Responses by the Staff of the Office of Environmental Health Hazard Assessment (OEHHA) to the March, 1997 Draft Technical Support Document (TSD) (Including Part B, "Health Risk Assessment for Diesel Exhaust") For Identification of Diesel Exhaust as a Toxic Air Contaminant (TAC)

Table of Contents

Comments of Association of American Railroads....2
Comments of U.S. Borax....9
Comments of California Trucking Association (August 22, 1997)....11
Comments of California Trucking Association (Anderson, September 25, 1997)....24
Comments of Dr. Kenny Crump on behalf of ICF Kaiser and Mercedes-Benz....26
Comments of John Duerr on behalf of Detroit Diesel....41
Comments of Engine Manufacturers Association (Keller, August 1, 1997)....45
Comments of Engine Manufacturers Association (French, August 22, 1997)....47
Comments of Engine Manufacturers Association (Gradient Corp)....72
Comments of Engine Manufacturers Association (Cox)....119
Comments of Dr. Erik Garshick....155
Comments of Daniel Greenbaum on behalf of Health Effects Institute....163
Joint Comments....170
Comments of Dr. Joe Mauderly of Lovelace Research Institute....173
Comments of Roger McClellan on behalf of Chemical Industry Institute of Toxicology....194
Comments of Mercedes-Benz....197
Comments of Dr. Suresh Moolgavkar....211
Comments of Joshua Muscat on behalf of American Health Foundation....216
Comments of Natural Resources Defense Council....223
Comments of Western States Petroleum Association....231

Comments of the Association of American Railroads, letter dated August 21, 1997 from Michael J. Rush to Genevieve Shiroma

Comment 1: The draft risk assessment's conclusion that diesel exhaust is carcinogenic is based on rat and epidemiology studies. At the July 1 workshop, Dr. Joe Mauderly explained that data from rat studies cannot be used to assess the effect of diesel exhaust on humans because rat lungs react very differently to diesel exhaust exposure than do human lungs. Dr. Mauderly cited recent studies finding that tumors in rats result from overloading the clearance mechanism of rat lungs, a phenomenon which does not occur in humans. Dr. Mauderly's findings are especially significant since the draft risk assessment relies extensively on his rat studies. (See also letter from Roger McClellan, Chemical Institute of Toxicology, to Genevieve Shiroma, California Air Resources Board (June 30, 1997), in which Dr. McClellan explains that "rat data are not relevant for use in assessing human lung cancer risk of ambient exposure to diesel exhaust or other particulate matter.")

Dr. Mauderly's presentation effectively eliminates any justification for reaching conclusions, based on rat data, about the effect of diesel exhaust on humans. There is nothing in the draft health risk assessment contradicting Dr. Mauderly's findings. Furthermore, diesel exhaust has not been found to cause cancer in other animals.

***Response**: As incorporated and discussed elsewhere in Part C, Dr. Mauderly has stated his view that "it is not appropriate to use existing lung tumor data from rats to generate quantitative estimates of unit human lung cancer risks from environmental exposures to diesel soot". It appears that Dr. Mauderly's views are not as broad as those indicated by the comment.*

With respect to the comments of Dr. Mauderly, we shared them with the Scientific Review Panel on October 16, 1997. There was a thorough discussion of the issue of using the rat data for quantitative human risk assessment at the meeting. The sense of the panel was that rat data and calculations provide useful information and should be left in the document, however, since human epidemiologic evidence was available on which to base the human risk estimate, that the human data should be used to form the range of risks. Therefore, OEHHA now bases the range of unit risk estimates only on the epidemiological information.

Staff do not agree that all available scientific evidence is consistent with the "particle overload" hypothesis. Sections 6 and 7 of this document state that the mechanism of action by which diesel exhaust induces lung tumors in rats is not established. One proposed mechanism is that exposure to diesel exhaust particulate matter at high concentrations exceeds pulmonary clearance capabilities and causes chronic inflammation. This inflammation leads to macrophage and/or neutrophil-induced oxidative DNA damage resulting in mutations which are instrumental in the induction of lung tumors, and also to cell proliferation which may be mechanistically important to the promotion of the rat lung tumors. This mechanism has also been invoked for carcinogenicity caused by other insoluble particles (e.g. carbon black, titanium dioxide). Rat lung tumor induction due to high dose (2.5 mg/m^3 or higher) exposure to diesel exhaust may share some commonality of mechanism with other carcinogenic insoluble particles; this possibility is discussed in the document. However, genotoxicity due to 1) the PAH and nitroPAH content of

diesel exhaust, and 2) possible oxidative DNA damage primarily due to diesel exhaust exposure may play a role in the induction of lung tumors in rats at lower levels of diesel exhaust. A recent report by Borm et al. (1997) indicates that incubating rat lung epithelial-derived cells with human polymorphonuclear lymphocytes (PMN) (either unactivated or activated by preexposure to phorbol myristate acetate) increases DNA adduct formation caused by exposure to benzo[a]pyrene; addition of more activated PMN in relation to the number of lung cells further increased adduct formation in a dose-dependent manner. The authors suggest that "an inflammatory response in the lung may increase the biologically effective dose of polycyclic aromatic hydrocarbons (PAHs), and may be relevant to data interpretation and risk assessment of PAH-containing particulates." These data raise the possibility that low dose diesel exhaust exposure may result in levels of neutrophil influx which would not necessarily be detectable via histopathological examination as acute inflammation but which might be effective at amplifying any potential diesel exhaust genotoxic effect. WHO (1996) has noted that modeling of human cancer risk from rat lung tumor data should take into account the effects of both particles (carbon core) and extractable organic matter (PAHs, nitro PAHs).

Additionally, some parameters of the "particle overload" hypothesis are incompletely characterized. No data exists on the claimed inadequacy of rat lung antioxidant levels. Alveolar type II cell epithelial hyperplasia has been noted after diesel exhaust exposure, but the measures of cell proliferation used were relatively crude and unsuitable for use in a quantitative estimate of cell proliferation as would be required for biologically-based modeling. It should also be noted that uncertainties exist regarding the magnitude and biological importance of particle overload for diesel exhaust-induced rat lung carcinogenicity. Hattis and Silver (1994) examined lung burden data from diesel exhaust rat carcinogenicity studies and came to the conclusion that "there is continuing accumulation of diesel-derived dust in the lungs of rats throughout life, even at low doses". They also found that this was not predicted by models developed to represent diesel exhaust particulate matter accumulation under "overload" versus nonoverload conditions. Finally, they have found that at high diesel exhaust exposure levels, the increase in the ration of internal diesel exhaust particulate matter burden to external exposure is not very large, being slightly larger than a factor of 2 at most, and state that "Although dust overloading is a real phenomenon, it is not a very large effect and thus would not be expected to give rise to dramatically lowered estimates of risk at low exposure levels."

Finally, Section 6.1.1.1 of this document lists two studies (Pepelko and Peirano, 1983; Heinrich et al., 1986a) in which the authors describe statistically significant lung tumor induction in mice in response to diesel exhaust exposure. Another study (Takemoto et al., 1986) is also listed in which the authors reported increased tumor incidences in mice which were not statistically significant, but which were later reported by IARC (1990) to be statistically significant after the data were reanalyzed. Section 6.1.3 has been changed to include a study by Ichinose et al. (1997) which describes lung tumor induction in mice after intratracheal exposure to diesel exhaust particulate matter. Of the six mouse diesel exhaust inhalation or intratracheal instillation carcinogenicity studies described in this document, three definitely contain positive results, one probably contains positive results, and two are negative. This indicates that the description of the mouse carcinogenicity results as "mixed" is justified, and it cannot be accurately stated that the only rodent species susceptible to diesel exhaust-induced lung tumors is the rat.

It is therefore premature to conclude that the carcinogenic response in rats to diesel exhaust is completely nonspecific, or that it is not relevant to identifying potential human cancer risk.

Comment 2: Nor do epidemiology studies support the proposed findings concerning the carcinogenicity of diesel exhaust. OEHHA's calculation of the unit risk of lung cancer from diesel exhaust is based on Dr. Roger McClellan's calculation of the expected increase in lung cancer deaths attributable to diesel exhaust. (See Office of Environmental Health Hazard Assessment, California Environmental Protection Agency, *Health Risk Assessment for Diesel Exhaust,* 7-16, 7-17 (May 1997)). Dr. McClellan's calculation was based on the Garshick case-control study of railroad workers. However, Dr. McClellan now believes his earlier conclusions are unsound because the data used by Garshick cannot support a calculation of cancer risk from diesel exhaust exposure.

On June 30, Dr. McClellan sent a letter to ARB explaining that reanalysis of the Garshick data shows that it cannot be used to estimate cancer risk from diesel exhaust:

> "In my professional opinion, the [OEHHA] report overinterprets the findings of the various epidemiological studies including the key studies by Garshick *et al.* … My earlier confidence in these studies led me to use them for developing quantitative estimates of the potency of diesel exhaust.
>
> "However, a rigorous reanalysis of the Garshick *et al.* data raises serious questions about the validity of the conclusions drawn by Garshick *et al.* … Although I and my colleagues have previously published [cancer potency estimates for exposure to diesel exhaust] and they are referenced in the report, these estimates are no longer valid because of the more recent findings…"

The reanalysis of the Garshick data referred to by Dr. McClellan is the work of Dr. Kenny Crump. Dr. Crump was hired by U.S. EPA to conduct a risk analysis of diesel exhaust. However, on the basis of Dr. Crump's analysis of the epidemiological evidence, EPA concluded that the data could not support a risk analysis. (See Environmental Protection Agency, *Health Assessment Document for Diesel Emissions: External Review Draft,* 1-22 (Sept. 1994) ("the evidence of carcinogenicity in humans falls short of being sufficient"). At the July 1 workshop, Dr. Crump cited two significant statistical anomalies in the data used by Garshick:

> "Lung cancer mortality was not significantly elevated among shopworkers in comparison to clerks and signalmen, despite the fact that shopworkers likely had the most intense exposures of any group;" and
>
> "Relative risk of lung cancer tended to decrease with increasing duration of exposure within exposed workers."

Consequently, Dr. Crump has concluded that the Garshick cohort study did not convincingly show that diesel exhaust caused lung cancer.

