996) have demonstrated the impact of misclassification in exposure categories on estimates of relative risks. If misclassification is random (that is, there is no overall tendency for over- or underestimation of exposure), estimated risks are likely to be lower than their "true" values. Their findings show that the magnitude of bias depends strongly on the proportion of misclassified subjects. Job categories of exposure overlap (Table 3) and such a multidirectional misclassification of exposure of an unknown proportion of workers seems likely.

Still another possible reason for the negative association might be the use of "duration of employment" as a measure of exposure. Calendar time and duration of employment are highly correlated, and separating out the effect of duration of employment could be difficult. Doll (1985) has reported examples of nonmonotonic increases in cancer risk with longer durations of employment in an occupationally exposed job, and indicates that using this variable alone could lead to findings requiring cautious interpretation.

One more possible source of bias in these data is the "healthy worker survivor effect" (Arrighi and Hertz-Picciotto 1994). That is, workers who are "healthier" and less susceptible to disease might stay in the work place longer, so that those employed for longer periods might show a smaller elevation in risk than those employed for a shorter duration.

Bias also could be introduced if lung cancer deaths were differentially or incompletely ascertained. If lung cancer were more likely to be underascertained among those employed longest in diesel-exposed jobs, the result would be a lower risk with longer duration of employment. However, such differential ascertainment seems unlikely.

The preceding critique emphasizes the challenges involved in correctly analyzing and interpreting these railroad worker data. The Panel's opinion is that discussion of

the uncertainties in these data should accompany any presentation of a quantitative risk estimate. However, despite the reason or reasons why the relative risks in these data decrease with duration of employment, the lack of a positive exposure-response association in the railroad worker cohort data substantially weakens that study's potential to provide a reliable quantitative estimate of risk of exposure to diesel engine emissions.

TEAMSTER STUDIES

As indicated, the teamster exposure-response analysis (Steenland et al. 1998) was published only after the Panel started working; therefore, the evaluation of this set of studies is less extensive than for the railroad worker studies. Also, the Panel did not request data from these investigators, thus its assessment was made on the basis of the published reports only.

DESIGN

1. Was the study design efficient, and did it specifically consider power, potential types of bias, and latency?

The teamster case-control study identified cases and controls from among members of the Teamsters Union who filed for pension benefits after at least 20 years in the Union, and who died in 1982 and 1983. The cases were deaths from lung cancer, and the published study (Steenland et al. 1990) did not indicate the distribution of causes of death for the controls. It is likely that, as in the railroad study, control causes of death were largely related to CVD. If this is the case, some control causes of death may have been related to diesel exposure, smoking, or both. Bias resulting from control subjects used would affect the estimate of risk and any QRA based on the study results.

In the teamster study, mechanic was the job category with the highest assigned diesel exposure; intermediate were long-haul and short-haul truck drivers; dock workers were assigned low exposure; and the unexposed category was composed primarily of dairy workers. It is not known whether the risk of dying from any of the control causes of death differed among the job categories or within strata of cigarette smoking, because of differences in diet, physical activity on the job, or other factors. If the risk differed, study findings could be biased.

Workers had to have 20 years of tenure in the industry to be eligible to apply for pension benefits, and data analyses included models with either 1960 or 1965 as the start of a worker's exposure to diesel exhaust. Deaths were ascertained between 1982 and 1983. If the development of lung cancers has a latency period of 20 years, it is likely that only the minimum time period passed for the development of the tumors, because the latent period can be as long as 40 years.

2. Were adequate quantitative exposure and covariate data planned and collected?

Next of kin provided data on smoking, diet, work history, and asbestos exposure for cases and controls. The investigators used these data to evaluate and control for possible confounding effects of these exposures on the association between diesel exhaust exposure and lung cancer. Work history data also were included in the pension applications available in the Teamsters Union records.

An industrial hygiene survey (Zaebst et al. 1991) measured EC_1 as a marker for diesel exhaust. The samples were taken in a time period after cases and controls would have been exposed, and might not represent actual exposures in space and time. However, the estimated exposures of teamsters to diesel exhaust are of greater relevance to public health than those of the railroad workers, because the teamster exposures are closer to the range of ambient levels of diesel emissions.

3. Was the literature reviewed for relevant study methods and previous research findings?

The literature reviews in the teamster studies are partial and brief. Because the original case-control study was not designated for QRA, a review of relevant methods would not be expected.

4. Were plans made for postpublication data scrutiny?

Planning for postpublication data scrutiny generally is not a major consideration of researchers. However, understanding of the railroad worker data has improved with each additional review. If the teamster data also are reassessed by the original investigators and other analysts, this is likely to yield additional insights and understanding. The exposure-response findings are still "young" in the literature and can benefit from critical peer review, examination by other researchers, and further exploration of possible biases.

EXPOSURE ASSESSMENT

1. Were detailed lifetime histories of occupational exposures collected?

Work histories were obtained for cases and controls from two sources, next-of-kin interviews and teamster pension applications. Next-of-kin interviews, administered by phone (20%) and mail (80%) to spouses (76%)

and others (24%), included information on work histories and several potential confounders, including smoking, diet, and asbestos exposure. Workers were classified into the job (exposure) category in which they had worked the longest. The categories were truck driver (diesel, gasoline, or both), diesel truck mechanic, and dock worker. The uncertainties associated with using questionnaire data and the impact of those uncertainties were not assessed. Although most pension applications reported that the worker had worked primarily in one job category, the impact of changes among job categories was not explored.

Work histories in the teamster pension applications, including all teamster jobs, were self-reported. These were also reviewed and used to assign workers into job (exposure) categories, although records did not differentiate between diesel and gasoline truck engines. Again, workers were assigned to the category in which they were employed the longest. The Teamsters Union assigned a U.S. census code for occupation and industry for each job a man included in his pension application. The four main occupations, based on records, were long-haul drivers, short-haul or city drivers, truck mechanics, and dock workers. The uncertainties associated with the exposure categories derived from the teamster records, and from combining the information from both the teamster records and the interviews, were not explored.

2. Were known chemical and physical characteristics of the main exposure specified?

The case-control study did not include estimates of exposure, but a companion industrial hygiene study was conducted after the epidemiologic study period ended. Steenland and colleagues (1992) derived job-specific exposure estimates from industrial hygiene measurements of EC_1 made by Zaebst and colleagues (1991). EC_1 is a reasonable marker for DPM, at least in terms of establishing relative exposures. It is more specific to diesel exhaust than RSP and relatively free of interference from ETS, an important particle source. On the basis of EC_1 measurements, Zaebst and colleagues (1991) provided some useful insights into the nature of diesel exposure for teamsters.

Relating EC_1 to DPM can be complicated, because the EC_1 fraction of DPM is variable and has probably increased over the period of exposure (Sawyer and Johnson 1995). This matter needs to be explored further if a quantitative risk is assigned. A carcinogenic fraction in DPM has not been identified, and there is no assurance that the proportion between that fraction and DPM or EC_1 has been constant. However, EC_1 is probably the best marker for DPM available at this time.

Nondiesel sources of elemental carbon and their ambient concentrations have changed over time, and geographic differences in concentration probably exist as well. Other sources of elemental carbon include gasoline engines, tire and brake wear, stationary combustion sources, and industrial processes; nondiesel sources were not considered in the assessment by Steenland and colleagues (1998). Although elemental carbon from stationary sources may have decreased over time, a recent study reports a substantial elemental carbon contribution from gasoline engine emissions that should be further considered (Northern Front Range Air Quality Study [NFRAQS] 1998) in the retrospective assignment of exposures by use of a marker for a complex mixture.

3. Were potentially confounding exposures measured or estimated?

In the original teamster study, next of kin of the cases and controls were interviewed about the subject's smoking, dietary, and asbestos exposures. Work histories included employment in any previous diesel-exposed job. The analysis used information from next of kin to control for smoking and asbestos exposure. Spatial and temporal variations in individual exposures to outdoor nondiesel respirable particles and residential indoor air were not assessed. However, these exposures would be of concern only if they related to both diesel exposure and lung cancer.

4. Were magnitude, duration, and variability of exposure determined?

Several aspects of exposure estimation could have considerable impact on the exposure-response analysis. First, in the teamster case-control study (Steenland et al. 1990), exposure was determined by assignment to one of five job (exposure) categories. Each worker was assigned to the category in which he was employed the longest. This method of assigning exposure necessarily restricts a subject to one job, and does not allow for the use of the complete work history or the potentially wide variability of exposure among workers within a job category over time and in different locations. Thus, information on the magnitude, duration, and frequency of exposure to diesel exhaust by the workers was not directly collected or assessed.

Second, recent in-use measurements suggest that newer engines on the road have higher emissions than the engine measurements used by Steenland and colleagues (1998). Extrapolations from recently reported measurements of particulate emissions from in-use heavy-duty vehicle diesel engines suggest that emissions were about 5 grams per mile through 1980 and then began to fall as new tech-

nology was introduced, with the fleet average falling to about 2 grams per mile in 1990 (Graboski et al. 1998). This differs from Figure 3 in Steenland and colleagues (1998), which shows lower emission estimates for each of three different scenarios of changing emissions over time.

Third, dieselization was assumed to begin in 1949. However, the proportion of the fleet that was diesel-powered increased gradually until the 1970s, by which time most heavy-duty trucks sold were diesel-powered. Exposures to diesel engine emissions in the early years would not have been as great as later in the study period.

Fourth, the emissions data for heavy-duty trucks were linked to the industrial hygiene data for the various job categories, and this exposure was scaled to heavy-duty truck VMT. These assumptions need to be explored further. Linking emissions from heavy-duty trucks to off-road job categories, such as dock workers and mechanics, is probably not the most appropriate method to estimate those workers' exposures. Working environments on loading docks and in repair shops, particularly ventilation and temperature, can affect exposure estimates and are not accounted for in this approach to estimating exposure. Also, increasing VMT does not necessarily mean increasing exposures for these workers; more likely it means hiring more workers handle the additional work. How the heavy-duty truck emissions and VMT assumptions interact and affect exposure estimates applied to other job categories requires further investigation.

Fifth, the estimation of relatively high roadside and background exposures may overestimate ambient exposures.

In summary, the estimated historical exposures to diesel emissions are associated with several major uncertainties (Steenland et al. 1998), the effects of which are difficult to estimate. The impact of varying the uncertainties associated with the assumptions needs to be addressed.

5. Were industrial hygiene data and historical data on use and repair of machinery and equipment obtained?

Historical exposures (Steenland et al. 1998) were estimated by combining an estimate of reduction in emissions from improved technology and fuels with the increase in emissions from the increased use of diesel engines and fuel. The reduction in emissions was taken from estimates made by Sawyer and Johnson (1995) of new engine particulate emissions for three broad time periods, the 1970s, 1980s, and 1990s. These data were incorrectly applied to the diesel truck fleet; because vehicle turnover time is long and particulate emissions increase with vehicle age, the fleet average emissions (which determine exposure) are always greater than the new engine technology. In-use truck emissions data are now available that could provide a better estimate of truck emissions (Graboski et al. 1998). Included are emission measurements from trucks that predate emission controls; these data could possibly be useful in establishing emissions relevant to the 1959–1983 exposures studied by Steenland and colleagues (1990).

6. Were personal exposure measurements obtained, and were they representative of the population studied?

The industrial hygiene study (Zaebst et al. 1991) provided information useful in identifying job categories in which teamsters might be at risk of exposure to diesel exhaust. Researchers developed a specific marker for diesel exhaust, EC_1, and demonstrated that ETS, a major source of respirable particle mass, did not interfere in its measurement.

The specific methods by which they selected breakbulk truck terminals and individual teamsters for air sampling were not described. The investigators appear to have collected a convenience sample rather than a random sample, so any inferences applied from these results to truck terminals and to teamsters across the country are uncertain. A statistically drawn sample from terminals and teamsters nationwide would have been desirable, although perhaps not feasible.

Individuals and sites were not sampled repeatedly, which makes it difficult to assess individual or fixed-site variability in exposure. Long-haul drivers in this study may not be representative of all such drivers, because only "short turn-around" drivers were sampled (drivers who returned to the originating terminal 10 to 12 hours after departing from it). Finally, the cut point of 10°C used in this analysis to represent climate effects on diesel exhaust exposure for teamsters is not well supported, and may introduce considerable uncertainty.

7. Were uncertainties in exposure assessment quantified?

The investigators used three different emissions models to represent changes in exposure over time, and they found that their primary conclusions did not change (Steenland et al. 1998). The major sources of error in estimating exposures to diesel exhaust have not been identified or critically evaluated. There is likely to be considerable uncertainty associated with the exposure categories into which workers were placed, and that uncertainty is likely to affect the findings of the case-control study as well as a QRA. The Panel did not extensively evaluate the

assumptions in the report because of time constraints; however, this is an aspect that others can pursue.

OUTCOMES

1. Were outcomes defined in specific and objective terms?

The investigators clearly described the outcome as lung cancer identified on death certificates as an underlying or contributory cause of death.

2. Was the full range of outcomes included in cohort studies?

This is not applicable, because the teamster study was a case-control study of lung cancer.

3. Were participants actively followed to determine outcome status and the date outcome occurred?

This is not a cohort study. However, if personnel records or deaths of teamsters were missing from the Central States Teamsters Union files, results could have been biased. Information on the completeness of these files was not given in the original publication of the case-control study (Steenland et al. 1990).

ANALYSIS

1. Were analytic methods specified a priori?

There is no indication that the results obtained by using various exposure measures were explored before exposure-response analyses were conducted or that observed outcomes influenced the analyses.