Related to the statistical anomalies found by Dr. Crump is the general lack of information on railroad worker exposure to diesel exhaust. This problem has led Dr. Garshick himself to conclude that "it is not possible to use the human epidemiologic data … to assign a unit risk with confidence due to the uncertainty of the exposure data," (letter from Eric Garshick to Randall Bond, May 30, 1995).

The draft risk assessment makes no attempt to be scientific in addressing the inconsistency between the shopworker data and the conclusions reached by the draft. The draft assessment simply excludes the shopworker data from its calculations. The assessment explains that shop workers were excluded because of the uncertainty concerning the exposure of shop workers to diesel exhaust. However, that explanation is unconvincing because, as Dr. Garshick has stated, the exposure data for all job categories is insufficient to support definitive conclusions. The reader is left with the conclusion that the shopworker data were excluded because the data do not support the thesis posed by the assessment, that diesel exhaust is carcinogenic.

The draft risk assessment also is very unscientific in interpreting the findings of the Health Effects Institute (HEI) in this area. The assessment states that "HEI found that the epidemiological data are consistent in showing weak associations between exposure to diesel exhaust and lung cancer." Inexplicably, the assessment fails to cite HEI's conclusion that "the lack of definitive exposure data for the occupationally exposed study populations precludes using the available epidemiologic data to develop quantitative estimates of cancer risk. " (Health Effects Institute, *Diesel Exhaust: A Critical Analysis of Emissions, Exposure, and Health Effects,* (April 1995). Incredibly, the draft risk assessment also implies that HEI supports the use of rat data to draw conclusions about the carcinogenicity of diesel exhaust: "HEI also found that the carcinogenicity of diesel exhaust had been convincingly demonstrated in rats." *Health Risk Assessment, supra* n. 2, at 15. In fact, HEI questions "the validity of using the rat bioassay data to characterize the potential human risk associated with ambient exposure to diesel emissions" because "the mechanism of lung tumor induction that appears to operate in rats continuously exposed to high concentrations of diesel exhaust ... may not be relevant to most humans…" *Diesel Exhaust* at 8.)

Thus, it is no surprise that neither U.S. EPA, the Health Effects Institute, nor any other organization has used epidemiology studies for risk assessment.(See Presentation of Peter A. Valberg, Gradient Corp., "Comments on OEHHA's Assessment of Diesel Exhaust (DE),Presentation to CARB Workshop, at 8 (July 1, 1997) (noting that neither the U.S. Environmental Protection Agency, the World Health Organization, U.S. EPA's Clean Air Scientific Advisory Committee, nor the Health Effects Institute has used the epidemiology results for risk assessment). Outside OEHHA, at least, the scientific consensus is that the epidemiology studies to date do not support a risk assessment.

Response*: With respect to a qualitative relationship between occupational exposure to diesel exhaust and lung cancer, OEHHA's present conclusions can be compared to the conclusions of the U.S.EPA, the HEI, and the WHO:*

The 1996 OEHHA draft states "The epidemiological studies concerning lung cancer risk and exposure to diesel exhaust provide evidence consistent with a causal relationship. The many associations found between lung cancer and diesel exposure are unlikely to be due to chance. Also, with the possible exception of the studies that did not take smoking into account, the findings reviewed above are unlikely to be due to confounding or bias. The results of various studies are consistent in the direction of an effect and are even somewhat similar in magnitude of effect. For example, all the studies of diesel railroad workers with adequate latency and more than 50 cases show evidence of an effect. Although the strength of the associations is weak, with low relative risk estimates being reported, several studies show clear exposure-response relationships. Finally, it is biologically plausible because of its mutagenic and carcinogenic constituents, that exposure to diesel exhaust would increase the risk of lung cancer. Therefore, a reasonable and likely explanation for the increased risks of lung cancer observed in the epidemiologic studies is a causal association between diesel exhaust exposure and lung cancer."

The 1994 U.S.EPA draft document states "Collectively, the epidemiology studies show evidence of an association between inhalation of diesel exhaust and lung cancer in humans. Although, the evidence for carcinogenicity in humans was in most cases positive, it is judged to be limited according to the EPA's weight of evidence guidelines, because the observed increases in risk were quite low and the influence of confounding factors could not be completely accounted for." Similarly, the HEI concluded in 1995 that "The available evidence suggests that occupational exposure to diesel exhaust from diverse sources increases the lung cancer rate by 20% to 40% in exposed workers generally and to a greater extent among workers with prolonged or intense exposure or both. These results are not readily explicable by confounding due to cigarette smoking or other known sources of bias." While the WHO in 1996 found "The relative risks for lung cancer as a result of exposure to diesel exhaust are generally low, and risks of this magnitude are more susceptible to chance and to the effects of unmeasured confounding factors and imprecision for adjusting for known confounding factors. As discussed above, the elevated risk for lung cancer in the four most informative studies is unlikely to be due to confounding by cigarette smoking and is probably due to exposure to diesel exhaust. Other studies, although limited primarily by the exposure ascertainment, support this assessment."

As noted by the commentator, with respect to the quantitative assessment of risk based upon the occupational epidemiology studies, OEHHA's position differs from those of a number of individual scientists, including the current position of Dr. McClellan, and also from the previous positions of several authoritative bodies such as the HEI and the WHO who generally have found that the inadequacy of exposure information limited the ability to conduct a quantitative risk assessment. However, in reaching their conclusions, they each relied to varying extents upon the results of the dose response analyses of the Garshick et al. 1988 cohort study data conducted by Dr. Crump. In those analyses, Dr. Crump did not obtain significantly positive dose response relationships between the incidence of lung cancer and diesel exhaust exposure. In its own later analyses using different methods, OEHHA obtained positive dose response relationships. This difference between some of the conclusions of OEHHA and these other authoritative bodies therefore depend in important part upon the comparative merits of the approaches of OEHHA and Dr. Crump. OEHHA, having characterized and reviewed the sources of this difference, finds its approach to be more appropriate. The differences are the subject of Appendix F of Part B, as

well as Part C here and previously, and have been highlighted for independent review by the Scientific Review Panel.

OEHHA feels that the comment overstates the U.S.EPA position on the use of epidemiological data in quantitative risk assessment. The 1994 U.S.EPA document does state that the human data are inadequate for quantitative risk assessment. However, within the context of the U.S.EPA and OEHHA documents, this inconsistency with OEHHA is only apparent. The U.S.EPA document found the epidemiological data inadequate in the context of deriving a preferred single point estimate for diesel exhaust's carcinogenic potency. The U.S.EPA therefore derived a point estimate from the animal cancer data. OEHHA did not develop a single point estimate.

OEHHA regards the U.S.EPA position as being consistent with use of epidemiological data in framing the range of cancer risks. This interpretation is consistent with the U.S.EPA presentation of the McClellan results in its Executive Summary and the subsequent statement found on page 1-26 of the U.S.EPA's 1994 draft document Executive Summary which states "In view of the uncertainty inherent in these types of calculations, the human and animal estimates should be viewed as complementary. For a bounding estimate intended to determine whether an exposure level as the potential to pose a hazard to human health, the published human estimates [the Harris, McClellan studies] may be practical for exposure levels in the range of observations in these studies. On the other hand projection of the public health impact of an exposure level may benefit from using estimates derived from animal experiments, because of the closely controlled conditions and their precisely measured levels, absence of many confounding factors, and narrow confidence limits around the tumor incidence rates. A unit risk estimate of $3.4 \times 10^{-5}/\ \mu/m^3$ for continuous life-time exposure, which is the geometric mean of the upper bound estimates calculated from the three rat experiments is therefore recommended. The proper use and understanding of these risk estimates is discussed in Chapter 12."

The OEHHA Executive Summary accordingly states "In their 1994 draft document, the U.S.EPA cited these same McClellan et al (1986) UCL values as being practical in assessing human risks involving exposures in the range of study observations." OEHHA took care in characterizing the U.S.EPA work. OEHHA requested the U.S.EPA review our draft language prior to its public release. They did not object to this characterization of their work. In discussions of the OEHHA analyses with U.S.EPA staff, we have found them to be generally supportive of including epidemiological data in framing a range of risks. At the July 1, 1997 workshop Dr. Koppikar indicated that the U.S.EPA would be using epidemiological data from the Garshick et al. (1987) case-control study in the next quantitative risk assessment it presents to the Clean Air Scientific Advisory Committee. Dr. Koppikar also stated that the resulting risk values were "pretty consistent and very similar with what Cal-EPA has presented here today." (Transcript of the Public Workshop for the Diesel Exhaust Identification Report, p.75.)

Risk assessment is one of the more complex analytical functions undertaken by government, and there is much opportunity for reasonable minds to differ with respect to the judgments involved. With regard to diesel exhaust, where the health effects literature is particularly voluminous and complex, there is much opportunity for honest disagreement. With respect to hazard identification, the meta-analysis used the Garshick et al. (1988) cohort relative risk value as it was reported.

With respect to dose response assessment, elimination of the shop worker data does not, in fact, substantially alter the conclusions of the OEHHA dose response analysis. Any conclusion that the "shopworker data were excluded because the data do not support the thesis posed by the assessment, that diesel exhaust is carcinogenic," is therefore without basis.

With respect to the comment's concerns regarding any selective citation of the HEI findings, our statement that the "HEI also found that the carcinogenicity of diesel exhaust had been convincingly demonstrated in rat," is found in the Executive Summary section (Section 1.3, Carcinogenicity) which addresses hazard identification. This section was not meant to reach the dose response assessment issue as suggested in the comment. Rather, the section of the Executive Summary summarizes, in traditional fashion, both the animal and the human evidence as to the carcinogenicity of diesel exhaust. The HEI position regarding use of the rat findings with respect to extrapolation to "ambient exposures to diesel exhaust" relates more to dose response assessment than hazard identification. It would have been appropriate to describe the HEI position on use of the rat data in quantitative risk assessment in Section 1.4.2 of the Executive Summary (Using the Rat Data to Predict Cancer Risk in Humans). This related concern will be taken into account in the course of updating the Executive Summary. However, the range of unit risk estimates no longer encompasses unit risk values derived from the animal data.

Comments on behalf of U.S. Borax Inc., letter dated August 22, 1997 to Genevieve Shiroma from Mark Ellis

Comment 1: Incorporation by Reference. U.S. Borax is a member of the National Mining Association and the California Mining Association. We have reviewed the comments submitted collectively by the Engine Manufacturers Association, the American Trucking Association and the National Mining Association, as well as those submitted by a coalition of business interests including the California Mining Association. We endorse the comments submitted by these organizations on the revised draft report and adopt them as our own.