2. Was the appropriateness of the statistical approach demonstrated and were potential biases explored?

Steenland and colleagues (1990) adjusted their statistical analyses by introducing categorical variables for age, race, smoking status, diet, and reported asbestos exposure. It appears that possible interactions among these variables were not investigated. All deaths occurred in 1982 and 1983, so there was no need to adjust for calendar year.

3. Were exposure-response relations statistically explored?

In the Steenland and colleagues (1998) analysis of an exposure-response association, average emissions per diesel vehicle were combined with historical information on heavy-duty truck VMT and the 1991 industrial hygiene measurements (Zaebst et al. 1991) to estimate past exposures for each job category. This assumption implies that estimating exposures in this way appropriately reflects exposures in the various job categories. However, the uncertainties of the assumptions involved were not adequately justified. Although the estimated exposures were reasonably proportional to fleet average emissions, it is not clear that they scale well with VMT, especially for off-road occupations. One might expect that the exposures of dock workers and mechanics would be independent of VMT, but related proportionally to the number of trucks coming to the dock and the number needing work by mechanics, if each truck driver is driving the same average distance per day. (The potential problems with the exposure estimation are discussed under Exposure Assessment, Question #4.) How the various assumptions regarding emission levels, timing of dieselization, and use of VMT interact and vary with, relate to, and affect the exposure-response analysis was not explored. It is possible that once the assumptions used to estimate exposure are reviewed and evaluated in the exposure-response analysis, the net impact of over- and underestimations will in fact be small; however, this issue needs to be explored. Figure 9 depicts the increasing use of diesel fuel as a fraction of all fuels used by vehicles since 1949.

To account for fleet turnover, the investigators conducted separate analyses with both zero- and five-year lags, and found similar results (Steenland et al. 1998). Actual fleet turnover time is likely to be much greater than five years, however; therefore, longer lag periods need to be explored. Steenland included a large number of covariates and reported goodness of fit of the final model, but effect modification was not explored. There is only limited discussion of how using other exposure-response functions might affect risk assessment, although the shape of the exposure-response function was evaluated in detail.

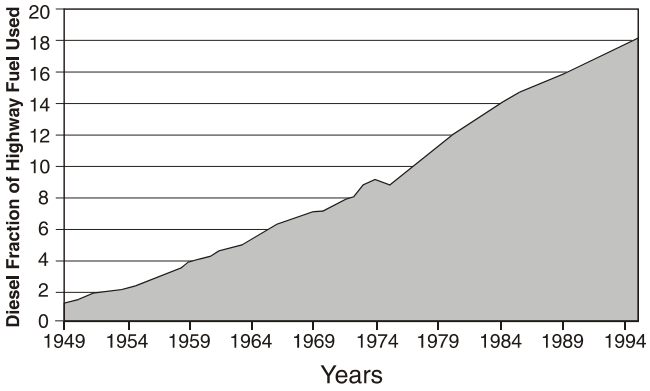

Figure 9. The increasing use of diesel fuel as a percentage of all fuels used by highway vehicles since 1949 in the U.S. (Adapted from Davis 1998.)

4. Were uncertainties in risk estimates quantified, especially those resulting from exposure measurement error?

The investigators discussed various sources of uncertainty, but did not evaluate them quantitatively, nor did they develop an overall quantitative estimate of uncertainty from all sources. How uncertainties in the emissions and exposure measurements propagate through the exposure model is complex, but methods exist to assess such uncertainty.

To a limited extent, the dependence of the estimated exposure-response relation on assumptions about the exposure model was explored by conducting analyses with increased exposures for long-haul drivers, and by using three assumptions about the decrease in emissions since 1970. The investigators' publication does not include analyses that fully reflect most potential biases and uncertainties in exposure estimates; however, it concludes with appropriate statements about the uncertainties in exposure estimates.

DISCUSSION OF USING TEAMSTER DATA FOR QUANTITATIVE RISK ASSESSMENT

The Panel concludes that a strength of the exposure-response analyses of the teamster data is in the apparent relevance of the diesel exhaust exposure levels to those of the general population (Zaebst et al. 1991), indicating that truck drivers are exposed annually to about 4 $\mu g/m^3$, roadside exposures are about 2.5 $\mu g/m^3$, and residential exposures are about 1 $\mu g/m^3$. The estimated annual exposure of long-haul truck drivers is not far above that of the general population. Heavy-duty trucks are one of the major sources of the general population's exposure to diesel emissions, so the assessment of risk for truck drivers is relevant to assessment of risk for the general population. Exposure assessment on the basis of elemental carbon provides a measure that is a reasonable marker for DPM, an improvement over previous assessments utilizing total respirable particulates.

The Panel believes that the assumptions used in the exposure assessment (Steenland et al. 1998) should be extended and validated, particularly to account for variations in diesel exhaust levels over time and improved estimation methods. The quantitative exposure portion of this study is new to the literature and, unlike the railroad worker studies, it has not undergone the extensive post-publication peer review that is required to develop a fuller understanding of the data. In fact, Steenland and colleagues (1998) concluded their findings by stating: "Our results should be regarded with appropriate caution because our exposure estimates are based on broad assumptions rather than actual measurements."

The Panel has some specific concerns regarding the exposure-response analysis assumptions and how the associated uncertainties in each assumption affect estimates of exposure. First, estimating exposure using previous 1990 emissions data may underestimate exposures given recent data from in-use vehicles that indicate emission levels of new diesel engines may in fact be higher than previous measurements made from an engine dynamometer, which were used in the study by Steenland and colleagues. Second, determining the onset of dieselization needs to be reconsidered because fewer diesel vehicles on the road in the early years would mean proportionally less diesel exhaust exposure during that time.

Third, the degree to which VMT accurately reflects the proportion of exposure for the various job categories needs more detailed exploration. More VMT by heavy-duty vehicles does not necessarily mean that exposures for dock workers and mechanics working on those trucks increase proportionately with the extra miles traveled. Instead, the VMT increase is likely to mean more dock workers and mechanics, rather than higher exposures per person.

Fourth, the analytic models did not explore the use of longer lag periods to account for the slow turnover of vehicles in use. A five-year lag is unlikely to be sufficient to account for fleet turnover, which may take closer to 20 years. Fifth, nondiesel sources of elemental carbon were not considered in the exposure assessment. Although elemental carbon from stationary sources may be decreasing in the air over time, recent findings indicate a substantial elemental carbon contribution from gasoline engines that should not be ignored (NFRAQS 1998). Finally, a background level of exposure was fairly similar to levels of exposure for the drivers, and sorting out the differences could be difficult.

The Panel speculated that controls selected for use in the case-control study of teamsters might have been biased in ways that could affect the findings. Exploring the distribution of causes of death among controls by smoking and exposure status could help clarify the possibility of selection bias and confounding in the data. Although no direct evidence of such bias is apparent, the Panel offers hypothetical examples showing the effect of potential bias in case-control studies with deaths from "other causes" (depicted as CVD because most of the "other causes" were probably CVD deaths) under different scenarios in Appendix D. Even though these examples suggest that case-control studies that use "other" causes of death could produce misleading results, it is important to

note that they are only hypothetical examples. Whether or not such bias is present will require further examination.

The postpublication peer review of a potentially important study can be a long, laborious, and uncertain process. It took several years of work by both the original investigators and independent analysts to learn much of importance about the railroad worker study. This process can and should be accelerated for the teamster study, particularly given the apparent relevance of the data to ambient exposures. When reviewed by the original investigators and other analysts, these data may reveal new issues and raise new points for discussion in the scientific and risk assessment communities.

Findings and Recommendations for Future Research

The Diesel Epidemiology Expert Panel was charged to (1) review the epidemiologic data that form the basis of current QRAs for diesel exhaust, (2) identify data gaps and sources of uncertainty, (3) make recommendations about the usefulness of extending or conducting further analyses of existing data sets, and (4) make recommendations for the design of new studies that would provide a stronger basis for risk assessment. The Panel was not charged to evaluate the broad epidemiologic literature concerning exposure to diesel exhaust and lung cancer for hazard identification purposes. The Panel's review of the epidemiology focused on the railroad worker (Garshick et al. 1987, 1988) and teamster (Steenland et al. 1990, 1992) studies because these studies have been used or considered for use in QRA, and other epidemiologic studies did not have associated quantitative exposure data. The Panel also reviewed published exposure-response analyses of the railroad worker data (Crump et al. 1991; OEHHA 1994, 1998; Crump 1999) and teamster data (Steenland et al. 1998), along with published industrial hygiene studies (Hammond et al. 1988; Woskie et al. 1988a,b; Zaebst et al. 1991). The Panel's findings are based on examination of these studies only, and not on the entire epidemiologic literature in this area.

The Panel recognized that no epidemiologic study can be perfect. Therefore, the Panel viewed its task as addressing the question: To what extent can limitations in the design and performance of a particular study affect its contribution to the body of epidemiologic knowledge under examination for QRA? The Panel also recognized that frequently it is very difficult to obtain retrospective data for estimating job-related work exposures, and that this process may require assumptions that cannot be validated. In the studies considered here, which form the core of the Panel's review, investigators made reasonable attempts to reconstruct past exposures to diesel engine emissions, using approaches that were feasible when the studies were conducted. These data have subsequently been used, in some cases, for purposes that were not envisioned by the original investigators.

Evaluation, reanalysis, and scrutiny of research reports bring new understanding. This report, including the evaluation, findings, and recommendations, is not meant to be the "last word" in this matter. Instead, it is meant to inform HEI and others about data needs and future research directions, including where improvements can be made on the basis of what has been learned from the limitations of currently available studies.

Important work is currently under way to study the effects of exposure to diesel exhaust among nonmetal miners in Germany (Säverin et al. 1998) and in the United States (NCI-NIOSH 1997). These studies were not reviewed because they are still in progress. However, the Panel heard presentations from these investigators at the HEI Diesel Workshop: Building a Research Strategy to Improve Risk Assessment (HEI 1999) at Stone Mountain, GA, March 7–9, 1999. In particular, the NCI-NIOSH study is large and appears to be well designed and comprehensive. It includes a cohort and nested case-control component, as well as extensive measurements of current exposure to diesel exhaust, detailed reconstruction of historical exposure, and biomarker development. When completed, these studies are likely to inform hazard identification, exposure estimation, and exposure-response analyses, all components of risk assessments.

The Panel recognizes that regulatory decisions need to be made in spite of the limitations and uncertainties of the few studies with quantitative data currently available. The findings described here and the systematic evaluation of these studies are designed to provide a means to weigh a study's strengths and limitations and to inform the QRA process.

FINDINGS

GENERAL

Enhanced exposure and epidemiologic data and analyses are needed for the purposes of QRA; these might come from further exploration of existing studies or from new studies.

RAILROAD WORKER STUDIES

At present, the railroad worker cohort study (Garshick et al. 1988), though part of a larger body of

hazard identification studies, has very limited utility for QRA of lifetime lung cancer risk from exposure to ambient levels of diesel exhaust for the following reasons.

- The various exposure-response analyses are limited by the scope and quality of currently available exposure data. Quantitative exposure data were not obtained during the cohort study period. Also, there is a paucity of qualitative data on individual exposures before 1959, and on the variation in exposure by railroad site, by season, and over time. The potential impact of concurrent exposures (for example, to grease, dust, other fumes, asbestos, and active and passive cigarette smoke) was not examined in depth. The diesel exhaust exposure data are suitable for a crude categorical measure of exposure by job category; but other measures, including duration of employment in a job category exposed to diesel exhaust, intensity of exposure concentration ($\mu g/m^3$), and lifetime exposure ([$\mu g/m^3$]-years), are not adequate to support quantitative exposure-response analyses.
- The Panel's analysis of the exposure-response association in the railroad worker data showed that the evidence for a positive association of lung cancer with cumulative exposure to diesel exhaust depends entirely on differences in risks among job categories. Train workers (with higher exposures) have higher risks compared with clerks (with low or no exposure). However, within all job categories, the relation of lung cancer risk to duration of employment is negative.
- Factors that might explain a negative association between duration of employment and lung cancer in these data include bias introduced by systematic differences in exposure misclassification among and within job categories; differentially incomplete ascertainment of lung cancer deaths by job category; lack of information on other occupational exposures and air pollutants; the presence of a healthy worker survivor effect; confounding by cigarette smoking; and analysis of relative risks rather than absolute risks. Also, in a case-control study, if causes of death among controls were associated with exposure to diesel exhaust, smoking, or both, the results could be biased.

TEAMSTER STUDIES

The investigators' analysis of the teamster data reported an exposure-response relation (Steenland et al. 1998) that may be useful for QRA; this relation will be better understood with further exploration of uncertainties and assumptions, particularly those relating to the reconstruction of past exposures and the selection of controls. Exposures of teamsters are more similar to ambient exposures of the public than are exposures of railroad workers, and the diesel exhaust to which teamsters are exposed comes from a source that is likely to be relevant to regulatory issues.

The Panel reviewed the teamster study without the benefit of additional analyses and interpretations, and its comments are not as detailed as those about the railroad worker studies. Understanding the teamster study will evolve with time; however, some conclusions can be drawn now.