***Response**: Comment noted. Responses to the statements of the referenced parties are provided with their comments.*

Comment 2: Public Review Process. The staff of ARB and OEHHA are to be commended on the open, positive and constructive manner in which they have conducted the public review process on the revised draft report and earlier iterations. I have had the opportunity to participate fully in this process in the context of my previous employment with the American Mining Congress and the National Mining Association, as well as on behalf of U.S. Borax. The process has been most revealing.

Particularly telling in this public discourse have been the participation and critical comments of Drs. Joe Mauderly and Eric Garshick, the authors of the principal studies that form the basis of the draft proposal. It is their position that their studies are not appropriate for use in quantitative human health risk assessment. Other prominent scientists and research organizations, including Drs. Kenny Crump, Roger McClellan and the Health Effects Institute agree that Cal-EPA has misinterpreted the data and misapplied their conclusions.

Conclusion. Based on the foregoing, ARB and OEHHA should accept the fact that inherent limitations in the underlying data cannot justify the quantitative risk assessment presented. Based on the best available scientific evidence, as detailed in the comments and testimony identified above, ARB and OEHHA should conclude that it is premature to conduct a reliable quantitative risk assessment of diesel exhaust at this time. Therefore, staff should withdraw its proposal pending further research developments and not recommend that the Board identify diesel exhaust as a toxic air contaminant.

***Response:** With respect to the comments of Dr. Mauderly, we shared them with the Scientific Review Panel on October 16, 1997. There was a thorough discussion of the issue of using the rat data for quantitative human risk assessment at the meeting. The sense of the panel was that rat data and calculations provide useful information and should be left in the document, however, since human epidemiologic evidence was available on which to base the human risk estimate, the human data should be used to form the range of risks. Therefore, OEHHA now bases the range of unit risk estimates only on the human epidemiological information.*

With respect to the comments of Dr. Garshick. It is important to separate his comments regarding the cohort study from those of the case-control study. The comments regarding the concern for

the dose-response relationship and consequently the cancer unit risk have been made regarding the cohort study. Regarding the effect that his railroad worker cohort epidemiologic data could not be reanalyzed to assign a unit risk with confidence, the draft presented a very broad range of unit risks and not a single value of unit risk. Thus, we have taken Dr. Garshick's comments into consideration as we revised the document.

With regard to the work by Dr. McClellan and the "retraction" of his calculations of potency based upon epidemiological data, our draft document was released prior to his "retraction" and could not have addressed it or its merits. However, the retraction does not appear to separate the issues regarding the dose-response relationship in the cohort study from the calculations used in the case-control study. We conducted a focused literature review to evaluate the estimates Dr. McClellan used and have revised the document accordingly.

Furthermore, to different extents, many of the prominent parties cited by the commentator have relied upon the work of Dr. Crump to reach the same or similar conclusions as Dr. Crump. OEHHA has closely examined the scientific data. In conducting its own analyses of the Garshick et al. (1988) cohort study and its own review of the related work by Dr. Crump, OEHHA has come to differ with the Crump approach. Therefore, OEHHA continues to use the Garshick et al. (1988) cohort study data in our analyses, as one of the pieces of information, to develop a plausible range of unit risk factors for diesel exhaust.

More detailed responses to the statements of the referenced parties are provided with their comments.

Comments on behalf of the California Trucking Association, August 22, 1997 on the Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant

Comment 1. The Draft report inaccurately describes diesel exhaust.
OEHHA reports that diesel exhaust is a "...biphasic mixture of substances...with a vapor phase and very fine particles' (1-1). Diesel exhaust is a triphasic mixture of gases, vapors and particles formed in the combustion processes and emitted into the environment.

***Response**: The term biphasic has been deleted.*

Comment: 2. The Draft report inaccurately reports scientific evidence.
Once diesel exhaust leaves the tailpipe, it disperses into its constituent parts and reacts with ambient air to form criteria pollutants and other chemicals. Simply put once it leaves the tailpipe it is no longer diesel exhaust. If OEHHA wants to label whole diesel exhaust a toxic air contaminant, their risk assessment must be conducted in engine test laboratories and research facilities with dilution tunnels, as only these locations can find exposures to whole diesel exhaust.

The Draft Report does not evaluate diesel exhaust. The report jumps from diesel exhaust to the particulate nature of diesel exhaust, and the particulate and gaseous phases of diesel exhaust. OEHHA conducts a health evaluation based on the concentration of the particulate matter, determining without any scientific justification that the mass of diesel particulate matter serves as a "practical surrogate for the source of carcinogenicity in diesel exhaust." OEHHA has conducted a particulate evaluation, not an evaluation of whole diesel exhaust. The Draft Report is fatally flawed as it does not evaluate diesel exhaust, yet names diesel exhaust as the culprit.-

***Response:** The document does not address exposures occurring within a tailpipe. It is concerned with how diesel exhaust emissions may affect public health.*

OEHHA considered both animal and human evidence regarding the potential adverse effects of diesel exhaust on health. The report presents the results of the inhalation animal studies concerning the health effects of whole diesel exhaust or filtered diesel exhaust or the diesel exhaust particulates. Where the results differed, OEHHA pointed out the differences. In addition, the epidemiological studies necessarily examined the effects of whole diesel exhaust as released into the work environment.

The use of the mass of particulate matter as a "practical surrogate" is widely accepted and has been used by the U.S.EPA, and the WHO. The document discusses the uncertainty involved with this approach. It acknowledges that the components of diesel exhaust, subsequent to release into the environment, do not always remain present in the same proportion to one another. It also states that the respirable particulate fraction constitutes a substantial part of the total diesel exhaust by weight and contains many identified carcinogenic components adsorbed onto the surfaces of the particulate matter. Use of the particulate fraction as a surrogate for total diesel emissions is therefore not expected to be a serious limitation.

Comment 3: OEHHA has refused to evaluate the actual chemicals that may be in diesel exhaust. This unusual selection can only be politically motivated and is arbitrary and capricious.

Response: *Diesel exhaust is a complex mixture containing many hundreds, if not thousands, of poorly studied or novel substances of potential concern. As indicated in Part A, a number or the constituents of diesel exhaust have been evaluated individually. Given the lack of knowledge for the myriad of individual components, however, it is not feasible to conduct a constituent-by-constituent approach to describing the toxicity of diesel exhaust. Such an approach could not capture the potential for synergism which exists when so many of the constituents are adsorbed onto the particle itself. OEHHA's approach to such a difficult subject parallels the approach of many researchers in the field.*

Comment 4: OEHHA has technologically discriminated against the compression ignition engine. The draft report links the fine particulates to diesel exhaust and attempts to determine, that the particulate of diesel is a carcinogen. Comparison of the various compounds in the exhaust of different types of engines show particulate phase pollutants from all carbon-based fuels including compressed natural gas, gasoline and diesel. Recent studies completed on particulate size found that the majority of lighter polycyclic aromatic hydrocarbons (PAH) were heavy-duty diesel vehicles, however, light duty vehicles were the major source of higher molecular weight and possibly more toxic PAHs. The authors also address the ultra-fine nature of the gasoline (less than .12 micrometers) and CNG particulate which produce considerably smaller particulates that diesel-fueled engines and carries much worse PAHs. (Miguel *et al.* July 1997), (Greenwood *et al.* 1996) (Greenwood *et al.* 1996)

In general, diesel vehicles have higher emission rates of particulates of the 1 microgram per cubic meter size. On the other hand, gasoline and natural gas engines have higher emissions of the finer, more toxic particulates. The evaluation of gasoline exhaust, based on OEHHA's assumptions regarding the oneness of diesel exhaust and particulate, have very dangerous implications for the same oneness with gasoline exhaust and particulate.

Response*: OEHHA was requested by the Air Resources Board to evaluate diesel exhaust health effects for the Toxic Air Contaminant program. The commentator correctly points out that other types of engines also produce exhaust that may be toxic. However, we are not evaluating those types of exhaust in this document. There is no intent by OEHHA to discriminate against the compression ignition engine. Under the Toxic Air Contaminant Program, OEHHA has conducted health assessments for 21 other substances including asbestos, methylene chloride, and vinyl chloride that may be released by a variety of other sources.*

Comment 5. OEHHA fails to evaluate the latest findings on the carcinogenicity of diesel exhaust in rodents, has misrepresented the animal data, ignores the warnings of principal authors and is at odds with the best available science. OEHHA used animal studies for quantitative risk assessment, disregarding new data showing that the rat data is an outlier. The principal researchers agree that rat data could not be extrapolated to mice, yet OEHHA has extrapolated the data to humans. Dr. Roger McClellan and Dr. Joe Mauderly, principal authors, commented during the July 1. 1997 workshop in opposition to OEHHA's findings. The relevant

scientific community on July 1, 1997 and in prior comments, has rejected CARB/OEHHA ,claim to legitimacy. Diesel Exhaust Identification Report, held July 1, 1997. (RT), p.237, L.15-17, p.228 L.15-17).

Using the rat data is inappropriate. Rats exposed to diesel exhaust, carbon black and carbon particulate without the organic chemical in diesel exhaust had the same tumor development. 'The scientific community has determined the genotoxicity of the material did not cause tumors in rats; instead, it was the process of lung overload which lead to tumor development.

Dan Greenbaum HEI stated at the workshop. "The current document does not go far enough in incorporating this mechanistic information into the risk assessment and dose-response assessments." OEHHA has ignored the facts that tumors are only found in rats after particle overload, and not just with diesel particles. Dr. Mauderly has represented his and other rat studies as unacceptable for generating quantitative estimates of unit human lung cancer risks from environmental exposures to diesel exhaust. OEHHA's conclusions ignore the principal author and do not reflect current or even good science. As stated by Dr. Mauderly, "It is simply inappropriate to generate human cancer risk from the present rat data. The rat data is not relevant to human risk and is invalid for risk assessment." CTA requests that it be removed from OEHHA's analysis.

Response: *With respect to the comments of Dr. Mauderly, we shared them with the Scientific Review Panel on October 16, 1997. There was a thorough discussion of the issue of using the rat data for quantitative human risk assessment at the meeting. The sense of the panel was that rat data and calculations provide useful information and should be left in the document, however, since human epidemiologic evidence was available on which to base the human risk estimate, that the human data should be used to form the range of risks. Therefore, OEHHA now bases the range of unit risk estimates only on the epidemiological information.*

It should also be noted that page 1-26 of the US EPA's 1994 draft document Executive Summary states "In view of the uncertainty inherent in these types of calculations, the human and animal estimates should be viewed as complementary. For a bounding estimate intended to determine whether an exposure level as the potential to pose a hazard to human health, the published human estimates [the Harris, McClellan studies] may be practical for exposure levels in the range ofobservations in these studies. On the other hand projection of the public health impact of an exposure level may benefit from using estimates derived from animal experiments, because of the closely controlled conditions and their precisely measured levels, absence of many confounding factors, and narrow confidence limits around the tumor incidence rates. A unit risk estimate of $3.4 \times 10^{-5}/\ \mu/m^3$ *for continuous life-time exposure, which is the geometric mean of the upper bound estimates calculated from the three rat experiments is therefore recommended. The proper use and understanding of these risk estimates is discussed in Chapter 12." US EPA clearly believes that the rat lung tumor data is relevant to human risk and is appropriate for use in risk assessment. Additionally, in its 1996 document, the World Health Organization (WHO) derived lung cancer risk estimates for diesel exhaust from the rat bioassay data.*

Staff do not agree that all available scientific evidence is consistent with the "particle overload" hypothesis. Sections 6 and 7 of this document state that the mechanism of action by which diesel

exhaust induces lung tumors in rats is not established. One proposed mechanism is that exposure to diesel exhaust particulate matter at high concentrations exceeds pulmonary clearance capabilities and causes chronic inflammation. This inflammation leads to macrophage and/or neutrophil-induced oxidative DNA damage resulting in mutations which are instrumental in the induction of lung tumors, and also to cell proliferation which may be mechanistically important to the promotion of the rat lung tumors. This mechanism has also been invoked for carcinogenicity caused by other insoluble particles (e.g. carbon black, titanium dioxide). Rat lung tumor induction due to high dose (2.5 mg/m^3 or higher) exposure to diesel exhaust may share some commonality of mechanism with other carcinogenic insoluble particles; this possibility is discussed in the document. However, genotoxicity due to 1) the PAH and nitroPAH content of diesel exhaust, and 2) possible oxidative DNA damage primarily due to diesel exhaust exposure may play a role in the induction of lung tumors in rats at lower levels of diesel exhaust. A recent report by Borm et al. (1997) indicates that incubating rat lung epithelial-derived cells with human polymorphonuclear lymphocytes (PMN) (either unactivated or activated by preexposure to phorbol myristate acetate) increases DNA adduct formation caused by exposure to benzo[a]pyrene; addition of more activated PMN in relation to the number of lung cells further increased adduct formation in a dose-dependent manner. The authors suggest that "an inflammatory response in the lung may increase the biologically effective dose of polycyclic aromatic hydrocarbons (PAHs), and may be relevant to data interpretation and risk assessment of PAH-containing particulates." These data raise the possibility that low dose diesel exhaust exposure may result in levels of neutrophil influx which would not necessarily be detectable via histopathological examination as acute inflammation but which might be effective at amplifying any potential diesel exhaust genotoxic effect. WHO (1996) has noted that modeling of human cancer risk from rat lung tumor data should take into account the effects of both particles (carbon core) and extractable organic matter (PAHs, nitro PAHs).

Additionally, some parameters of the "particle overload" hypothesis are incompletely characterized. No data exists on the claimed inadequacy of rat lung antioxidant levels. Alveolar type II cell epithelial hyperplasia has been noted after diesel exhaust exposure, but the measures of cell proliferation used were relatively crude and unsuitable for use in a quantitative estimate of cell proliferation as would be required for biologically-based modeling. It should also be noted that uncertainties exist regarding the magnitude and biological importance of particle overload for diesel exhaust-induced rat lung carcinogenicity. Hattis and Silver (1994) examined lung burden data from diesel exhaust rat carcinogenicity studies and came to the conclusion that "there is continuing accumulation of diesel-derived dust in the lungs of rats throughout life, even at low doses". They also found that this was not predicted by models developed to represent diesel exhaust particulate matter accumulation under "overload" versus nonoverload conditions. Finally, they have found that at high diesel exhaust exposure levels, the increase in the ration of internal diesel exhaust particulate matter burden to external exposure is not very large, being slightly larger than a factor of 2 at most, and state that "Although dust overloading is a real phenomenon, it is not a very large effect and thus would not be expected to give rise to dramatically lowered estimates of risk at low exposure levels."

Finally, Section 6.1.1.1 of this document lists two studies (Pepelko and Peirano, 1983; Heinrich et al., 1986a) in which the authors describe statistically significant lung tumor induction in mice in

response to diesel exhaust exposure. Another study (Takemoto et al., 1986) is also listed in which the authors reported increased tumor incidences in mice which were not statistically significant, but which were later reported by IARC (1990) to be statistically significant after the data were reanalyzed. Section 6.1.3 has been changed to include a study by Ichinose et al. (1997) which describes lung tumor induction in mice after intratracheal exposure to diesel exhaust particulate matter. Of the six mouse diesel exhaust inhalation or intratracheal instillation carcinogenicity studies described in this document, three definitely contain positive results, one probably contains positive results, and two are negative. This indicates that the description of the mouse carcinogenicity results as "mixed" is justified, and it cannot be accurately stated that the only rodent species susceptible to diesel exhaust-induced lung tumors is the rat.

It is therefore premature to conclude that the carcinogenic response in rats to diesel exhaust is completely nonspecific, or that it is not relevant to identifying potential human cancer risk.

Comment 6: "Meta-analysis … is not an appropriate tool to use for environmental epidemiology if exposure data is unknown." The commentator cites a report of a workshop on guidelines for environmental epidemiology in support of this statement (Blair *et al.* 1995) The commentator cites this article to the effect that " 'The components of exposure such as **intensity, frequency and duration** as well as cumulative exposure should be examined.' OEHHA data is nothing more than years of service and job title. From these vague categories, exposure intensity, frequency and duration were assumed by the OEHHA. The meta-analysis is invalid." (Emphasis in original comment)

***Response**: In occupational and environmental epidemiology, detailed exposure measurements for the subjects under study are the exception rather than the rule. Therefore, in these types of studies, indirect measures or surrogates of exposure are often used. The authors of the article cited by the commentator recognize this limitation, but do not conclude that a meta-analysis using such studies is invalid. Rather, they indicate that combining studies together that have different measures of exposure can contribute to heterogeneity (i.e., substantial differences among the results of the individual studies that preclude combining them into a summary estimate of effect). More specifically, Blair et al. write (p. 194):*

> "*In environmental epidemiology, exposure measures often involve surrogates, while specific exposures are not clearly identified…[E]exposure should be specified as narrowly as possible to translate positive findings into effective risk reduction activities*…Type and length of duration of employment for occupational exposures or proximity to a source are often used as a surrogates (sic) for general population exposure levels obtained through direct measurements. The way in which these surrogates are used often varies from study to study and can limit the ability to create comparable categories for analysis across studies. Commonly used surrogates include job titles (citations), length or duration of employment (citations) and proximity from an alleged or hypothesized source (citation)…*Differences in exposure metrics (between studies or over time) can be an important source of heterogeneity*." (Italics in original)

The OEHHA meta-analysis was based on occupational epidemiological studies involving diesel exhaust in which there were no concurrent industrial hygiene measurements. Therefore, OEHHA staff had to rely on surrogate measures of exposure as assessed primarily by job titles and duration of employment. Recognizing that this would likely be a significant source of heterogeneity, OEHHA staff stratified the data based on occupational categories in one set of analyses (e.g., truck drivers, railroad workers), in order to examine the effect that this would have on the heterogeneity among the pooled estimates in these occupational subsets.

The full text of the citation adduced by the commentator reads as follows: "The assumptions behind the exposure assessment should be explained carefully. ***To the extent possible****, components of exposure, such as intensity, frequency, duration, as well as cumulative exposure should be examined" (Blair et al., op. cit., p. 195).*

Again, given the absence of concurrent exposure measurements, a meticulous analysis of these components of exposure, while certainly desirable, is not possible. OEHHA has clarified these issues regarding exposure in the methods section of the revised text of the meta-analysis.

Comment 7: OEHHA has violated the guidelines set forth by the scientific community [in the article by Blair *et al.*, noted in the prior comment for meta-analysis." In addition to the problem of exposure assessment, other design flaws articulated by the commentator include: (1) Failure to include studies of different designs; (2) Inclusion of redundant information such as more than one study conducted on the same cohort.

Response*: The guidelines described in the article by Blair et al. provide a summary of a workshop focusing on methodological desiderata for meta-analyses of environmental epidemiological studies. Some issues addressed in this article are clearly more important than others, and some remained unresolved among the workshop participants. Therefore, while these guidelines are certainly useful, they cannot be interpreted as determinative of the validity of a given meta-analysis. Having said that, OEHHA staff do not concur with the commentator's assertions that the diesel exhaust meta-analysis violates these guidelines regarding specific design issues, as follows: (1) Studies with several different designs are actually included in the meta-analysis, specifically case-control, nested case-control, prospective cohort, and historical or retrospective cohort approaches (among others). (2) Seven studies were excluded from the meta-analysis (as noted in Table D.2, in Appendix D) because they were redundant with those included, i.e., they purported to study the same populations as studies included in the meta-analysis. Had these redundant studies been included, they would have given undue weight to the mortality experience of those populations, i.e., this essentially would have resulted in "double-counting" with the inclusion of multiple pooled RR estimates for these study populations. However, there were six studies included in the meta-analysis from which multiple estimates of effect (of the risk of lung cancer from exposure to diesel exhaust) were extracted and used in the meta-analysis. In these instances, such estimates of effect were derived from different occupations with different workforces examined in the same study, in which overlapping exposures or person-time experience would be very unlikely. For example, in the cohort study reported by Boffetta et al. (1988), OEHHA staff extracted risk estimates for truck drivers, railroad workers, and heavy equipment operators, which essentially represent three different sets of exposures among three occupational*

populations described in one scientific report. These are not redundant reports on the same population.