- The set of teamster studies may provide reasonable estimates of worker exposure to diesel exhaust, but significant further evaluation and development are needed. The marker for diesel exhaust that was selected for study by Steenland and associates, EC_1, is more sensitive and specific than RSP adjusted for environmental tobacco smoke, but has several limitations (e.g., the contribution of diesel emissions to ambient EC_1 concentrations has not been constant over time). The industrial hygiene study, which was conducted after the period when workers in the case-control study were exposed, identified a range of exposures for various job categories, but did not consider (1) site-to-site variations, (2) seasonal variations, (3) concurrent exposures to other agents, (4) historical ambient particle concentrations, or (5) intra- and interindividual variability. The estimation of historical exposures needs to incorporate recent data on diesel emissions from vehicles in use, reassessment of when dieselization occurred, alternatives to estimating exposure by vehicle miles traveled, and historical regional ambient pollution data.
- The exposure-response relation reported in the teamster study increases in a linear manner. However, more can be learned from other analysts examining these data using different approaches.
- Neither a roster of the study population nor an alternative method of selecting controls to represent it was available to the researchers. It cannot be established with certainty whether the causes of death used for controls adequately represent the joint distribution of exposure to diesel exhaust and smoking in the case-control study. If smoking, or diesel exhaust exposure as determined by job category, or both were associated with causes of death used for controls, results could be biased.

RECOMMENDATIONS

The Panel's recommendations reflect its general understanding, as expressed in its framework for evaluating studies, of what constitute adequate data for QRA. They also reflect the preceding evaluation of the studies of railroad workers and teamsters. The Panel is aware that research currently in progress will respond to some of these research needs; however, results are not yet available, and it is not yet clear whether all of the proposed needs will be met.

COMPLETED STUDIES

1. The Panel recommends against using the current railroad worker data as the basis for QRA in ambient settings.
2. Further scrutiny of the teamster data, including estimation of uncertainty in both the exposure estimates and selection of controls, is recommended in order to improve the use of these data in QRA. Strengths of the teamster study include the relevance of exposure levels to the general population and the use of an exposure marker for diesel engine emissions that was an improvement over RSP. The teamster study exposure-response analysis is relatively new, and its further review and analysis by both the original investigators and others should be accelerated. Alternative retrospective exposure models need to be developed that use the alternative assumptions described above and in more detail in the body of the text.

NEEDS FOR NEW TECHNIQUES AND DATA

3. Better measures of exposure to constituents of diesel emissions, with careful attention to selection of the sample studied, are needed. Of particular importance are the selection and validation of a chemical marker of exposure to the complex mix of diesel exhaust emissions. 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.
4. 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.
5. Although biomarker technology was not available when the studies reviewed were conducted, appropriate, validated, and specific biomarkers of diesel exposures, health outcomes, and susceptibility are needed.

DESIGN NEEDS FOR NEW STUDIES OF EXPOSURE-RESPONSE ANALYSES

6. Exposures should be adequately and accurately characterized with respect to magnitude, frequency, and duration, rather than solely by duration of employment. 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.
7. 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.
8. 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.

NEEDS FOR NEW STUDIES

A prospective epidemiologic study of the development of lung cancer in exposed and unexposed individuals could have many strengths. Information on confounders and exposures could be more complete than for a retrospective study, and many of the biases and uncertainties discussed in this report could be eliminated or reduced. These advantages, however, need to be weighed against the disadvantages, which include high costs and a long period of follow-up. Other study designs that include retrospective components, are possible for a new epidemiologic study of lung cancer, but they are likely to include uncertainties and sources of bias that investigators will need to explore completely and acknowledge in their reporting.

9. The Panel recommends that a new, large, epidemiologic study of diesel exhaust emissions and lung cancer be considered after (1) currently ongoing or existing studies, including HEI's feasibility studies (to be completed in the spring of 2000), are evaluated, and (2) attempts to retrofit improved exposure assessments to existing epidemiologic studies are evaluated, including whether they can provide sufficiently accurate, complete, and relevant exposure data to support QRA.
10. Studies of lung cancer risk in general populations exposed to ambient diesel exhaust particulate matter will be difficult to conduct; however, such studies could usefully investigate other, noncancer health effects that occur in a shorter time after exposure.

Appendices

Appendix A. Diesel Epidemiology Expert Panel Workshop

WORKSHOP AGENDA FOR APRIL 20, 1998

Introductory Remarks *John Bailar* and *Kathleen Nauss*

Summary of the Design and Results of the U.S. Railroad Worker Studies and Issues Raised Using These Data for Dose-Response Analysis *Eric Garshick* and *Thomas Smith*

Questions on the Study Design and Results of the Railroad Worker Study

Analyses of Dose-Response Based on the U.S. Railroad Worker Cohort Study

- Discussion of Key Assumptions and Main Findings of Original Analysts *Kenny Crump* and *Stanley Dawson*
- Questions for Drs. Crump and Dawson
- Perspectives of Other Analysts *Duncan Thomas* and *Suresh Moolgavkar*
- New Analyses *Leslie Stayner* and *Dale Hattis*
- Questions for Drs. Stayner and Hattis

General Discussion of Approaches for Developing Dose-Response Estimates for Diesel Exhaust and Lung Cancer Based on the Railroad Worker Studies

Alternatives to the Railroad Worker Studies: U.S. Truck Driver Studies *Kyle Steenland*

- Questions and Discussion of Truck Driver Analysis

General Discussion

WORKSHOP PARTICIPANTS

DIESEL EPIDEMIOLOGY EXPERT PANEL

John C. Bailar, III (Chair)
Department of Health Studies, The University of Chicago, Chicago, IL, and the HEI Health Review Committee

Ethel Gilbert
Radiation Epidemiology Branch, National Cancer Institute, Rockville, MD

David Hoel
Department of Biometry and Epidemiology, Medical University of South Carolina, Charleston, SC

Brian Leaderer
Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, CT

Charles Poole
Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, NC

Robert Sawyer
Department of Mechanical Engineering, University of California, Berkeley, CA, and the HEI Health Research Committee

G. Marie Swanson
Cancer Center and Department of Family Practice, Michigan State University, East Lansing, MI

INVESTIGATORS

Eric Garshick
Brockton/West Roxbury Veterans' Medical Center, West Roxbury, MA

Thomas Smith
Harvard School of Public Health, Boston, MA

Frank Speizer
Harvard Medical School, Boston, MA, and the HEI Health Research Committee

Kyle Steenland
National Institute for Occupational Safety and Health, Cincinnati, OH

ANALYSTS

Kenny Crump
ICF Kaiser, Ruston, LA

Stanley Dawson
Office of Environmental Health Hazard Assessment, California EPA, Sacramento, CA

Dale Hattis
Clark University, Worcester, MA

Suresh Moolgavkar
Fred Hutchinson Cancer Research Center, Seattle, WA

Leslie Stayner
National Institute of Occupational Safety and Health, Cincinnati, OH

Duncan Thomas
University of Southern California, Los Angeles, CA

OBSERVERS AND DISCUSSANTS

Nick Barsic
John Deere, Waterloo, IA

Lester Grant
National Center for Environmental Assessment, U.S. EPA, Research Triangle Park, NC

Aparna Koppikar
National Center for Environmental Assessment, U.S. EPA, Washington, DC

Michael Spallek
Volkswagen AG, Hannover, Germany

Charles Yarborough
Caterpillar, Inc., Peoria, IL

HEI STAFF

Aaron J. Cohen
Senior Scientist

Maria G. Costantini
Senior Scientist

Daniel S. Greenbaum
President

Diane J. Mundt
Consultant

Kathleen M. Nauss
Director for Scientific Review and Evaluation

Robert M. O'Keefe
Director of Program Strategy

Mary Stilwell
Administrative Assistant

Jane Warren
Director of Research

Appendix B. Framework for Systematically Evaluating Epidemiologic Studies for Quantitative Risk Assessment

The Panel developed a framework of epidemiologic principles, with a focus on exposure assessment, for the systematic evaluation of the strengths and limitations of the railroad worker and teamster epidemiologic studies. It designed a set of questions representing an ideal research strategy, and emphasized that a "perfect score" was not the objective. The Panel did not assign overall priorities to these principles, because insisting that any one be given absolute priority over another may interfere with efforts to manage or address other principles in a more balanced fashion.

The Panel recognized that others have written on the use and limitations of epidemiologic studies for QRA (Gordis 1988; Hertz-Picciotto 1995; Samet et al. 1998). However, it chose to develop this series of questions in order to address specifically the issues it believed would be most relevant to systematic assessment of gaps in current knowledge that might be filled by future research efforts. The Panel also recognized that few, if any, epidemiologic studies are designed specifically for QRA, and the framework allows potential users of available studies a somewhat parallel approach to evaluating the strengths and limitations of each. This Appendix offers details and some rationale for including each of the items within the framework.

This framework also provides some basic concepts of epidemiology for readers unfamiliar with such material. For a more complete understanding of epidemiologic study designs and methods, several general texts in epidemiology are available (Monson 1980; Schlesselman 1982; Checkoway et al. 1989; Hulka et al. 1990; Selvin 1991; Gordis 1996; Kelsey 1996; Rothman and Greenland 1998). The Panel assumed that any epidemiologic study would follow standard epidemiologic practices as discussed in these texts.

DESIGN

1. Was the study design efficient, and did it specifically consider power, potential types of bias, and latency?

Epidemiologic studies need time to recruit subjects, an adequate latent period for the health outcome of interest to develop, and sufficient time for the accrual of substantial follow-up during any period of elevated risk. Study design considerations include cohort and case-control designs, the choice of study population, and the selection of controls. An efficient study design that allows more detailed characterization of exposure in all or a subgroup of the cohort, as well as collection of personal data such as smoking history, generally may be useful for a QRA. A simple design is the nested case-control study, but more complex designs also merit consideration. Although these efficient designs allow collection of details that might be otherwise unattainable, epidemiologic studies without them can still be useful in QRA.

Determination of the statistical power needed to address questions of interest is an important aspect of study design; adequate power for analyses of subgroups within the population also needs consideration. The design, including sample size, must make adequate allowance for factors that reduce power (such as losses to follow-up and uncertainties in exposure measurements) and at the same time minimize potential biases. If the sample size is insufficient, statistical power is low and the association under investigation is unlikely to be detected. It is useful to calculate the statistical power for detecting effects of various magnitudes, and to indicate the precision with which effects can be estimated, including the size of the effect that might be excluded (upper confidence limit). Exposure measurement errors, which can reduce both power and estimates of precision, need to be taken into account. Reasonable assumptions about latency are a part of most power calculations. For cohort studies, it is desirable to present power as a function of the duration of follow-up. Cohort studies that continue follow-up as long as subjects provide usable new information can be particularly beneficial to QRA.

The study design should reduce or control for several important biases including confounding, information bias, and selection biases. The contribution of an individual epidemiologic study to the body of literature for hazard identification depends on whether the statistical power of the study is sufficiently large to detect an effect and provides a degree of precision to that estimate as indicated by the width of the confidence interval. It also depends on

whether the study design, data collection, and statistical analyses have reduced the potential sources of bias.

Confounding is a form of bias that occurs when a factor is associated with both the exposure and the outcome of interest. This factor is referred to as a confounding variable or confounder. In studies of lung cancer associated with exposure to diesel exhaust, the potential confounders of importance include age, cigarette smoking, ETS, asbestos exposure, other occupational exposures, and possibly nondiesel outdoor air pollution, but only if the factor is related to both diesel exposure and lung cancer risk. The effects of demographic variables such as age can generally be removed during the analysis, but adjustments for variables such as smoking and exposure to asbestos are not possible in the absence of adequate data on these factors. Information on the degree of possible confounding sometimes can be gleaned from considering causes of death other than the one of primary interest.

The direction of the effect of confounding depends on the directions of the associations between the confounder and the exposure, and between the confounder and the disease. Cigarette smoking, for instance, is widely acknowledged to be positively associated with lung cancer. Thus, if smoking is more common among persons exposed to diesel exhaust than among unexposed persons, the confounding would lead to overestimation of the risk of cancer associated with diesel exhaust. Conversely, if the group exposed to diesel exhaust smokes less than those unexposed, the confounding would lead to underestimation of risk.

Information bias refers to differential errors in measurement of exposure, outcome, or confounder information between groups being compared. For example, if the extent of information available or obtained for exposed persons differs from that for unexposed persons, an information bias would occur.

Misclassification is another form of potential bias in epidemiology that occurs, for example, when a control subject in fact has the outcome of interest. Misclassification can be either differential or nondifferential. In differential misclassification, the rate of misclassification differs by study group; for example, the fact of exposure may be wrong more often for cases than for controls. Nondifferential misclassification is the inaccurate classification of exposure or disease that is not dependent on any factor, such as case or control status. Differential misclassification can bias the estimate of risk in any direction, but nondifferential misclassification most often leads to an overall underestimation of effect. Some forms of misclassification can distort the shape of the exposure-response curve.

Selection bias results from a systematic difference in characteristics of subjects included in a study and is most likely to occur when choosing controls for case-control studies. Selection bias also is possible in cohort studies if, for instance, a cohort is assembled in such a way that completeness of follow-up for lung cancer differs by exposure. (Refer to Appendix D for further discussion of this matter.)

Submittal and selection of a study for publication in a journal, as well as the failure of a study to be published, can result in a form of selection bias. Editors as well as investigators often are reluctant to publish results that do not appear to show an effect, and they are extremely reluctant to publish results that appear to show implausible associations (such as inverse associations between exposure and health outcome). In a research field consisting of dozens of studies, however, a few such results are inevitable.

It is well established that persons in the work force tend to be "healthier" than persons not employed, and therefore healthier than the general population. Worker mortality tends to be below average for all major causes of death. Mortality in the general population thus is generally not a suitable standard for comparison in occupational studies, and can lead to biased findings. Because of this "healthy worker effect," investigators often compare morbidity or mortality of heavily exposed workers with workers in the same industry who have little or no exposure.

Some investigators also have reported a "healthy worker survivor effect" (Arrighi and Hertz-Picciotto 1994), in which mortality of long-term employees in some occupations is less than in those with shorter employment duration. This may seem counterintuitive because, if an association exists, one would expect mortality to increase with longer employment, which then becomes a surrogate for both age and accumulated exposure. Reasons for this survivor effect are not clear, but may include early demise or departure from the work force of some susceptible individuals.

The study design chosen needs to allow for an adequate latent period for developing the health outcome of interest after exposure to the risk factors studied. For some cancers, the latent period may be 20 to 40 years (Cold Spring Harbor Laboratory 1981). Latency period, timing of exposures, duration of exposures, and exposure-response measures are all interlinked, and all are essential to a complete assessment of risk. Exposure duration might provide a reasonable surrogate for exposure, although the investigator might not have data to convert duration into

units of exposure. An additional limitation of this approach is that exposure duration does not reflect any variation of the intensity of exposure with time or among subjects.

Addressing the biases above strengthens a study's possible contribution to both hazard identification and QRA.

2. Were adequate quantitative exposure and covariate data planned and collected?

The Exposure Assessment section below has more details on exposure assessment and data collection. Exposure data generally are collected as part of epidemiologic studies, but the degree of quantification of exposure varies. Crude exposure distinctions can be adequate for some purposes, such as hazard identification, but detailed planning for collection of quantitative exposure data is particularly important if a study is to be used for QRA. The exposure assessment plan also informs the analytic plan.

Planning to include data for each study subject with accurate information about the primary exposure of interest, as well as exposure to other known and suspected risk factors, is optimal. In occupational studies without individual monitoring, surrogate measures or markers to estimate exposure can provide useful information. A range of exposures is needed if effects of different magnitude, duration, or frequency are to be compared.

It is important that the study design consider incorporating the variability of personal exposure over time, including multiple samples for the same individual. To the extent that the study design limits exposure measurement for a subsample of the population, it is important to demonstrate that the individuals selected for monitoring reasonably represent the range of activities, behaviors, and exposures of the study population for which the exposure is assessed retrospectively. Exposure data that include the range relevant to regulatory decision-making would be particularly useful for QRA.

3. Was the literature reviewed for relevant study methods and previous research findings?

The literature review can serve to refine and focus a study's hypotheses, support study design decisions, explain analytic approaches, evaluate exposure measures, define the outcome of interest, and identify questions that remain to be answered. A QRA will raise specific issues about exposure measurement, analytic techniques, and outcome measurements that go beyond what is included in the reports from most epidemiologic studies. The literature review needs to specifically address these QRA issues. A thorough review of the literature regarding a possible hazard includes a critical analysis of all epidemiologic studies concerned with the suspected hazard, and a summary of relevant animal studies, exposure studies, and case studies. Information on possible biologic mechanisms of action or variations in individual susceptibility is also important. A summary of the literature surrounding other known and suspected risk factors for the outcome of interest, as well as a summary of current issues within the research topic, can be useful. If unusual, recent, or critical methodological issues are of importance to the area of research, a summary of the pertinent literature also can be useful.

4. Were plans made for postpublication data scrutiny?

While a study is under way, interim reports, scientific critiques midway through the study, and periodic external independent reviews can provide the investigator with useful feedback. After publication, although not an epidemiologic principle, sharing copies of the data and related documentation with colleagues can assist the scientific community and regulatory agencies in understanding the details of a particular study, and can provide a scientific "second opinion" that further evaluates the pertinent issues in an objective manner. Sharing data with other investigators for reanalysis can allow other analytic approaches to be developed, which can be particularly important when published studies do not produce a clear consensus about the magnitude, or sometimes even the direction, of an effect. Such postpublication analyses are most effective when both the original investigator and the recipient of this information hold to the highest standards of collegiality, ensure the integrity of any promise of confidentiality of study subjects, and respect each other's intellectual property rights. The original investigators should have some rights in these circumstances to see any results of further analyses before they are made public.

EXPOSURE ASSESSMENT

1. Were detailed lifetime histories of occupational exposures collected?

Quantitative measures of exposures are important in any epidemiologic study used for QRA. The greater the detail regarding specific exposure, including how much, for how long, and at what concentration, the more useful the study is for this purpose. Frequently, however, individual measures of exposure are not available, and surrogate measures or markers of exposure are used. For example, the most general surrogate measures of exposure in occupational epidemiologic studies are job classification

and work location. Starting with a standard classification scheme for job and work location generally is preferred, unless there are specific reasons to replace or supplement the standard with an ad hoc classification. Standard schemes include the U.S. Census Bureau Classified Index of Industries and Occupations, the U.S. Labor Department Dictionary of Occupational Titles and Definitions, and the U.S. Office of Management and Budget Standard Industrial Classification and Standard Occupational Classification. These standard codes can be grouped according to possible exposures. The advantage of beginning with the standard classification systems when quantitative exposure data are not available is that many published mortality studies use such systems, and can provide a basis for comparison. A description of job duties frequently can refine industry and job titles.

Information on the magnitude, frequency, and duration of exposures within each job is desirable and preferable to duration only, but is often unavailable. In many occupational studies, the only exposure information available is duration of employment in specific jobs, which provides a poor estimate of exposure. Thus, the exposure scale in exposure-response analyses of such studies is essentially a detailed analysis of duration of employment. In addition, risk can be related to duration of employment for reasons other than the exposure of interest, thus biasing the exposure-response relation; an example is concurrent exposure to other possibly toxic agents.

The goal of a QRA for regulatory purposes is to estimate an effect of lifetime exposure to the agent of interest. Uncertainties are introduced when exposure magnitude, frequency, and duration are not quantified among job categories; when individual exposure varies primarily because of differences in exposure duration within job category; or when exposures in all of the job categories examined are similar. Although measures of cumulative exposure, which depend on exposure duration, are important for QRA, a study is more informative if available exposure data also allow the effects of exposure intensity and duration to be evaluated separately.

2. Were known chemical and physical characteristics of the main exposure specified?

A detailed characterization of the sources of an agent, valuable to understanding the main exposure of interest, would include specific information (chemical and physical characterization) on the nature and source of the exposure, historical trends, aspects of source use that affect the exposure, and the factors that control transport and transformation of the emitted contaminants indoors and outside. Mixtures of substances are particularly difficult to characterize because they may vary from time to time and place to place in ways that may be difficult to document.

3. Were potentially confounding exposures measured or estimated?

The main exposure of interest often is not isolated in the environment where subjects are exposed. That is, subjects are likely to be exposed to additional agents that may confound or modify the association under investigation; for example, smoking or exposure to ETS would be such factors in many studies.

Some studies of occupational exposures have investigated confounding by smoking. Results have been mixed—some produce evidence of confounding and some do not. A nonsignificant p value does not demonstrate that no confounding occurred, though confidence limits on a confounder effect might provide useful information about the greatest degree of confounding that would be generally consistent with the data. For example, smoking might not appear to be a confounder for a particular occupational exposure, but a study is open to criticism if no smoking data are collected and the association between exposure and outcome is weak. In addition to being a confounder, smoking may be an effect modifier, in that the estimated increase in risk of lung cancer from diesel exhaust may differ between smokers and nonsmokers. Smoking data are necessary to allow effect modification to be investigated.

Measurements of confounders are needed to understand their potential effects on the analysis of the exposure of interest. Failure to account for possible confounding or effect modification by other exposures might mask a real (positive or negative) association. When the magnitude of the association of interest is weak, uncontrolled confounding, particularly from a strong confounder such as cigarette smoking, can have a major impact on the study's results and on the credibility of their use.

4. Were magnitude, duration, and variability of exposure determined?

Studies with quantified individual exposure data contribute more to a QRA than studies without such data. Because personal exposure varies over time, sampling the same individual multiple times to capture various seasons, activities, and work place settings is preferable to obtaining a single sample. Multiple measurements during episodes of peak exposure, as well as collecting data for subgroups of interest, are often important.

Historical or current source emission data, when combined with factors that govern the dispersal and removal of contaminants in the environment (meteorological variables, ventilation, dry or wet deposition, chemical transformations, etc.), can be modeled to assess air contaminant concentrations in particular air environments, which in turn can be used to assess individual or group exposures over time. Air contaminant measurements that have been made in several different environments can be used to assess such exposures directly. If an historical record of environmental measurements exists, past exposures associated with a particular environment can be assessed. Estimates of total exposure can be based on either measured or estimated concentrations within each outdoor or indoor environment, combined with time-activity observations or questionnaire data. Uncertainty is associated with using environmental measurements to assess exposure, however, and the source and magnitude of that uncertainty needs to be specified.

5. Were industrial hygiene data and historical data on use and repair of machinery and equipment obtained?

A critical review of the historical records on the nature and chemical characterization of exposure and environmental monitoring data can add to the exposure assessment and improve historical reconstruction estimates. In addition to historical records of measured exposures (specific to an employer), records on the purchase and maintenance of equipment, the industrial processes resulting in exposure, and repair shop operations can provide information to help define historical exposure levels and trends by industry.

6. Were personal exposure measurements obtained, and were they representative of the population studied?

Personal monitoring for inhaled exposures is conducted using either passive (diffusion) or active (pump) samplers worn in the breathing zone to measure exposure to contaminants of interest as people go about their daily activities. Subjects typically wear the monitors for periods of a few hours to several days to provide an integrated measure of exposure over space and time. Personal monitoring can reduce some of the uncertainty in exposure assessment; however, continuous monitoring cannot distinguish among exposures that occur in different indoor and outdoor environments or among different activities in which the subjects are involved.

It usually is not practical or cost-effective to conduct personal monitoring on all subjects. Supplementary questionnaire-derived information, environmental models, and daily activity diaries usually are needed to estimate the contribution of specific environments, activities, or sources to the total exposure.

Special care must be taken in selecting a sample of subjects for personal monitoring because exposures may vary widely within job categories and for an individual over time. A random sample of subjects can be selected in various ways to ensure that it is representative of the entire population under study (e.g., a simple random sample of individuals, stratified sampling based on job environments or job categories, random cluster sampling). Nonrandom samples, especially convenience or quota samples drawn to address specific questions, are sometimes useful but always difficult to interpret, and they produce results of greater uncertainty than properly designed random samples. The process used to select the sample and to document that it is representative of the study population need to be stated clearly and in some detail, especially for nonrandom samples.

7. Were uncertainties in exposure assessment quantified?

Exposure assessment should include quantification of the uncertainty in each aspect of the assessment, as well as the overall uncertainty in estimates of exposure. These uncertainties are likely to vary among subjects (or categories of subjects). Sources of uncertainty that are the same (or strongly correlated) for subjects within a category should be separated from sources that are independent. For many components of the exposure assessment, quantitative estimates of uncertainty are necessarily based on subjective judgments. To ensure that the resulting evaluation of uncertainty is useful for interpreting the exposure-response analyses, it is suggested that exposure assessors work closely with the epidemiologists and statisticians analyzing the epidemiologic data. (Refer to Analysis Question 3 below.)

OUTCOMES

1. Were outcomes defined in specific and objective terms?

Criteria for health outcomes in a study are less subject to bias if the definitions are accurate, specific, attainable, and as objective as possible. A study is less likely to be biased if the outcome is not subject to possible misclassification. For example, the operational definition for cancer can be a tissue specimen or pathologic diagnosis (most

accurate), a clinical diagnosis (less accurate), or a death certificate (least accurate).

2. Was the full range of outcomes included in cohort studies?

Cohort studies rely on various sources of information and record-keeping systems to define and identify subjects having the outcome of interest. A by-product of this is obtaining information about a range of other health conditions. For example, if death certificates are examined for a possible relation between some exposure and death from a particular cancer, the certificates of all deceased subjects must be searched to find those that report the cancer. With little additional effort, the full range of causes of death can be examined. Relevant information on death certificates may include any underlying causes of death, as well as the primary cause of death.

3. Were participants actively followed to determine outcome status and the date outcome occurred?

A clear statement of the cohort follow-up methods is needed. Active follow-up of study participants to determine their health outcome and its date, as well as the determination of a subject as "outcome free," or the date a participant was last known to be "outcome free," can be important for reducing bias. Active follow-up strengthens a study's findings because confirmation of the outcome is sought, and the information acquired is more complete and generally less biased than with passive follow-up. Subjects determined to be "lost to follow-up" are usually assigned a date corresponding to the last date they were confirmed to be "outcome free." If the follow-up is passive, subjects not identified as having the outcome are generally assumed to remain "outcome free." For example, absence of a claim for death benefits may not be adequate evidence that the study subject is still alive. In this situation, a special small study might be undertaken to estimate the number of outcomes that were missed by passive follow-up.

If the percentage of follow-up completed is high, bias tends to be reduced. If a significant number of participants are nonrespondents or lost to follow-up, results can be questionable because of possible misclassification biases, and the number of subjects with unknown outcome may be sufficiently great that the validity of the study findings is questioned.

ANALYSIS

1. Were analytic methods specified a priori?

A study protocol commonly specifies the fundamental hypotheses, but it is equally important to specify the primary methods of analysis the investigators plan to use. Additional analytic techniques, especially those suggested by the data, can supplement the primary methods of analysis, but they lack full statistical justification. A general concern is that analytic methods not specified a priori may be chosen to emphasize some aspect of the findings and, therefore, bias the results. An example would be performing multiple statistical tests of variations and reporting only the most extreme p values.

The statement of primary analytic methods specifies the general analytic approaches, software packages to be used, cut off points for continuous variables, exposure metrics, methods of adjustment for age and other confounders, subsets of the data to be analyzed, and exposure-response models.

Exposure data may include several types of measurements, and several possible approaches are likely for assigning exposure estimates to individual subjects. Although analyses based on more than one approach can be informative, it is important to specify the primary exposure metrics before the exposure-response relation is examined.