Comment 8: "[T]he studies deemed 'excluded' from the meta-analysis are suspect. Based on author bias throughout the report, it is likely that OEHHA's decision to exclude or include came after extensive analysis in an effort to prove a point. The California Trucking Association, through freedom of information act, has requested all runs completed by OEHHA in determining study design for the meta-analysis. This is our seventh request for this data."

***Response**: The criteria for including and excluding studies for the meta-analysis were established a priori, i.e., before OEHHA staff began reading the studies and extracting data from them. These criteria are indicated on pp. D-1 and D-2 of the current version of the report. Tables D.1 and D.2 (pp. D-12 - D-14) list the studies included and excluded, respectively. Table D-2 also notes the reason(s) for excluding a given study. Contrary to the commentator's belief, OEHHA staff did not design the meta-analysis "after extensive analysis." OEHHA staff are aware of only one prior request submitted by the commentator for the alleged "runs completed by OEHHA in determining study design for the meta-analysis." No such runs were done, and to the best of OEHHA staff members' knowledge, this information was previously communicated to the commentator.*

Comment 9: "Relative risks in a study population are calculated by dividing rate of exposure among cases by prevalence of exposure among controls. The relative risk means the exposed persons are x times greater than those without cancer. The rule of thumb in textbooks for interpreting the size of risks is as follows:

Greater than 3 is a strong association
2-3 is a weak association
1-2 is a very weak association
1 is no association and
less than one is a negative association

Since the relative risk is a statistical association, it is an apparent relationship between the exposure and the disease lung cancer. The statistical association is not science and statistics cannot determine whether the association identified through epidemiology are (sic) real. OEHHA has no exposure data, no biological mechanism and a very weak association. OEHHA has not met the scientific test of labeling the diesel exhaust a toxic air contaminant."

***Response**: Relative risk (RR) is defined as "The ratio of the risk of disease or death among the exposed to the risk among the unexposed." (Last 1988). For rare diseases, the RR can be approximated by the odds ratio (OR), which is what the commentator has defined. It is not clear what the source of the commentator's guidelines for interpreting the magnitude of RRs is: most epidemiology textbooks in the possession of OEHHA staff have no such rules of thumb. However, Table 5.1 in the text by Monson (1990) presents a "Guide to Strength of Association," which resembles the commentators rule of thumb, as follows:*

Rate Ratio		*Strength of Association*
0.9 - 1.0	*1.0 - 1.2*	*None*
0.7 - 0.9	*1.2 - 1.5*	*Weak*
0.4 - 0.7	*1.5 - 3.0*	*Moderate*
0.1 - 0.4	*3.0 - 10.0*	*Strong*
<0.1	*>10.0*	*"Infinite"*

By Monson's criteria, most of the point estimates in the diesel lung cancer epidemiology studies would be considered to be weak to moderate associations. As Monson notes, "[t]he cutoff points are clearly arbitrary and subject to redefinition." (p. 88) However, neither Monson nor the other texts consulted by OEHHA staff make a categorical assertion that weak associations cannot represent causal relationships. The principal concern with weak associations is that they may be due to uncontrolled confounding. This does not mean that uncontrolled confounding cannot also affect the interpretation of studies in which large relative risks are reported; rather that a confounder that could provide the entire explanation for a strong observation should be easier to identify, measure, and control. Regardless of the magnitude of relative risk, however, a variety of biases (e.g., observation bias, selection bias) may affect the interpretation of the observed associations.

Furthermore, if the implications of the commentator's statement were true, that is, if so-called "weak" or "very weak" associations could not serve as a basis for causal inference, then much of what passes for clinical knowledge in medicine today is without epistemological foundation. Weak associations are found in many areas of medicine and, in conjunction with other aspects of causal inference, constitute the basis for a variety of clinical and public health interventions to prevent or ameliorate disease. To place the commentator's "rule of thumb" in context, the following table represents estimates of relative risk of death from cardiovascular disease due to cigarette smoking in several prospective epidemiological studies:

Study on cigarette smoking	*Estimate of RR of death from cardiovascular disease*
British doctors	*1.6*
Males in 25 states	
ages 45-64	*2.08*
ages 65-79	*1.36*
U.S. Veterans	*1.74*
Japanese study	*1.96*
Canadian veterans	*1.6*
Males in nine states	*1.70*
Swedish males	*1.7*
Swedish females	*1.3*
California occupations	*2.0*

Source: U.S. Department of Health and Human Services (1989)

According to the commentator's classification scheme, all of these reported associations between cigarette smoking and heart disease should be classified as either "weak" or "very weak". Yet active cigarette smoking is widely recognized as one of the principal causes of heart disease; several years ago, the Centers for Disease Control estimated that 156,000 deaths/year were due to heart disease caused by cigarette smoking.

Similarly, several large national reviews of the evidence have concluded that environmental tobacco smoke (ETS) exposure causes lung cancer in nonsmokers, even though most pooled RRs are in the range of about 1.2 to 1.9 (National Research Council, 1986; U.S. Department of Health and Human Services 1986; National Institutes of Health 1993).

The commentator correctly implies that statistics alone cannot serve as the basis for causal inference. Other factors that should be considered are discussed at length in section 6.2.4 of the March 1997 version of the document, and include the consistency of the findings, the possibility that findings are due to bias or chance, the evidence for exposure-response relationships, the temporal sequence and biological plausibility of the associations.

OEHHA assessed causal inference using standard criteria. These criteria included 1) the consistency of the findings; 2) the strength of the associations, 3) the possibility that the findings were due to bias, 4) the probability that the findings were due to chance, 5) evidence of exposure response relationships, 6) temporality of the associations, and 7) biological plausibility of the associations. The great majority of the epidemiological studies find an association. The small magnitude of the relative risk increases the potential for confounding. However, the number and diversity of the occupations studied, and the various analyses of sources of confounding (e.g. smoking, ETS exposure, recall bias, informational bias) do not indicate that confounding or chance accounts for the observed results. While limited exposure intensity information was available, based upon duration of exposure, there was evidence of an exposure response trend.

While biological plausibility is not required for causal inference, there is biological evidence to support the association: 1) diesel exhaust contains many mutagens, 2) diesel exhaust causes lung cancer in animal studies, 3) diesel exhaust contains many substances which occur in other complex mixtures which are respiratory carcinogens in the human, and 4) diesel exhaust contains known and probable human carcinogens.

Comment 10: OEHHA cancer numbers are inconsistent with federal EPA.
OEHHA has determined that ambient exposures to diesel exhaust at very low concentrations (2.2 micrograms/cubic meter) will cause 2,143 lung cancers per year in California. OEHHA stands alone on this calculation as this is a higher number of cancers than the entire fine particulate numbers used by EPA in developing the National Ambient Air Quality Standards.

***Response**: OEHHA has not determined that diesel exhaust will cause 2,143 lung cancers per year in California. We estimated that up to 10 to 2,000 Californians exposed to ambient levels of diesel exhaust over their lifetime could develop lung cancer per year. The 200-fold range of estimated life-time cancer risk provided by OEHHA captures much of the uncertainty involved in those estimations. The commentator's value pertains to the top of this range. The NAAQS for fine particulates were not based upon cancer. Rather, the NAAQS largely addresses the association of short-term exposures to fine particulates with increased morbidity and mortality. In setting the NAAQS, therefore, the U.S.EPA did not reach the risk assessment issues addressed by OEHHA.*

Comment 11: OEHHA has developed human risk assessment and dose response assessments by manufacturing data to meet a biased unsupported conclusion.
Garshick *et al* (1987a, 1988), a cohort and case control of U.S. Railroad Workers, was used to estimate the risk of lung cancer to the general population due to diesel exhaust exposure. OEHHA states these studies were used for their 1) "quality", 2) finding of a relationship of cancer rate to duration of exposure" 3) "availability of measurements of diesel exhaust among similar railroad workers from the 1980s in other studies" and 4) "direct information on smoking rate. were available." Although this is OEHHA's best shot, it falls far short of the mark. So much so, the author has stated the study is inappropriate for risk assessment because of the lack of dose-response data.

No exposure data was collected, these people were already deceased when the study that supplemented the exposure data was released. In addition the same professors from the Garshick study wrote the exposure study (Woskie *et al.* 1988b).

Assumptions were made by job classification as to level of exposure (none, low, medium, high) with much uncertainty. For example, in raw numbers, clerks were no exposure, but had 106 cases of lung cancer out of 1,694 deaths. Signalmen labeled as low exposure, reported 276 lung cancer cases. The highest group of lung cancers were yard brakeman at 312 only reported medium exposure. The highest exposure categories did not have expected corresponding lung cancers. How do the models take this fact into consideration?

The most interesting comparison in the human data is the lung cancer deaths against the 1,509 men who died of pregnancy complications (Eight Revision of the International Classification of Disease codes 630, 640, 660, and 670). Overall, the numbers for lung cancer deaths and pregnancy complications deaths were similar and based on OEHHA's logic, it is likely they could prove that diesel exhaust causes men to die from complications in pregnancy.

JOB	Exposure	Lung Cancer	% of Lung Cancers	Pregnancy Complications Men[1] (PCM)	% of (PCM)
Clerk	None (1)	106	6.3	104	6.9
Signalman	Low (4)	276	16.3	283	18.8
Passenger Engineer	Medium (5)	26	1.5	39	2.6
Freight Engineer	Medium (5)	179	10.6	151	10.0
Yard Engineer	Medium (5)	155	9.1	93	6.2
Yard Brakeman	High (6)	312	9.7	249	16.5
Freight Brakeman	High (6)	179	10.7	165	10.9
Conductor Brakeman (Passenger)	High (6)	25	1.5	41	2.7
Freight Conductor	High (6)	34	1.1	22	1.5
Hostler	High (6)	25	1.5	28	1.9
Electrician	High (7)	132	7.8	121	8.0
Machinist	High (7)	220	13	176	11.7
Shop Supervisor	High (7)	25	1.5	37	2.5

The cohort was 55,407 white men

The statistical analysis on the human epidemiological data excluded the shop workers, the most exposed category. There is no scientific reason for removal of this group of workers and also no evidence that workers would have heterogeneous exposures. During the workshop, Dr. Dawson stated the basis for eliminating the railroad studies was the author's conversations with railroad workers. CTA requests the underlying document relied on by OEHHA to determine that shop exposure would have heterogeneous exposures. In the opinion of companies who operate truck shops, this is a flawed assumption and should be the most consistently exposed group of workers.