Defining group categories from continuous data is also important, and these categories should be specified in the protocol rather than determined on the basis of statistical significance levels after the data have been obtained.

2. Was the appropriateness of the statistical approach demonstrated, and were potential biases explored?

An early step in the analysis is to evaluate the appropriateness of the statistical methods specified in the protocol for the primary analyses. This usually involves comparing the fits of various models to establish that analyses based on the model selected do not seriously distort results. The specifics of this evaluation vary from study to study; discussion of a few common types of assumptions follows.

Because lung cancer risk depends on gender, age, calendar year period, and smoking habits, and because estimates of exposure might also vary by these factors, it is important to provide adequate control for these factors in developing quantitative risk estimates. Age and calendar year period can be related strongly to cancer risks as well as to many other health endpoints. In studies (such as occupational studies) in which exposures accumulate over time, exposure also is likely to be related to age and

calendar year period. For this reason, analyses need to provide tight control for these variables, and use minimal assumptions regarding the form of their relation to risk of disease. Less flexible models (for example, those that treat age or calendar year as continuous, linear, or exponential variables, or do not include interaction terms) can be considered, but only if it has been demonstrated that they provide an adequate fit to the data.

Analyses are often based on the assumption that the relative risk is constant across strata. This assumption can be evaluated by fitting models that allow relative risks to vary by factors such as age, calendar year period, gender, and time since initial employment. Differences in relative risks might simply reflect differences in the magnitude of exposure; alternatively, other factors may modify sensitivity to a specific exposure. This modification of risk can be important in analyses that quantify the relation between exposure and risk (for example, by estimating the risk per unit of exposure). Risks can be compared among subgroups defined by these variables to investigate these dependencies, or parameters can be introduced to estimate and test the dependencies, or both. However, statistical power and precision for adequately evaluating modifying factors are often limited. Also, whether a factor is considered to modify the risk depends on the metric of risk. For example, if relative risks for lung cancer are similar for smokers and nonsmokers, absolute risks for the two groups can differ substantially.

Absolute risk models often are more difficult to fit; however, they give a direct measure of impact and, therefore, their wider use is suggested. Neither a relative risk model (and its surrogate, the odds ratio) nor an absolute risk model might mirror the biological situation; there is little reason to expect nature to follow any predetermined mathematical structure. Many statistical modeling approaches are available, and the discussion here is not meant to be inclusive. Other types of models, such as nonparametric models (including Kaplan-Meier plots), multistage models of carcinogenesis, or other biologically based models might also be appropriate to consider. However, in the absence of an understanding of the mechanisms for exposure-related carcinogenic risk, it may be difficult to interpret the results of multistage models or other biologically based models. Even when the fit is good, such models generally should be reported as approximations. If evidence of substantial variation is found, it is especially important to fit models that allow for this variation and to present separate results for various subgroups of data.

3. Were exposure-response relations statistically explored?

The exposure-response relation is important to the development of a QRA. In the analysis, the investigators may explore the shape of the relation (e.g., linear), possible modifying factors, and alternative measures of exposure.

Flexible models permit a better understanding of the possible relation between exposure and health outcome. This understanding is valuable for estimating the impact of a particular exposure on a population's health both present and future, as well as for the risk estimation needed to assess and control occupational and environmental hazards.

The effect of a given exposure on risk is likely to depend on the length of time that has passed since the exposure occurred. First, for example, recent exposures are not likely to have affected the risk of diseases with a long latent period (e.g., cancer), so it is often desirable to exclude exposure received during some specified period before the time at risk, known as the lag period. The lag period need not be the same for all endpoints. Even if the protocol has specified lag periods, it is desirable to evaluate their appropriateness by conducting analyses based on alternative lag periods. Second, to evaluate more completely the effect of time since exposure, it might be desirable to include separate variables for the effects in each of several "windows" of exposure (for example, categories of time since exposure such as 5 to 15, 15 to 25, and 25+ years), or to use continuous variables to examine a change in risk with time since exposure. Measures that weight exposures in the various "windows" might also be considered. Third, exposures that vary from time to time, or from one person to another in the same category, may be weighted to reflect actual exposures, but this requires data that are rarely available.

If a specific exposure-response model is proposed (such as a linear model), the assumption of the proposed response form also needs validation. This can be accomplished by comparing the fit with more general parametric models or categorical models. For example, the fit of a linear model can be compared with the fit of a linear-quadratic model or a model in which the risk is related to an estimated power of the exposure.

Often more than one approach is available for defining the exposure metric to be analyzed. Most studies can benefit from collaboration among the data analysts and the persons assessing exposure. Factors to consider include possible bias and uncertainty in the proposed metric and the ability of the metric to discriminate among subjects. The final choice can involve a trade-off of these factors. In

an occupational study, for example, duration of employment might be estimated accurately, but exposures among subjects with similar employment duration cannot be differentiated. Measures that use industrial hygiene data to assign quantitative measures to various job categories provide better discrimination among subjects, but are likely to be subject to large uncertainties. Several exposure measures for each subject can be computed and examined for variability and the degree to which they are correlated.

4. Were uncertainties in risk estimates quantified, especially those resulting from exposure measurement error?

Exposure estimates are subject to both systematic bias and random errors. Systematic bias in exposure estimates can affect risk estimates that are expressed per unit of exposure. Random errors in exposure estimates also can result in bias. In general, random errors bias estimated regression coefficients toward the null, although this may not necessarily be true in the special case in which the only error is the substitution of group means for individual measurements (Berkson error). Random errors can also result in both underestimation of uncertainty and distortion of the shape of the exposure-response curve. Statistical analyses can take into account the uncertainties in exposure estimates; however, this is often difficult or impossible in practice, in part because statistical methods for doing so are complex (possibly requiring extensive software development), and in part because it is difficult to quantify potential bias and uncertainty.

Fairly simple procedures sometimes can provide at least some information on the sensitivity of analyses to uncertainties in exposure estimation. One such procedure is to conduct restricted analyses that exclude subjects or job categories thought to have exposure estimates that are especially uncertain or particularly subject to bias. Another is to restrict analyses to workers initially employed after some specified date when exposure assessment is thought to have become more adequate.

Methods for accounting for random errors have been developed, but these procedures are often difficult to apply and can require computer simulations. They also require knowledge of the nature and magnitude of exposure measurement errors, including the extent to which errors are correlated among subjects. Application of such methods might be important even when they are based on overly simple assumptions; sensitivity analyses can then be conducted under several alternative assumptions. Such analyses do not decrease the uncertainty resulting from imperfect dosimetry (only improvements in dosimetry can do that), but they can provide an assessment of the additional uncertainty in the estimated exposure-response relation resulting from this problem.

Sampling error is another source of uncertainty in risk estimates. Confidence intervals and *p* values that describe this uncertainty often are based on the assumption that various statistics follow normal (Gaussian) distributions. Statistical theory provides assurance that this assumption is appropriate provided the sample size is sufficiently large. However, exposure distributions are often highly skewed (nonnormal), and analyses can be affected strongly by a small excess of observations in higher exposure categories. In this situation, the assumption of normality might not be adequate or appropriate. Although they are computationally more cumbersome, *p* values and confidence intervals based on a likelihood ratio statistic often provide better approximations of uncertainty than other approaches (such as the use of the asymptotic standard error). Software that allows reasonably easy implementation of this approach is now available. In some cases, exact methods or computer simulations can be considered.

Appendix C. The Panel's Exploration of Railroad Worker Data

This appendix describes results of the Panel's examination of certain aspects of the original railroad worker data, as presented by Garshick and colleagues (1988). These analyses were undertaken to assist the Panel in evaluating the usefulness of the data for QRA, to help clarify reasons for differences in results obtained by other analysts, and to verify and better understand these other analyses. The Panel's analyses were limited to these objectives, and were not intended to explore or evaluate completely the railroad worker data or to provide a model for QRA. This discussion of results and stated conclusions describes the Panel's findings from its analysis in conjunction with its review of the relevant literature.

DESCRIPTION OF THE DATA

Dr. Eric Garshick provided the basic data in the form of computer records on over 55,000 railroad workers. The Panel used the DATAB module of the software package EPICURE (Preston et al. 1991) to collapse data into cells that contained person-years of exposure, and numbers of deaths from all causes and from lung cancer. Deaths from lung cancer were defined as those where the underlying cause was coded ICD 162.0 or 162.1. Categories of cells included single calendar years (1959 through 1976), attained age in five-year intervals to age 80, and age in 1959 in five-year intervals. A few lung cancer deaths over age 80 and the associated person-years of follow-up were excluded because of a concern about the quality of death certificate information in this age group. The last four years of follow-up (1977–1980) were excluded because follow-up appeared to be incomplete during this period (Crump et al. 1991; Garshick 1991).

Duration of exposure to diesel exhaust (which was duration of employment) was categorized into 12 intervals (in months: 0–29, 30–44, 45–59, 60–74, 75–89, 90–109, 110–129, 130–149, 150–174, 175–199, 200–224, 225–249, and ≥ 250) with a ramp of linearly increased exposure over the 15 years prior to 1959. Lags of 5 and 10 years were considered. The analysis was limited to men whose job categories in 1959 were classified as (1) clerks, (2) signalmen, (3) engineers and firers, (4) conductors and brakemen, (5) hostlers, and (6) shop workers. These groups were subsequently combined for analysis into three groups consisting of clerks and signalmen (groups 1 and 2), train workers (groups 3 through 5), and shop workers (group 6).

METHODS USED FOR EXPOSURE-RESPONSE ANALYSES

Analyses were based on Poisson regression using the AMFIT module of the software package EPICURE (Preston et al. 1991). Previous analyses (Crump et al. 1991; OEHHA 1998; Crump 1999) also were based on Poisson regression. The variables of employment duration and job category were grouped as described above. Unless noted otherwise, analyses were stratified by attained age (five-year categories) and calendar year (single-year categories). The STRATA command of AMFIT was used to include all interactions among these variables. Duration of employment and measures of cumulative exposure were lagged for five years; alternative analyses using a ten-year lag yielded very similar results. Analyses used deaths in which the underlying cause was lung cancer; analyses that included deaths in which lung cancer was coded as a secondary cause produced similar results.

The results presented are based on the log-linear model (used by OEHHA) in which the relative risk is calculated by

$$RR = e^{\alpha GRP + \beta D},$$

where *GRP* is an indicator variable for groups defined by job category, and *D* is a measure of duration of employment or exposure. Both the estimated coefficients (α and β) and the associated relative risks are presented. Other analyses, based on the linear model (used in many of the analyses conducted by Crump and associates [1991])

$$RR = e^{\alpha GRP}[1 + \beta D],$$

led to conclusions similar to those from the log-linear model. For graphic presentation, relative risks by job category and exposure duration were calculated. Confidence intervals were based on the asymptotic standard errors of α or β. Statistical tests were based on the likelihood ratio statistic. All reported *p* values are two-tailed.

RESULTS OF EXPOSURE-RESPONSE ANALYSES

Other analysts have assigned exposure rates to job categories and multiplied these by various measures of exposure duration. For this reason, the Panel's initial analyses addressed the effects of both duration of employment and job category. For these analyses, three broad job category groups were defined as indicated in Table C.1.

Results of various models are shown in Table C.2. Model 1, which included variables for both job category (*GRP*) and employment duration (*D*0) fit substantially better than a model that included only employment duration ($p < 0.001$; model not shown). The coefficient for *D*0 (expressed per 10 years of employment duration) was negative when the *GRP* variable was not included. In Model 2, which included a separate employment duration coefficient for each of the three job categories, all three coefficients were negative, and this model did not fit significantly better ($p = 0.49$) than Model 1, which used a single employment duration coefficient. For graphic presentation, relative risks for each category defined by duration of employment (0–4, 5–9, 10–14, 15–17, and 18+ years) and job category were calculated with the 0–4-year category for clerks serving as the referent category. These relative risks are shown in Figure C.1.

The next set of analyses explored variables similar to those used by OEHHA in which one constant exposure rate was assigned to clerks and signalmen and another rate to train workers. The exposure rates given in Table 7-2 of the California OEHHA report (1998) were 39 $\mu g/m^3$ for clerks and 82 $\mu g/m^3$ for train workers. When background exposure (the value for clerks) was subtracted, the result was 0 $\mu g/m^3$ for clerks and 43 $\mu g/m^3$ for train workers. Because one of the issues of interest was the effect of subtracting background exposure, two exposure variables were defined: *D*1 with no background correction, which equaled 39 $\mu g/m^3$ for clerks and 82 $\mu g/m^3$ for train workers; and *D*2, with the background exposure subtracted, which equaled 0 $\mu g/m^3$ for clerks and 43 $\mu g/m^3$ for train workers. Shop workers were excluded from these analyses to be consistent with those of other analysts. (Again, the Panel's analyses were aimed at understanding the data and the factors that influenced risk estimates obtained by others, and not at developing a model for QRA.)

The analyses by OEHHA used exposure measures similar to the *D*2 variable described above; they used exposure intensities of 40, 50, or 80 $\mu g/m^3$ for train workers, which were intended to reflect different assumptions regarding baseline exposure concentrations (as described in section 7.2.1.4 of OEHHA 1998). Because the Panel used durations and intensities of exposure that were not identical to those used by the OEHHA, the quantitative risk estimate analyses are not directly comparable. However, the Panel believes that the direction and statistical significance of the effect for the *D*2 variable should be reasonably comparable with the OEHHA results.