***Response**: Risk assessment is one of the more complex analytical functions undertaken by government, and there is much opportunity for reasonable minds to differ with respect to the judgments involved. With regard to diesel exhaust, where the health effects literature is particularly voluminous and complex, there is much opportunity for honest disagreement.*

OEHHA has not "manufactured" data as suggested by this comment. OEHHA's analyses are based upon published findings in the scientific literature. Where OEHHA interpreted scientific studies, OEHHA used standard methods of scientific analysis. Where assumptions were required in the interpretation of experimental or epidemiological findings, OEHHA explicitly stated the assumption, the rationale underlying the assumption, and the uncertainties introduced by use of the assumption. Our work is subject to replication and verification.

With respect to the comparison of the signalmen and yard brakemen, the 3,548 signalmen comprised 25% of the unexposed study population. The 9,126 yard brakeman comprised 31% of the exposure cohort. Given the potential for multiple comparisons between subgroups and the related inequalities which result, and the reduced statistical power resulting from the smaller size of the subgroup populations, the interpretation of such subgroup comparisons is greatly limited. With respect to OEHHA's modeling of the signalmen and yard brakeman, both the signalmen and yard brakeman were included in the OEHHA analyses.

OEHHA acknowledges that there is limited exposure information available for the epidemiological studies. However, OEHHA feels that there is sufficient exposure information to support a reasonable estimation of potency. The CTA is correct in stating that the Garshick et al. (1988) cohort study did not collect personal exposure data. Such information is not often available in epidemiological studies. OEHHA therefore relied upon the Woskie (1988b) study of occupational exposures in the same industry, but at a later time, to estimate the earlier exposures. OEHHA considers this use of the data a reasonable one. OEHHA points out that both industry and EPA scientists have also looked to this same exposure data in conducting their own risk assessments for the Garshick cohort study.

With respect to the exclusion of shopworkers from the OEHHA analyses, OEHHA has provided its rationale in Section 7.3.3. paragraph 5. In a letter to the U.S.EPA, Dr. Garshick clearly stated concerns regarding uncertainties in the shopworkers' exposures. He wrote "The problem with the shop worker exposure group as defined based on the ICC 3 digit job code is that the shop workers who worked in the diesel repair shops shared job codes with workers in non-diesel shops where there was no diesel exhaust or little to no asbestos exposure. Apparent exposure as a shop worker based on the job code was then diluted with workers in the same job code but without true exposure, making it less likely to see an effect in the shop worker group. In addition, workers in the shop worker group of job codes tended to have less stable career paths and move to an exposed job category over 20 years compared to other diesel exposure categories. These comments appeared in the cohort paper (1) in regards to the failure to find an effect of apparent asbestos exposure in the shop workers." (p.5, Letter to Dr. Chao Chen, August 15, 1991). The scientific presentations and discussion at the January 1996 meeting also indicated that shop workers were heterogeneously exposed. Some shops entailed diesel exhaust exposures in their

operations, others did not. In the original Garshick et al. (1988) cohort study, the investigators reported results including and excluding the job classifications of shopworkers and hostlers. They reported that with both, shopworkers and hostlers, excluded from the analysis, the effect of diesel exhaust exposure remained significant and of comparable magnitude to the whole cohort. Similarly, in the Crump et al. (1991) analysis the presence of an effect did not seem to be affected by the inclusion or exclusion of shopworkers.

OEHHA recognizes that the limited exposure information available does contribute to the overall uncertainty of the dose response risk assessment for diesel exhaust based upon the epidemiological findings. However, the overall magnitude of the associated uncertainty is not unduly large. The greater than usual uncertainty in the exposure estimates is substantially offset by the much smaller than usual range of extrapolation from the occupational exposures of interest to the ambient levels of concern here. OEHHA provided a broad range of risk so as to fairly capture the scope of the uncertainty in these analyses.

The Garshick et al. cohort study (1988) did not find 1,509 men whose deaths were classified under deaths from complications from pregnancy. An assertion that the study did find such men and, any adverse implication that such an assertion would have for the study's reliability, is in error. The study author, Dr. Garshick, has assured us that, as men do not die of pregnancy, he recycled the pregnancy codes to keep track of other aspects of the study participants.

Comments on behalf of the California Trucking Association, letter to Mr. Peter Rooney dated September 25, 1997, from Joel D. Anderson

Comment: Based on our preliminary findings, the science used in OEHHA's current draft appropriates every lung cancer in California to ambient diesel exhaust exposure. Just a few years ago, these same lung cancers were attributed to exposure to second hand tobacco smoke.

Most important was our analysis of the human data used by OEHHA. The same study that reported 1,694 lung cancers associated with job classification found 1,509 men whose deaths were classified under complications from pregnancy. The additional case control study is of great interest to CTA, and we have not received the underlying data.

***Response**: Each year about 450 out of every million Californians die of lung cancer. Using the OEHHA range of 95% UCL unit risk factors for diesel exhaust proposed in our most recent draft, we find up to 0.3 to 63 of these lung cancer deaths may be attributed to diesel exhaust exposure. This range is quite wide. In other words, at the most, one out of seven lung cancer deaths would be attributable to ambient diesel exhaust exposure; at the lower end, 1 out of 1400 would be attributed to this exposure. OEHHA's epidemiological findings with regard to Environmental Tobacco Smoke (ETS) indicate that about 12 of these lung cancer deaths each year in California may be attributed to ETS. These analyses show that we have not appropriated every lung cancer in California to diesel exhaust exposure. The following table provides further information on this point. The calculations can be complex and it is important not to compare life-time risks to annual risks.*

Comparison of Attributed Risks to Population Experience

*Lung Cancer**	*California 1990 - 17,333 additional cases, 13,275 deaths* *597 additional cases/10^6 Californians, 453 deaths/10^6 Californians*
	California 1992 - 17,368 additional cases, 13,364 deaths *564 additional cases/10^6 Californians, 430 deaths/10^6 Californians*
	California 1994 - 16, 850 additional cases, 13,682 deaths *524 additional cases/10^6 Californians, 420 deaths/10^6 Californians*

** Based upon statistics reported by the California Department of Health Services.*

Diesel Exhaust: *22 - 4,400 potential additional cases (lifetime)/10^6 Californians at 1990 diesel exhaust population adjusted average ambient levels*
0.3 - 63 potential additional cases/year/10^6 Californians at 1990 diesel exhaust population adjusted average ambient levels

Risk Source	*Unit Risk* $(\mu g/m^3)^{-1}$	*cases/year/10^6 Californians*	*deaths/year/10^6 Californians*
human/animal geometric mean (95% UCL)	*2.00E-04*	*6*	*5*
Garshick/McClellan low concentration (95% UCL)	*2.00E-03*	*63*	*52*
Garshick/McClellan high concentration (95% UCL)	*5.00E-04*	*16*	*13*
Table 7.10 95% UCL	*1.37E-03*	*43*	*35*
Table 7.10 calculated MLE unit risk	*1.00E-03*	*31*	*26*

Deaths calculation assume 1994 case/death ratio.

ETS *360 estimated total **deaths** at pre-1992 levels*
a rough calculation of deaths and cases/year/10^6 Californians assuming 1.32 cases/death and a State population of 29 million is 16 and 12 cases and deaths, respectively, per 10^6 Californians.

Comments from ICF KAISER, letter dated July 2, 1997
from Kenny Crump to Genevieve Shiroma

Comment 1: OEHHA's Use of the Garshick *et al.* (1988) Cohort Study in Section 7.3
Despite the extensive analyses of the Garshick *et al.* data conducted by OEHHA, OEHHA's current risk assessment emphasizes an approach that does not rely upon this work, but is a straightforward adaptation of an analysis presented in the original Garshick *et al.* (1988) report (Section 7.3.2 of OEHHA document). OEHHA's risk assessment based on this analysis is quite similar to its use of the Garshick *et al. (1*988) cohort data in the 1994 document. The main difference is that the present report relies upon a Garshick *et al.* analysis that eliminated shopworkers and hostlers, whereas the 1994 OEHHA report (California OEHHA, 1994) utilized a Garshick *et al.* analysis that included the entire cohort.

Figure 1 depicts the two analyses from Garshick *et al.* (1988) relied upon by OEHHA in their 1994 draft (top graph) and in the present document (bottom graph), respectively. These analyses are Cox regressions that investigated how relative risk of lung cancer varied with total years of work in a job that involved DE exposure, compared to clerks and signalmen, who were assumed to be unexposed. Calendar year was used to define the time axis of the Cox regression and age was controlled by modeling age at the beginning of follow-up. The graph on the top (relied upon by OEHHA in its 1994 draft) is based on the complete cohort, and the graph on the bottom (relied upon in current OEHHA document) eliminated shopworkers and hostlers.

I have reproduced these analyses using the same method of analysis as Garshick *et al.* (Cox regression) and Figure 1 actually shows my analysis, which agreed very closely with the results reported in Garshick *et al. (1*988) (e.g., all relative risks agreed to two significant figures). Figure 2 shows the same analyses, but with two changes. First, whereas the Garshick *et al.* analyses counted any exposure in a year as a full year of exposure (Dr. Eric Garshick, personal communication), my analysis accounted for the reported number of months worked per year. Second, whereas Garshick *et al.* controlled for age at the start of the study (1 959), my analysis controlled for attained age. Controlling for attained age provided a better fit to the data than controlling for age in 1959 (producing a smaller deviance by 26 using all data and 18 with shopworkers and hostlers omitted, while estimating the same number of parameters in both instances). These graphs demonstrate that when age is controlled for using a method that more accurately models the underlying age pattern, the trend relied upon by OEHHA is not present. Instead, the relative risk decreases with increasing duration of exposure within the exposed group and there is no evidence of any increased risk among train riders with more than 10 years of exposure to DE.