Table C.1. Number of Lung Cancer Deaths and Person-Years of Follow-Up by Job Category from Railroad Worker Cohort Data

Job Category	Deaths from Lung Cancer	Person-Years of Follow-Up
Clerks and signalmen	307	220,802
Train workers (including engineers, firemen, conductors, brakemen, and hostlers)	752	462,951
Shop workers	321	190,874

Model 3 included only the *GRP* variable and Models 4 and 5 included only the quantitative cumulative exposure measure *D*1 or *D*2, respectively. Without the *GRP* variable in Models 4 and 5, the coefficients for both *D*1 and *D*2 were positive, but the *p* value for statistical significance was smaller when *D*2 (with background correction) was used (Model 5: $p = 0.005$) than when *D*1 (with no background correction) was used (Model 4: $p = 0.10$). When *GRP* was added (Models 6 and 7), the fit was improved significantly and the coefficients for both *D*1 and *D*2 became negative. In addition, the relative risks for train workers compared with clerks and signalmen became larger than in Model 3, where the variable adjusting for

Figure C.1. Panel's analysis depicting consistently elevated risk of lung cancer for train workers compared with clerks for each time period, but decreasing risk by job category over duration of employment.

Table C.2. Estimated Coefficients and Relative Risks with 95% Confidence Intervals for Job Category Groups, Duration of Employment (*D*0), and Cumulative Exposure (*D*1 and *D*2)

Variable	Coefficient α or β[a,b] (95% CI)	Relative Risk (e^{α} or e^{β})[c,d] (95% CI)	Two-Sided *p* Value
Model 1			
GRP			
Clerks and signalmen	0.0	1.0	
Train workers	0.26 (0.13, 0.40)	1.3 (1.1, 1.5)	< 0.001
Shop workers	0.12 (–0.04, 0.27)	1.1 (1.0, 1.3)	0.14
*D*0 (Duration of employment)	–0.29 (–0.50, –0.09)[a]	0.7 (0.6, 0.9)[c]	0.006
Model 2			
GRP			
Clerks and signalmen	0.0	1.0	
Train workers	0.25 (–0.11, 0.60)	1.3 (0.9, 1.8)	0.18
Shop workers	–0.08 (–0.49, 0.33)	0.9 (0.6, 1.4)	> 0.5
*D*0 (Duration of exposure)			
Clerks and signalmen	–0.35 (–0.64, –0.06)[a]	0.7 (0.5, 0.9)[c]	0.02
Train workers	–0.33 (–0.57, –0.10)[a]	0.7 (0.6, 0.9)[c]	0.007
Shop workers	–0.18 (–0.46, 0.10)[a]	0.8 (0.6, 1.1)[c]	0.22
Model 3			
GRP			
Clerks and signalmen	0.0	1.0	
Train workers	0.26 (0.12, 0.39)	1.3 (1.1, 1.5)	< 0.001
Model 4			
*D*1 (No background correction)	0.17 (–0.04, 0.37)	1.2 (1.0, 1.5)	0.10
Model 5			
*D*2 (Background correction)	0.33 (0.10, 0.57)	1.4 (1.1, 1.8)	0.005
Model 6			
GRP			
Clerks and signalmen	0.0	1.0	
Train workers	0.48 (0.27, 0.69)	1.6 (1.3, 2.0)	< 0.001
*D*1 (No background correction)	–0.43 (–0.75, –0.10)[b]	0.7 (0.5, 0.9)[d]	0.011
Model 7			
GRP			
Clerks and signalmen	0.0	1.0	
Train workers	0.50 (0.19, 0.81)	1.6 (1.2, 2.2)	0.002
*D*2 (Background correction)	–0.47 (–1.01, 0.07)[b]	0.6 (0.4, 1.1)[d]	0.09

[a] Coefficient β per 10 years.

[b] Coefficient β per 10^3 (μg/m^3)-years.

[c] Relative risk e^{β} per 10 years.

[d] Relative risk e^{β} per 10^3 (μg/m^3)-years.

exposure duration was not included. Analyses based on linear models resulted in a similar pattern, where the sign of the coefficient changed from positive to negative when the *GRP* variable was added to the model.

In Table C.2, the differences in relative risk between train workers and clerks/signalmen, as measured by the parameter α, were assumed not to depend on age or calendar year period. That is, the relative risk of lung cancer for train workers compared with clerks and signalmen was assumed to be the same for all age and calendar year periods.

Further exploration of these data revealed that these relative risks were not constant; risks for train workers increased with calendar year period (after adjustment for age and duration of employment), whereas risks for clerks and signalmen showed little evidence of such an increase. Although the Panel has not explored these differences in detail, they nevertheless illustrate the complexity of these data and the difficulties of providing an adequate summary measure of effect. Calendar year and cumulative exposure are highly correlated, which makes it especially difficult to sort out their separate effects; correlation coefficients (calendar year and cumulative exposure), weighted by person-years, were 0.90 for clerks and signalmen, 0.94 for train workers, and 0.87 for shop workers.

Because of the dependencies noted above, analyses using separate calendar year and age adjustments for each of the three job category groups were also conducted by stratifying on age, calendar year, and *GRP*. Using this approach, the coefficients (with 95% CI) for duration of employment (analogous to those presented in Table C.2, Model 2) were –0.01 (–0.47, 0.44) per 10 years for clerks, –0.55 (–0.84, –0.27) per 10 years for train workers, and –0.09 (–0.46, 0.29) per 10 years for shop workers. The coefficients for *D*1 and *D*2 (analogous to analyses based on Models 6 and 7) were –0.62 (–0.95, –0.28) per 10^3 ($\mu g/m^3$)-years for *D*1, and –1.28 (–1.94, –0.62) per 10^3 ($\mu g/m^3$)-years for *D*2.

Several approaches for adjusting for age and calendar year were explored, including (1) substituting age in 1959 for age at risk, (2) not including interactions of age and calendar year, and (3) treating age as a continuous variable (log-linear) instead of as a categorical variable. Results of fitting these various methods of adjustment to Models 5 and 7 (with *D*2) are shown in Table C.3. With all methods of adjustment, the coefficients from Model 5 (with *D*2 alone) were positive and differed significantly from zero. Also in all cases, these coefficients became negative when a variable measuring differences between train workers and clerks/signalmen (*GRP*) was added (Model 7). However, in some cases (not shown), interaction terms improved the fits of various models, especially if age in 1959 was used instead of age at risk. Also, treating age at risk (or age in 1959) categorically improved the fit over treating these variables continuously.

DISCUSSION

The analyses described above demonstrate that the lung cancer risks for train workers were higher than for clerks and signalmen, and that shop workers had intermediate risks. Within the exposed groups (train workers and shop workers), lung cancer risk decreased with increasing dura-

Table C.3. Estimated Coefficients and 95% Confidence Intervals for Simple Measures of Exposure (*D*2) Based on Different Adjustments to Models 5 and 7 for Age and Calendar Year

Method of Adjusting for Age and Calendar Year[a]	Coefficient β per 10^3 ($\mu g/m^3$)-Years	
	Model 5 Without *GRP* in Model	Model 7 With *GRP* in Model
Age at risk[b], Year, Interaction	0.33 (0.10, 0.57)	–0.47 (–1.01, 0.07)
Age at risk[b], Year	0.32 (0.08, 0.55)	–0.50 (–1.03, 0.04)
Age at risk[c], Year	0.41 (0.17, 0.64)	–0.09 (–0.62, 0.45)
Age in 1959[d], Year, Interaction	0.33 (0.09, 0.56)	–0.49 (–1.04, 0.05)
Age in 1959[d], Year	0.37 (0.14, 0.61)	–0.22 (–0.76, 0.32)
Age in 1959[c], Year	0.41 (0.17, 0.64)	–0.09 (–0.62, 0.45)

[a] "Year" is treated as 18 single-year categories from 1959 through 1976. "Interaction" includes all possible interaction terms for the two variables indicated.

[b] "Age at risk" is treated as 8 five-year categories 40–44, 45–49, . . . , 75–79.

[c] Age is treated as a continuous variable.

[d] "Age in 1959" is treated as 5 five-year categories 40–44, 45–49, . . . , 60–64.

tion of employment, although this decrease was statistically significant only for the train workers. If exposure to diesel exhaust increased the risk of lung cancer in a monotonic fashion, such decreases would not be expected.

Simple measures of exposure, determined by multiplying duration of employment by exposure intensities assumed to be constant for the clerks/signalmen and train worker groups, were also analyzed. With no adjustment for job category, a positive slope was obtained, which was statistically significant when zero intensity was assumed for the clerks and signalmen group. To clarify the extent to which this positive response was due to the difference in baseline risks for train workers compared with clerks and signalmen, the variable *GRP* to measure this difference was included in the model. When this was done, the direction of association with exposure became negative; this result reflects entirely the decreasing association with increasing employment duration for train workers, and demonstrates that the positive exposure-response relation obtained with Models 4 and 5 was entirely due to the baseline differences between the two job categories.

The exposure measure defined for Model 5 above was categorized as > 0–465, 466–626, and 627 or more (μg/m^3)-years. Compared with the 0 category, the relative risks (and 95% CIs) are 1.4 (1.2, 1.7), 1.3 (1.1, 1.6), and 1.1 (0.9, 1.3), respectively. The clerks and signalmen comprised the 0 category, and the positive exposure categories were equivalent to duration of employment categories for train workers. For the train workers, the cut points 466 and 627 (μg/m^3)-years correspond to 10.8 and 14.6 years (or 130 and 175 months in the original categorization of the data).

The exposure-response function indicated by the relative risks noted in the previous paragraph and the decrease in risk with increasing duration of employment shown in Figure C.1 is not consistent with a positively increasing association between cumulative exposure to diesel exhaust and lung cancer risk. Even if the estimated exposure intensities were incorrect, a positive trend with duration of employment would be expected for the exposed groups (including train workers) if lung cancer risk increased consistently with increasing cumulative exposure. Furthermore, if the difference between train workers and clerks/signalmen were primarily due to differences in exposure, one would expect the relative risk for train workers, compared with clerks and signalmen, to be smaller or even eliminated after adjusting for exposure. However, adjusting for exposure increased this relative risk, as shown with Models 6 and 7 (compared with Model 3) in Table C.2.

The above analyses consider the ramp pattern of exposure (see Figure 7 in the chapter Summary of Railroad Worker and Teamster Studies); however, analyses using the roof pattern of exposure produced similar results. For Models 4 and 5 using the roof pattern, the slopes for cumulative exposure *D*1 and *D*2 increased to 0.64 and 0.97, respectively. When *GRP* was included, as in Models 6 and 7, the slopes for *D*1 and *D*2 became negative (–0.82 and –0.86, respectively). Thus, the coefficients for cumulative exposure using the roof pattern responded in the very same manner as they did using the ramp pattern, changing from positive to negative when the variable for job category (*GRP*) was included.

The OEHHA also applied a multistage model to the data to make quantitative risk estimates. This involved assigning exposure concentrations to job categories for modeling. As a correction for background exposure, the clerks and signalmen were given values of zero exposure rate as was done in Model 5 of Table C.2. The OEHHA reported a positive dose-response relation, as is also shown with Model 5. This multistage approach is likely to have the same difficulties with cumulative exposure as were pointed out in the discussion of Models 5 and 7.

The differences between train workers and clerks/signalmen and the negative association with employment duration within job category (especially the train workers) indicate a strong likelihood of bias due to unmeasured variables, such as smoking, or to differential follow-up among the groups being compared, or to both. The reasons for these biases are not known, but it is the Panel's opinion that their probable presence makes these data unsuitable for QRA.

In addition, the strong correlation between calendar year and duration of employment makes it very difficult to sort out the separate effects of these variables. In a study with excellent data on exposure and on other risk factors for lung cancer, this might be possible. However, this study did not include data that clearly distinguish exposure rates among workers, and did not indicate changes in these rates over time for individual workers. Data on potential confounders such as smoking were also unavailable. Although the railroad worker data do not show an increasing relation between exposure to diesel exhaust and risk of lung cancer, the possibility that strong biases in these data have masked a true association cannot be ruled out. It is also possible that the follow-up period has been too short for effects to become evident, or that the true exposure-response relation in these data is nonmonotonic. With these limited data, it is not possible to determine whether the exposure-response relation is, in fact, nonmonotonic, or the result of bias in the data.

Model 5 is similar to the model used as the basis for the California OEHHA's risk assessment (1998). This model

does not fit the data, as was demonstrated by the strong improvement in fit when the variable reflecting the difference between train workers and clerks/signalmen (*GRP*) was added (Model 7). Analyses using Model 5 differ from OEHHA's analyses in that (1) the employment duration measure OEHHA used did not take into account the portion of the year (months) workers were employed, and (2) the methods used to adjust for age and calendar year period were slightly different. Although these differences could affect the quantitative estimates, it seems unlikely that they would affect qualitative conclusions.

Crump (Crump et al. 1991; Crump 1999) also has conducted many analyses of the railroad worker data, and has investigated duration of employment as well as more complex measures of exposure than the Panel used in its analyses. However, Crump's results are likely to reflect the negative association of duration of employment within job categories. Crump has noted that the OEHHA positive slope is driven by the difference between train workers and clerks/signalmen (Crump 1999), and the Panel's analyses confirm this. The Panel used results of the analyses with and without terms for job category, in an attempt to clarify the roles of job category (as a measure of exposure intensity) and employment duration, and to understand to a greater extent than in previous discussions of these analyses the differences in results obtained by OEHHA and by Crump and colleagues.

Methods to adjust for age and calendar year, particularly regarding the choice between attained age or birth cohort (equivalent to age in 1959), have received considerable discussion. Because any two of the variables (attained age, birth cohort, and calendar year) determine the third, this choice in itself should not matter provided that an interaction among the variables is allowed. However, different results might arise because of the number and definition of the categories used, or because interactions were not included in the model. The California OEHHA (1998) has conducted analyses, applying many different treatments of age and calendar year, with little evidence of an important impact on results. The Panel's investigation of this issue confirms this conclusion, in that analyses were not greatly affected by the method of adjusting for age and calendar year.