We first demonstrated this inverse trend in a Poisson regression analysis included in our original report (Crump *et al.,* 1991). (Due to computer memory limitations we were unable at that time to perform the Cox regressions shown in Figures 1 and 2.) 1 presented this result at the 1994 OEHHA workshop, the 1996 San Francisco diesel workshop, and in four letters (Crump, 1995ab, 1996ab). Dr. Dawson (1 995) hypothesized that the inverse trend was due to an 'instability" in our analysis and proposed a modified analysis that utilized U.S. sex- and race-specific lung cancer rates to

control for age and calendar year. However, when I implemented his recommended analysis, I found that it produced the same inverse trend as my original analysis (Crump, 1995ab).

To better understand the lung cancer trends with increasing exposure to DE among exposed workers (i.e., train riders and shopworkers), I have conducted analyses similar to those presented in Figure 2, but with clerks and signalmen omitted. (These analyses were Poisson regressions that used categorical control for age and calendar year.) Figure 3 shows how the relative risk of lung cancer varies with increasing duration of exposure among all exposed workers, and among train riders (i.e., excluding shopworkers and hostlers). Both of these analyses indicate a decreasing trend in lung cancer relative risk with increasing duration of work, which is consistent with the analyses shown in Figure 2. This decreasing trend is mainly due to a strong decreasing trend among conductors and brakemen ($p = 0.01$). Similar results were obtained when these analyses were repeated with the last 4 years of follow-up omitted. In particular, the decreasing trend was statistically significant with shopworkers and the last four years of follow-up omitted. These results are consistent with the general decreasing trends I have found using other measures of exposure (Crump, 1996a).

Thus, three basically different types of analyses have been used to address the trend in lung cancer relative risk with increasing years of exposure: Poisson regression using internal controls (our original method of analysis), Poisson regression using external controls (recommended by Dr. Dawson, 1995), and Cox regression (used by Garshick *et al.).* All of these analyses show the same inverse trend in risk with increasing duration of exposure among workers exposed to DE. Despite having had this inverse trend brought to their attention on numerous occasions, OEHHA continues to rely upon the original Garshick *et al.* analysis depicted in Figure 1, while making vague and unsubstantiated charges regarding the analyses that show an inverse trend. On page 7-19 of the OEHHA draft report, OEHHA states: 'in analyses that (a) utilized a different model than was reported **by** Garshick *et al.,* (b) assumed exposure only subsequent to the start of follow-up, and (c) incorporated the full cohort and full follow-up, there was a failing phase of risk with categories of increasing cumulative exposure." With regard to point (a), as noted above the inverse trend is also found when the Garshick *et al.* model (Cox regression) is applied. With regard to point (b), the analysis is based on years of exposure since 1959 because this is the exposure measure used by Garshick *et al.,* and because the Garshick *et al.* analysis is the one relied upon by OEHHA in its risk assessment. With regard to point (c), as pointed out above the inverse trend is even stronger when the analysis is limited to train riders and the last four years of follow-up is omitted. Further down on page 7-19 OEHHA offered the following criticism of the analyses that demonstrate the inverse trend: "... numerical investigation by staff suggests that some of this difference [i.e., reason for the inverse trend] was due to effect that the models had on the (random) structure of the data in the cross-tabulation and that this difference appeared to be accentuated by the different correlation among time variables and by the small number of cancer deaths in the highest exposure category, ..." I find it very difficult to respond to such vague and unsubstantiated claims. In order that I may provide an adequate response, I request that OEHHA specifically explain (a) what the phrases " effect the models had on the (random) structure of the data" and "different correlation among time variables" actually mean, (b) what objective basis OEHHA has for making these claims, and (c) why do these claims, whatever they mean, apply to my analyses and not to any of the analyses presented by OEHHA in Appendix E. The phrase "different correlation among time variables"

sounds slightly similar to a claim made earlier by Dr. Dawson (1995a) to the effect that an "instability is suggested by high values of the correlation coefficients among parameters" stemming from estimation of "so many parameters". If this is the same claim, I would like to know if OEHHA believes that it also applies to the Poisson regression analysis using external controls (which was proposed by Dr. Dawson as a remedy for this 'problem' and which, as noted above but never publicly acknowledged by OEHHA, provided the same inverse trend) and, if so, what objective basis does OEHHA have for this claim. I would also like to know whether (and, if not, why, and based on what objective information) OEHHA believes this same criticism would apply to its analyses in Appendix E that estimate even more parameters. Similarly, I would like to know if OEHHA believes that these claims also apply to the Cox analysis that shows the inverse trends, what objective basis they have for this belief, and why their claim would apply to this analysis and not to any of the analyses conducted by OEHHA. Finally, I note that it is obvious from Figure 2 that the inverse trend is not due to "small numbers of deaths in the highest exposure category" since the inverse trend is evident even without including the highest exposure category.

***Response**: We are very appreciative of the work you have done with regard to the original reanalysis (Crump et al., 1991) and subsequent reanalyses. We have found your comments on previous drafts of our document helpful. The focus of our 1997 document was to provide a range of risks instead of identifying a single best estimate as we had suggested in 1994. Furthermore, there appears to be considerable confusion in comments we received regarding the analyses and reanalyses of the cohort data set. For this reason we thought it would be helpful to provide a straight-forward adaption of the Garshick et al. (1988) cohort data in our risk assessment. However, in consideration of this comment and those of other commenters, the current revision of the TSD does not use the published results of the Garshick et al. (1988) cohort study. Instead, Chapter 7 reports an analysis of the individual data that is very similar to that of that described in these comments, except that it uses continuous forms of the covariates, attained age and calendar year, instead of categorical forms, and it uses elapsed time of exposure to diesel exhaust instead of cumulative months of exposure. The slope obtained for relative hazard with duration of exposure is quite statistically significant. The resulting appearance of the categorical trend of risk with exposure duration is within the overlap between the trend found in the present comments and the trend which Garshick et al. reported. See Appendix F for an account of considering confidence intervals to define an overlap. See Figure 7-3 for a display of the trend found in the revised TSD.*

The passages which the commenter has described as vague and unsubstantiated have been removed from the TSD in preparing the current draft.

Comments on Analyses of Garshick *et al.* (1988) Cohort in Appendix E

<u>General Comments</u>

Appendix E contains results of a number of statistical exposure-response analyses conducted by OEHHA based upon the underlying data from the Garshick *et al.* cohort study. Although these analyses are not emphasized in the main report, OEHHA does cite them in support of the range for the 95% upper confidence limit on the unit risk.

As I noted earlier, the relative risk of lung cancer was significantly elevated among train riders in the Garshick *et al.* (1988) cohort study relative to clerks and signalmen. This can be demonstrated in a straightforward comparison of lung cancer rates in the two groups, without including any additional information on DE exposure. Given this, it is likely that dose response analyses comparing the two groups would find a significantly positive trend, even if there is no trend of increasing risk with increasing exposure. This is illustrated by the inverse trends shown in Figure 2. In fact it has been shown by Dr. Dawson (Crump, 1995a, 1996b) if one ignores the inverse trend in these figures and fit a straight line to these data, anchored at a relative risk of 1 for no exposure, one can obtain a statistically significant positive slope despite the inverse trend.

All of the analyses in Appendix E represent different ways of comparing the lung cancer mortality of train riders to that of clerks and signalmen. Since train riders have an elevated lung cancer mortality compared to clerks and signalmen, these analyses are prone to demonstrate a significant linear trend even if there is no additional evidence for a dose response. To illustrate this, I repeated one of OEHHA's analyses from Appendix E (7-stage Armitage-Doll model, 6th stage affected by DE, roof exposure, omitting shopworkers), except rather than using the reported months of work per year, I assigned months worked by exposed workers at random. This analysis also produced a statistically significant trend between lung cancer mortality and exposure to DE. Similarly, one would expect analyses of the type presented in Appendix E to show a statistically significant trend regardless of whether there was a gradient towards increasing risk with increasing exposure among workers exposed to DE.

Response*: The OEHHA results found significant linear trends for a variety of analysis methods. Some are influenced by the large unexposed group having lower cancer rates than corresponding members of the exposed group, consistent with the comment. However, it is important to note that the current fit is not anchored at a relative risk of 1. Even those slopes that proved significant only with the unexposed group included in the analysis are still relevant in view of the finding in the meta-analysis of Appendix D that there was a general elevation of lung cancer among studies of workers exposed to diesel exhaust. Furthermore the slopes obtained are in general agreement with slopes from the other analyses.*

Specific Comments

Comment 1: **Comments on the "general models" implemented by OEHHA in Appendix E:** The "general models" developed by OEHHA in Appendix E, like all of the models presented in this Appendix, represent different ways of comparing the lung cancer mortality of train riders to that of clerks and signalmen. As explained above, significant linear trends would be expected from these models as well by virtue of the fact that train riders have a higher lung cancer mortality rate than clerks and signalmen, even in the absence of a progressively increasing risk with progressively increasing exposure. However, despite these significant trends, these analyses do not present a dose response trend consistent with an effect of DE exposure.

Figure 4 shows the dose response trend in the analysis based on "roof exposure" that was preferred by OEHHA, based on national lung cancer rates and controlling for birth year ("NRATE + Byr"

from OEHHA's Table E-2). (This graph is based on OEHHA's computer output, which was kindly provided to me by Dr. Dawson.) This graph indicates that lung cancer risk peaks at an intermediate exposure and decreases as exposure continues to increase. The risks among the -highest exposed workers are no different from those of unexposed workers. This dose-response pattern was highly inconsistent with the linear trend assumed by OEHHA ($p < 0.001$). The other analyses presented OEHHA, including those based on ramp exposure, also show dose response patterns that are not indicative of an effect of DE exposure.

***Response**: For the general models, the categorical trends of risk with cumulative exposure have a rising and a falling phase, discussed in the comment and shown in the TSD. Traditional exposure-response relationships would be expected simply to rise with exposure. Although such a pattern could be a random effect over a rather small range of exposure, that does not seem very likely because the effect appears to persist across different choices of variables and data. Assuming that the data selected are not seriously flawed, what seems more likely is that the model used does not fully reflect the cancer process and that more biologically accurate models are needed, such as those in the next comment.*

The comment states that the dose-response pattern was highly inconsistent with the linear trend assumed by OEHHA and cites a p-value. This comment does not completely describe the information provided by OEHHA in Table E-2. First, the data referred to (NRATED + Byr) are a good fit to a positive, linear dose-response slope. This is shown by the p-value in column 6 which is $p < 0.001$. The p-value referred to in the comment is an additional test conducted by OEHHA. In addition to testing for the fit of a linear trend through the points, the OEHHA report also tested for a comparison fit between the categorical trend and linear trend (the categorical trend refers to connecting the points by the line segments). This second test indicates that the categorical trend provides a yet better fit to the data in many of the analyses OEHHA conducted, such as the one pointed out in the comment. Consequently, if one asks the question, is there a significant linear positive dose-response to the data? Table E-2 indicates the answer is yes. If one further asks can a non-linear model provide a better fit to the data in some cases? The answer is also yes. It is also important to point out that in the multistage model analyses, there is a significant positive dose-response, and the categorical trend does not provide a better fit.