ACKNOWLEDGMENTS

The Panel greatly appreciates the helpful cooperation of Drs. Eric Garshick, Kenny Crump, and Stanley Dawson in providing data and other information regarding their analyses of the railroad worker data.

Appendix D. Potential Impact of Control Selection Bias

The Panel, having raised the possibility that bias could have been introduced during the selection of control groups for the case-control studies by Garshick and colleagues (1987) and Steenland and coworkers (1990), offers hypothetical examples showing the potential for bias in case-control studies with deaths from "other causes." "Other causes" in these examples are depicted as CVD mortality because it is the leading cause of death among men; the problems arise because CVD often is associated strongly with smoking as well as with choice of occupation. The Panel presents five different scenarios concerning the association between (1) diesel exhaust and lung cancer, (2) diesel exhaust and "other" causes of death, (3) smoking and lung cancer, (4) smoking and the "other" causes of death, and (5) diesel exhaust, smoking, and lung cancer. These examples demonstrate that case-control studies using "other" causes of death as the control group have the potential to produce misleading results.

In all the examples it is assumed that:

- smoking has a strong, positive effect on lung cancer mortality risk: RR = 7.00; and
- smoking has a weaker, positive effect on CVD mortality risk: RR = 2.00.

The examples differ by:

- whether exposure to diesel exhaust has a positive effect (RR > 1.00) or no effect (RR = 1.00) on lung cancer risk;
- whether smoking is correlated with diesel exhaust exposure, with 65%, 50%, or 35% of smokers in the exposed group; and
- whether CVD risk is increased (RR = 1.54) or decreased (RR = 0.65) among exposed persons, within strata of cigarette smoking.

In each example, a cohort study that lacked information on smoking would provide a biased estimate of the crude relative risk (in the row in each section of the tables labeled "Total"). A case-control study with CVD deaths as controls and with information on smoking, as conducted by Garshick and colleagues (1987) and Steenland and coworkers (1990), would provide estimates of the crude and stratum-specific relative risks in the last column of each table.

Table D.1 includes data for the first three examples. In all three, exposure to diesel exhaust *has no effect on* lung cancer, and exposed persons are *at lower risk* of CVD than unexposed persons within strata of smoking. Smoking is associated with diesel exhaust in a different direction for each example.

In the first example (section A of Table D.1), smoking is positively associated with diesel exposure. A cohort study of this population would report a crude RR = 1.58, which would be due entirely to confounding by smoking. A case-control study with CVD deaths for controls would report a crude estimate of RR = 1.99, which would appear to be due only partially to confounding by smoking. The smoking-adjusted estimate from the case-control study would be RR = 1.54, very similar to the unadjusted estimate of 1.58 from the cohort study. The adjustment for smoking in the case-control study would be falsely reassuring because of control selection bias.

In the second example (section B of Table D.1), smoking is inversely associated with exposure. If the case-control study only were performed, as in the case of the study by Steenland and colleagues, the implausible RR = 0.63 in the cohort, unadjusted for smoking, would never be seen. The case-control study would find a weak association between exposure and lung cancer (RR = 1.19) before adjusting for smoking, which would become stronger (RR = 1.54) when smoking is controlled. However, the scenario assumes that diesel exposure conditional on smoking status has no effect on lung cancer (RR = 1.00).

In the third example (section C of Table D.1), smoking is unassociated with exposure. As in the second example, suppose that only the case-control study is conducted. Control selection bias resulting from the use of CVD controls would produce an upwardly biased estimate of RR = 1.54, whether or not smoking is controlled for in the analysis.

Table D.2 includes data for the fourth and fifth examples. For both hypotheses, exposure to diesel exhaust *increases* lung cancer risk, and exposed persons are *at higher risk* of CVD than unexposed persons within the strata of smoking. The association of smoking and diesel

Table D.1. Hypothetical Examples in Which Exposure to Diesel Exhaust Has No Effect on Lung Cancer Risk, Exposed Persons Are at Lower Risk of Cardiovascular Disease Than Unexposed Persons Within Strata of Smoking, and the Association Between Smoking and Exposure Changes

			Risk		Cases		Relative Risk of Lung Cancer[a]	
Smoking	Diesel Exposure	Number of Persons	Lung Cancer	Cardio-vascular Disease	Lung Cancer	Cardio-vascular Disease	Cohort	Case Control
A. Smoking Is Positively Associated With Exposure								
Yes	Yes	65,000	0.00280	0.01300	182	845	1.00	**1.54**
	No	35,000	0.00280	0.02000	98	700	1.	1.
No	Yes	35,000	0.00040	0.00650	14	228	1.00	**1.54**
	No	65,000	0.00040	0.01000	26	650	1.	1.
Total	Yes	100,000	0.00196		196	1,073	**1.58**	**1.99**
	No	100,000	0.00124		124	1,350	1.	1.
B. Smoking Is Inversely Associated with Exposure								
Yes	Yes	35,000	0.00280	0.01300	98	455	1.00	**1.54**
	No	65,000	0.00280	0.02000	182	1,300	1.	1.
No	Yes	65,000	0.00040	0.00650	26	423	1.00	**1.54**
	No	35,000	0.00040	0.01000	14	350	1.	1.
Total	Yes	100,000	0.00124		124	878	**0.63**	**1.19**
	No	100,000	0.00196		196	1,650	1.	1.
C. Smoking is Unassociated with Exposure								
Yes	Yes	50,000	0.00280	0.01300	140	650	1.00	**1.54**
	No	50,000	0.00280	0.02000	140	1,000	1.	1.
No	Yes	50,000	0.00040	0.00650	20	325	1.00	**1.54**
	No	50,000	0.00040	0.01000	20	500	1.	1.
Total	Yes	100,000	0.0016		160	975	**1.00**	**1.54**
	No	100,000	0.0016		160	1,500	1.	1.

[a] Results from a cohort study without smoking information and from a case-control study with CVD controls and smoking information are shown in bold.

exhaust is assumed to go in opposite directions for each example.

In the fourth example (section A of Table D.2), smoking is positively associated with exposure. The cohort study with no smoking information gives too high an estimate (RR = 3.16) of the actual effect of exposure to diesel exhaust on lung cancer risk (RR = 2.00). Because of control selection bias, the estimates from the case-control study are too low, regardless of whether they are computed with (RR = 1.30) or without (RR = 1.68) adjustment for smoking.

In the fifth example (section B of Table D.2), smoking is inversely associated with exposure. As in the fourth example, the unadjusted result of the cohort study and the unadjusted and adjusted results of the case-control study underestimate the effect of exposure on lung cancer risk.

These examples show that a severely distorted estimate of the association between exposure to diesel exhaust and lung cancer, and a severely distorted picture of the direction and degree of confounding by cigarette smoking, can come from case-control studies in which the controls are a collection of "other deaths" composed largely of CVD deaths.

Table D.2. Hypothetical Examples in Which Exposure to Diesel Exhaust Increases Lung Cancer Risk, Exposed Persons Are at Higher Risk of Cardiovascular Disease Than Unexposed Persons Within Strata of Smoking, and the Association Between Smoking and Exposure Changes

			Risk		Cases		Relative Risk of Lung Cancer[a]	
Smoking	Diesel Exposure	Number of Persons	Lung Cancer	Cardio-vascular Disease	Lung Cancer	Cardio-vascular Disease	Cohort	Case Control
A. Smoking Is Positively Associated with Exposure								
Yes	Yes	65,000	0.00280	0.02000	182	1,300	2.00	**1.30**
	No	35,000	0.00140	0.01300	49	455	1.	1.
No	Yes	35,000	0.00040	0.01000	14	350	2.00	**1.30**
	No	65,000	0.00020	0.00650	13	423	1.	1.
Total	Yes	100,000	0.00196		196	1,650	**3.16**	**1.68**
	No	100,000	0.00062		62	878	1.	1.
B. Smoking is Inversely Associated with Exposure								
Yes	Yes	35,000	0.00280	0.02000	98	700	2.00	**1.30**
	No	65,000	0.00140	0.01300	91	845	1.	1.
No	Yes	65,000	0.00040	0.01000	26	650	2.00	**1.30**
	No	35,000	0.00020	0.00650	7	228	1.	1.
Total	Yes	100,000	0.00124		124	1,350	**1.27**	**1.01**
	No	100,000	0.00098		98	1,073	1.	1.

[a] Results from a cohort study without smoking information and from a case-control study with CVD controls and smoking information are shown in bold.

Literature Cited

Arrighi HM, Hertz-Picciotto I. 1994. The evolving concept of the healthy worker survivor effect. Epidemiology 5:189–196.

Bhatia R, Lopipero P, Smith AH. 1998. Diesel exhaust exposure and lung cancer. Epidemiology 9: 84–91.

Boffetta P, Jourenkova N, Gustavsson P. 1997. Cancer risk from occupational and environmental exposure to polycyclic aromatic hydrocarbons. Cancer Causes Control 8:444–472.

Checkoway H, Pearce NE, Crawford-Brown DJ. 1989. Research Methods in Occupational Epidemiology. Oxford University Press, New York, NY.

Cohen AJ, Higgins MWP. 1995. Health effects of diesel exhaust: Epidemiology. In: Diesel Exhaust: A Critical Analysis of Emissions, Exposure, and Health Effects. A Special Report of the Institute's Diesel Working Group. Health Effects Institute, Cambridge, MA.

Cold Spring Harbor Laboratory. 1981. Banbury Report 9: Quantification of Occupational Cancer. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Crump KS. 1999. Lung cancer mortality and diesel exhaust: Reanalysis of a retrospective cohort study of U.S. railroad workers. Inhalation Toxicol 11:101–107.

Crump KS, Lambert T, Chen C. 1991. Assessment of risk from exposure to diesel engine emissions: Report to U.S. Environmental Protection Agency. Contract 68-02-4601 (Work Assignment No. 182, July). Office of Health Assessment, U.S. Environmental Protection Agency, Washington, DC.

Davis SC. 1998. Transportation Energy Data Book (18th ed.). Oak Ridge National Laboratory, Oak Ridge, TN.

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

Doll R, Peto R. 1981. The Causes of Cancer: Quantitative Estimates of Avoidable Risks of Cancer in the U.S. Today. Oxford University Press, New York, NY.

Dosemeci M, Stewart PA. 1996. Recommendations for reducing the effects of exposure misclassification on relative risk estimates. Occup Hyg 3:169–176.

Garshick E. 1991. Personal letter of August 15 to Chao Chen, Office of Health Assessment, U.S. Environmental Protection Agency, Washington, DC. [Source: Eric Garshick, M.D., M.O.H., Pulmonary and Critical Care Section, Brockton/West Roxbury Veterans Administration Medical Center, 1400 VFW Parkway, West Roxbury, MA 02132; garshick_MD.eric_z+@brockton.va.gov.]

Garshick E. 1998. Comments to the California State Senate Transportation Committee, August 4, 1998, Sacramento, CA. [Source: Eric Garshick, M.D., M.O.H., Pulmonary and Critical Care Section, Brockton/West Roxbury Veterans Administration Medical Center, 1400 VFW Parkway, West Roxbury, MA 02132; garshick_MD.eric_z+@brockton.va.gov.]

Garshick E, Schenker MB, Muñoz A, Segal M, Smith TJ, Woskie SR, Hammond SK, Speizer FE. 1987. A case-control study of lung cancer and diesel exhaust exposure in railroad workers. Am Rev Respir Dis 135:1242–1248.

Garshick E, Schenker MB, Muñoz A, Segal M, Smith TJ, Woskie SR, Hammond SK, Speizer FE. 1988. A retrospective cohort study of lung cancer and diesel exhaust exposure in railroad workers. Am Rev Respir Dis 137:820–825.

Gordis L, ed. 1988. Epidemiology and Health Risk Assessment. Oxford University Press, New York, NY.

Gordis L. 1996. Epidemiology. WB Saunders Co., Philadelphia, PA.

Graboski MS, McCormick RL, Yanowitz J, Ryan L. 1998. Heavy-Duty Diesel Vehicle Testing for the Northern Front Range Air Quality Study. Colorado Institute for Fuels and High-Altitude Engine Research, Colorado School of Mines, Golden, CO.

Hammond SK, Smith TJ, Woskie SR, Leaderer BP, Bettinger N. 1988. Markers of exposure to diesel exhaust and cigarette smoke in railroad workers. Am Ind Hyg Assoc J 49:516–522.

Health Effects Institute. 1995. Diesel Exhaust: A Critical Analysis of Emissions, Exposure, and Health Effects. A Special Report of the Institute's Diesel Working Group. Health Effects Institute, Cambridge, MA.

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

Hertz-Picciotto I. 1995. Epidemiology and quantitative risk assessment: A bridge from science to policy. Am J Public Health 85:484–491.

Hulka BS, Wilcosky TC, Griffith JD. 1990. Biological Markers in Epidemiology. Oxford University Press, New York, NY.

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 Life Sciences Institute Risk Science Institute. 1999. Workshop Consensus Report: Relevance of the rat as a model for hazard and risk characterization of inhaled poorly soluble, nonfibrous, nongenotoxic particles. Inhalation Toxicol (in press).

Kelsey JL, Whittemore AS, Evans AS, Thompson WD. 1996. Methods in Observational Epidemiology (2nd ed.). Oxford University Press, New York, NY.

Monson RR. 1980. Occupational Epidemiology. CRC Press, Boca Raton, FL.

Morgan WKC, Reger RB, Tucker DM. 1997. Health effects of diesel emissions. Ann Occup Hyg 41:643–656.

Muscat JE, Wynder EL. 1995. Diesel engine exhaust and lung cancer: An unproven association. Environ Health Perspect 103:812–818.

National Academy of Sciences. 1983. Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington, DC.

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 Nonmetal Miners (NCI-NIOSH Interagency Project Protocol, June 20, 1997). [Source: Michael Attfield, NIOSH/RDS/FSB MS P-163, 1095 Willowdale Road, Morgantown, WV 26505; (304) 285-5737; MDA1@CDC.GOV.]

National Institute for Occupational Safety and Health. 1988. Carcinogenic Effects of Exposure to Diesel Exhaust. NIOSH Current Intelligence Bulletin 50. DHHS (NIOSH) Publication No. 88-116. Centers for Disease Control, Atlanta, GA.

National Toxicology Program. 1998. Report on Carcinogens: Background Document for Diesel Exhaust Particulates. U.S. Department of Health and Human Services, Research Triangle Park, NC.

Northern Front Range Air Quality Study (NFRAQS). 1998. Final Report. DRI Document No. 6580-685-8750.1F2. Desert Research Institute, Reno, NV.

Office of Environmental Health Hazard Assessment. 1994. Health Risk Assessment for Diesel Exhaust: Preliminary Draft, June 1994. California Environmental Protection Agency, Sacramento, CA.

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

Preston DL, Lubin JH, Pierce DA. 1991. EPICURE User's Guide. HiroSoft International Corporation, Seattle, WA.

Railroad Facts. 1940, 1944, 1946, 1948, 1950, 1953, 1954, 1957, 1964, 1967, 1970. Association of American Railroads, Washington, DC.

Rothman KJ, Greenland S. 1998. Modern Epidemiology (2nd ed.). Lippincott-Raven, Philadelphia, PA.

Säverin R, Bräunlich A, Enderlein G, Heuchert G. 1998. A cohort study on lung cancer and diesel exhaust in potash miners (abstract). Dortmund, Berlin, Germany.

Samet JM, Schnatter R, Gibb H. 1998. Invited commentary: Epidemiology and risk assessment. Am J Epidimiol 148:929–936.

Sawyer R, Johnson J. 1995. Diesel emissions and control technology. In: Diesel Exhaust: A Critical Analysis of Emissions, Exposure, and Health Effects. A Special Report of the Institute's Diesel Working Group. Health Effects Institute, Cambridge, MA.

Schlesselman JJ. 1982. Case-Control Studies. Oxford University Press, New York, NY.

Stöber W, Abel U. 1996. Lung cancer due to diesel soot particles in ambient air? A critical appraisal of epidemiological studies addressing this question. Int Arch Occup Environ Health (Suppl) 68:S3–S61.

Selvin S. 1991. Statistical Analysis of Epidemiologic Data. Oxford University Press, New York, NY.

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 NK, Silverman DT, Hornung RW. 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. Exposures to diesel exhaust in the trucking industry and possible relationships with lung cancer. Am J Ind Med 21:887–890.

U.S. Department of Labor. 1972. Bureau of Labor Statistics: Railroad Technology and Manpower in the 1970s. U.S. GPO Bulletin 1717:71. U.S. Government Printing Office, Washington, DC.

U.S. Environmental Protection Agency. 1994. Health Assessment Document for Diesel Emissions: External Review Draft. EPA 600/8-90/057-a. Office of Research and Development, Washington, DC.

U.S. Environmental Protection Agency. 1997. National Air Pollutant Emissions Trends Report 1990–1996. EPA-454/R-97-011. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

U.S. Environmental Protection Agency. 1998. Health Assessment Document for Diesel Emissions: SAB Review Draft. EPA/600/8-90/057C. Office of Research and Development, Washington, DC.

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

Woskie SR, Hammond SK, Smith TJ, Schenker MB. 1989. Current nitrogen dioxide exposures among railroad workers. Am Ind Hyg Assoc J 50:346–353.

Woskie SR, Smith TJ, Hammond SK, Schenker MB, Garshick E, Speizer FE. 1988a. Estimation of the diesel exhaust exposures of railroad workers: I. Current exposures. Am J Ind Med 13:381–394.

Woskie SR, Smith TJ, Hammond SK, Schenker MB, Garshick E, Speizer FE. 1988b. Estimation of the diesel exhaust exposures of railroad workers: II. National and historical exposures. Am J Ind Med 13:395–404.

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

Ziskind R, Carlin T, Ballas J. 1978. Evaluation of gas hazards inside heavy-duty diesel truck cabs. Proceedings of the American Chemical Society, Fourth Joint Conference on Sensing Environmental Pollutants, pp. 377–383. American Chemical Society, Washington, DC.

Abbreviations

AEM	adjusted extracted mass
ARP	adjusted respirable particles
CI	confidence interval
CVD	cardiovascular disease
DPM	diesel particulate matter
EC_1	elemental carbon with an aerodynamic diameter of 1 µm or less
EPA	Environmental Protection Agency (California or U.S.)
ETS	environmental tobacco smoke
IARC	International Agency for Research on Cancer
ICD 8	International Classification of Disease, Version 8
ILSI	International Life Sciences Institute
NCI	National Cancer Institute
NFRAQS	Northern Front Range Air Quality Study
NIOSH	National Institute of Occupational Safety and Health
NO_x	oxides of nitrogen
NTP	National Toxicology Program
OEHHA	Office of Environmental Health Hazard Assessment
QRA	quantitative risk assessment
RR	relative risk
RRB	Railroad Retirement Board
RSP	respirable-sized particles
TEX	total extractable material
VMT	vehicle miles traveled
WHO	World Health Organization

Related HEI Publications: Diesel Engine Exhaust

Research Reports

Report No.	Title	Principal Investigator	Publication Date
Constituents of Diesel Exhaust: Metabolism, Biomarkers, and Carcinogenicity			
2	Disposition and Metabolism of Free and Particle-Associated Nitropyrenes After Inhalation	J. Bond	1986
4	The Metabolic Activation and DNA Adducts of Dinitropyrenes	F. Beland	1986
7	DNA Adducts of Nitropyrene Detected by Specific Antibodies	J.D. Groopman	1987
16	Metabolism and Biological Effects of Nitropyrene and Related Compounds	C.M. King	1988
17	Studies on the Metabolism and Biological Effects of Nitropyrene and Related Nitro-Polycyclic Aromatic Compounds in Diploid Human Fibroblasts	V.M. Maher	1988
19	Factors Affecting Possible Carcinogenicity of Inhaled Nitropyrene Aerosols	R.K. Wolff	1988
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.A. Beland	1989
32	Respiratory Carcinogenesis of Nitroaromatics	R.C. Moon	1990
33	Markers of Exposure to Diesel Exhaust in Railroad Workers	M.B. Schenker	1990
34	Metabolic Activation of Nitropyrene and Diesel Particulate Extracts	A.M Jeffrey	1990
46	Role of Ring Oxidation in the Metabolic Activation of 1-Nitropyrene	F.A. Beland	1991
55	Mutations Induced by 1-Nitrosopyrene and Related Compounds During DNA Replication in Human Cells and Induction of Homologous Recombination by These Compounds	V.M. Maher	1993
61	Methods Development Toward the Measurement of Polyaromatic Hydrocarbon–DNA Adducts by Mass Spectrometry	R. Giese	1993
64	Biomonitoring of Nitropolynuclear Aromatic Hydrocarbons via Protein and DNA Adducts	K. El-Bayoumy	1994
66	The Effects of Copollutants on the Metabolism and DNA Binding of Carcinogens	P. Howard	1994
68	Pulmonary Toxicity of Inhaled Diesel Exhaust and Carbon Black in Chronically Exposed Rats		
	Part I: Neoplastic and Nonneoplastic Lung Lesions	J.L. Mauderly	1994
	Part II: DNA Damage	K. Randerath	1995

(Continued on next page)

Research Reports

Report No.	Title	Principal Investigator	Publication Date
72	DNA Adduct Formation and T-Lymphocyte Mutation Induction in F344 Rats Treated With 1,6-Dinitropyrene	F.A. Beland	1995
84	Evaluation of the Potential Health Effects of the Atmospheric Reaction Products of Polycyclic Aromatic Hydrocarbons	J. Arey	1999
Dosimetry			
10	Predictive Models for Deposition of Inhaled Diesel Exhaust Particles in Humans and Laboratory Species	C.P. Yu	1987
40	Retention Modeling of Diesel Exhaust Particles in Rats and Humans	C.P. Yu	1991
Emissions			
5	An Investigation into the Effect of a Ceramic Particle Trap on the Chemical Mutagens in Diesel Exhaust	S.T. Bagley	1987
56	Characterization of Particle- and Vapor-Phase Organic Fraction Emissions from a Heavy-Duty Diesel Engine Equipped with a Particle Trap and Regeneration Controls	S.T. Bagley	1993

Special Reports

	Diesel Exhaust: A Critical Analysis of Emissions, Exposure, and Health Effects (In press)	1995

HEI Communications

7	Diesel Workshop: Building a Research Strategy to Improve Risk Assessment	1999

Copies of these reports can be obtained by writing or calling the Health Effects Institute, 955 Massachusetts Avenue, Cambridge, MA 02139. Phone (617) 876-6700 Fax (617) 876-6709 E-mail pubs@healtheffects.org www.healtheffects.org

HEI

The Board of Directors

Archibald Cox *Chairman*
Carl M. Loeb University Professor (Emeritus), Harvard Law School

Douglas Costle
Chairman of the Board and Distinguished Senior Fellow, Institute for Sustainable Communities

Alice Huang
Senior Councilor for External Relations, California Institute of Technology

Donald Kennedy
President (Emeritus) and Bing Professor of Biological Sciences, Stanford University

Susan B. King
Fellow, Sanford Institute of Public Policy, Duke University

Richard B. Stewart
Professor, New York University School of Law

Robert M. White
President (Emeritus), National Academy of Engineering, and Senior Fellow, University Corporation for Atmospheric Research

Health Research Committee

Bernard D. Goldstein *Chairman*
Director, Environmental and Occupational Health Sciences Institute

Glen R. Cass
Professor of Environmental Engineering and Mechanical Engineering, California Institute of Technology

Seymour J. Garte
Professor of Environmental and Community Medicine, Environmental and Occupational Health Sciences Institute

Rogene Henderson
Senior Scientist, Lovelace Respiratory Research Institute

Meryl H. Karol
Professor of Environmental and Occupational Health, University of Pittsburgh Graduate School of Public Health

Stephen I. Rennard
Larson Professor, Pulmonary and Critical Care Medicine Section, Department of Internal Medicine, University of Nebraska Medical Center

Jonathan M. Samet
Professor and Chairman, Department of Epidemiology, School of Hygiene and Public Health, Johns Hopkins University

Robert F. Sawyer
Class of 1935 Professor of Energy (Emeritus), Professor of the Graduate School, University of California, Berkeley

Frank E. Speizer
Edward H. Kass Professor of Medicine, Channing Laboratory, Harvard Medical School, Department of Medicine, Brigham and Women's Hospital

Gerald van Belle
Professor, Departments of Environmental Health and Statistics, University of Washington

Health Review Committee

Daniel C. Tosteson *Chairman*
Professor of Cell Biology, Dean Emeritus, Harvard Medical School

John C. Bailar III
Professor, Department of Health Studies, Biological Sciences Division, The University of Chicago

A. Sonia Buist
Professor of Medicine and Physiology, Oregon Health Sciences University

Ralph D'Agostino
Professor of Mathematics/Statistics and Public Health, Boston University

Thomas W. Kensler
Professor, Division of Toxicological Sciences, Department of Environmental Sciences, Johns Hopkins University

Edo D. Pellizzari
Vice President for Analytical and Chemical Sciences, Research Triangle Institute

Donald J. Reed
Distinguished Professor of Biochemistry, Department of Biochemistry and Biophysics, and Environmental Health Sciences Center, Oregon State University

David J. Riley
Professor of Medicine, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School

Sverre Vedal
Professor of Medicine, University of British Columbia

Officers and Staff

Daniel S. Greenbaum *President*
Richard M. Cooper *Corporate Secretary*
Howard E. Garsh *Director of Finance and Administration*
Kathleen M. Nauss *Director for Scientific Review and Evaluation*
Robert M. O'Keefe *Director of Program Strategy*
Jane Warren *Director of Research*
Aaron J. Cohen *Senior Scientist*
Maria G. Costantini *Senior Scientist*
Alison Geyh *Staff Scientist*
Bernard Jacobson *Staff Scientist*
Debra A. Kaden *Senior Scientist*
Diane J. Mundt *Staff Scientist*
Martha E. Richmond *Staff Scientist*
Geoffrey H. Sunshine *Staff Scientist*
Gail V. Allosso *Office and Contracts Manager*
Thomas Atwood *Manager of Publications and Information*
Julia F. Campeti *Publications Assistant*
Terésa Fasulo *Senior Administrative Assistant*
L. Virgi Hepner *Senior Scientific Editor*
Darlene Jones *Senior Administrative Assistant*
Judith Lopez *Receptionist*
Francine Marmenout *Senior Executive Assistant*
Teresina McGuire *Accounting Assistant*
Beverly Morse *Administrative Assistant*
Jacqueline C. Rutledge *Controller*
Hope Steele *Acting Managing Editor*

HEI *A Partnership of the U.S. Environmental Protection Agency and Industry*

Health Effects Institute

Phone 617.876.6700
Fax 617.876.6709
E-mail *pubs@healtheffects.org*
Web *http://www.healtheffects.org*

955 Massachusetts Avenue
Cambridge, MA 02139