Comment 2: Comments on the Armitage-Doll cancer models implemented by OEHHA in Appendix E. Through the courtesy of Dr. Dawson of OEHHA I obtained copies of programs used to conduct the analyses presented in Appendix E, including a program that implements the Armitage-Doll model of cancer. I discuss this model in more detail later in this report. OEHHA's program for implementing the Armitage-Doll model was clearly written and documented, which enhanced my ability to understand and use the program. Using this program I made a number of exposure-response investigations, which included models considered by OEHHA in Appendix E, as well as some models that were not included there. My investigation covered all 24 combinations of the following choices for a 7-stage Armitage-Doll model: "roof' or "ramp" exposure pattern; 5 or 10-year lag from occurrence of first malignant cell until death (The 10-year lag was used in order to be able to compare with results in Appendix E, and I am not sure that a lag of this magnitude is biologically plausible.); first or sixth stage affected by DE exposure (I did not follow OEHHA's example of including a model with the seventh stage dose-related since I do not believe that this

version of the Armitage-Doll model provides a biologically plausible exposure metric -- see below); including all the cohort data, eliminating shopworkers, or eliminating shopworkers and the last 4 years of follow-up.

The dose responses obtained in these 24 analyses are presented in Figures 5-10 at the end of this report. (Although OEHHA's analyses in Appendix E used exposures derived from assuming a fixed concentration for all exposed workers of 50 $\mu g/m^3$ DE in 1980,, the specific value of 50 represents a multiplicative factor that is irrelevant to the statistical analysis. Consequently, although I assumed the same exposures as OEHHA in my analyses, I have expressed exposure in the equivalent terms of weighted years of DE exposure.) None of these graphs indicate a progressive increase in lung cancer relative risk with increasing exposure. Rather they tend to indicate higher risks for small values of the exposure variable, with a decrease in risk as the highest exposure values. In this regard, they are qualitatively similar to the dose responses shown in Figure 2. Most of these dose responses are incompatible with the linear response assumed by OEHHA (as noted graphically by the fact that a straight line anchored at 1 for zero exposure cannot be drawn through all of the (90%) confidence bounds). Despite these trends, all but one of these models obtained a statistically significant coefficient for DE exposure, as did the models in Appendix E developed by OEHHA. However, this appears to be due solely to the overall higher incidence of lung cancer among train riders, since none of these models indicate a progressively increasing risk with progressively increasing values of the exposure variable.

Some similar graphs are presented by OEHHA in Appendix E. Two of OEHHA's graphs should be directly comparable to two of mine: Figures E-4 and E-6 in Appendix E compared to my graph with stage 6 affected by DE, excluding shopworkers and last 4 years of follow-up (Figure 10); 10-year lag with roof or ramp exposure. However, OEHHA's graphs look somewhat different from mine. These differences appear to be due to differences in how age and calendar year were controlled and how exposure category boundaries were selected. My analyses used category variables to control for age and calendar year, defined nine exposure categories by placing all clerks and signalmen in the zero exposure group, and set category boundaries so that equal numbers of cancers were included in each of the remaining eight categories when the entire cohort was included in the analysis. These eight categories were collapsed into four when creating the graphs. OEHHA, on the other hand, did not control at all for calendar year and controlled for age using a single continuous age variable. It is not clear why OEHHA chose this approach, since, as noted earlier, adequate control for these variables is very important when analyzing these data, and OEHHA exercised much more extensive control for age and calendar year in their "general models" in which they utilized four or five category variables to control for age and five category variables to control for calendar year. Moreover, I found that using categorical variables to control for age and including control for calendar year produced a highly significant improvement in fit in the Armitage Doll analyses over the method used by OEHHA ($p < 0.0001$). Consequently, it appears that OEHHA's Armitage-Doll models did not control adequately for age and calendar year. Further comments on controlling for age in the Armitage-Doll model are made below in connection with more detailed comments on the model.

The lack of a positive exposure response gradient within the DE exposure group demonstrated in Figures 1 and 2, and also in my graphs of the exposure response derived from the Armitage-Doll

model, is consistent with many other analyses I have conducted on the Garshick *et al. (1* 988) data. I presented earlier the results of 40 dose-response analyses (Crump, 1996a) that controlled for job category, and which consequently were testing for increases in lung cancer risk with increases in exposure within job groups. None of these analyses found a significant positive trend, although a number of significant negative trends were identified. I have also conducted analyses like those shown in Figures 1 and 2 (using months of work in a diesel-related job as the exposure variable), except restricting the analysis to particular groups of workers. Results of these analyses were as follows: non-significant positive trend among engineers and firers (p = 0.52), significant negative trend among conductors and brakemen (p = 0.01), non-significant trend among shopworkers (p = 0.56), borderline significant negative trend among all train riders (p = 0.08, Figure 3), and non-significant trend among all exposed workers (p = 0.46, Figure 3). Very similar results were obtained with the last 4 years of follow-up omitted: non-significant trend among engineers and firers (p = 0.66), significant negative trend among conductors and brakemen (p = 0.005), non-significant trend among shopworkers (p 0. 1 9); significant negative trend among all train riders (p = 0.04**),** and non-significant trend among all exposed workers (p = 0.50). Thus different methods of analysis' consistently show no evidence of an exposure-response gradient except for the fact that train riders had a higher lung cancer mortality than shopworkers and signalmen.

Response*: Following his assertion of a lack of a positive exposure-response gradient within the DE group, the commenter mentions four supporting points, none of which are presently sustained upon investigation. (1) In regard to Figures 1 and 2 in the comments, the TSD now includes, as pointed out in the first response above, a new but closely related analysis showing a categorical exposure-response trend that is plausible, especially in view of the large confidence interval on the relative hazard for the highest exposure group. See also Appendix F for further discussion of the comparison of Figures 1 and 2. However, in contrast to what the comment states, the trends described in Figures 1 and 2 can be fit to positive, statistically significant linear slopes. (2) Another part of the comment objects to OEHHA's implementation of the Armitage-Doll model because of the use of only a single continuous variable to control for age parametrically and the lack of control for age using a categorical variable and for calendar year. The control for external effects, as would be implemented by controlling for calendar year is a useful idea. The current analyses found that control for external effects using age-at-start-of-study gave a least as good a fit as using calendar year and gave similar trends of hazard with exposure. The effect of either of these controls is to reduce the still significant slope, compared to not using either control, and to produce a less clear upward categorical trend of hazard with the weighted exposure. The change of deviance indicated that the addition of either of these controls, using the corrected calculations, was only marginally significant. In controlling for age, the comment's claim that it is more appropriate to implement the relative risk predicted by the Armitage-Doll model and to model the background risk nonparametrically is apparently based on the argument that the deviance is reduced with this approach. But that calculation apparently has the error in it. A counter argument is (a) that the use of the pure Armitage-Doll model, controlled only for calendar year or birth cohort, would appear to be more self-consistent and (b) that too much flexibility in the model, such as in using the categorical form of age to control the background cancer rate, can cause undesirable fluctuations in the result while decreasing the deviance. So the current TSD does not go further than adding age-at-start-of-study as a control. (4) In regard to the commenter's graphs of exposure-response derived from the Armitage-Doll model, the next*

comment below correctly points out that there was a slight error in OEHHA's implementation of the model. Nevertheless, the commenter performed calculations using this erroneous implementation, so it is inappropriate to rely on those results. As the commenter points out in this connection, "the error should be corrected before applying these results quantitatively in risk assessment". (5) The commenter's earlier results of 40 dose-response analyses (Crump, 1996a) that did not find a significant positive trend were not appropriate for testing the hypothesis of an exposure-response relationship for diesel exhaust because they did not subtract the background concentration of the unexposed group when using the exposure measurements. (6) Analyses like those shown in Figures 1 and 2 that did not show an effect in particular workers may, with a more secure analysis method, show a finding similar to that in Point 1, although establishing any effect in a smaller group is generally less likely.

Comment 3: Fitting the Armitage-Doll Model to Data. The Armitage-Doll multistage model of cancer implemented by OEHHA assumes that lung cancer is initiated when a cell has passed through k stages, and the rate in one of the stages is a linear function of current DE exposure. This model does not incorporate proliferation of intermediate cells, such as occurs with the MVK two-stage clonal expansion model (Moolgavkar *et al.,* 1988). This model predicts the hazard of mortality from lung cancer to be

$$h(t) = \begin{cases} [(a_1 \ldots a_k (t-L)^{k-1} / (k-1)!] [\, 1 + (b_i/a_i)\, I(t-L; k, i)\,] & i = i,\ldots,k-1 \\ [(a_1 \ldots a_k (t-L)^{k-1} / (k-1)!] [1 + (b_i/a_i)\, c(t-L)\,] & i = k, \end{cases}$$

where $l(t;k,i) = \int_t (t-u)^{k-i-l} u^{i-i} c(u)\, du\, (k-1)\, ! / [t^{k-1}\, (k-i-1)\, (I-1)\, !\,]$, and where t indicates age, where i ($1 < i < 7$) is the stage assumed to be sensitive to DE, and L is the lag from the occurrence of the first malignant cell until death. In addition to the two references provided by OEHHA for this expression (Whittemore, 1977; Thomas, 1982), it is also derived in Crump and Howe (1984). A 7-stage model (k = 7**)** was assumed by OEHHA and a lag of either 5 or 10 years was assumed to occur between the appearance of the first malignant cell and death from the resulting tumor. It appears that there was a slight error in OEHHA's implementation of the model in that the lag is incorporated in the upper limit of the integral but not elsewhere in the hazard expression. Nevertheless, I don't expect that this error will make a large qualitative difference, and, to facilitate comparisons with OEHHA's analyses I did not correct this error before making my computer runs. However, the error should be corrected before applying the results quantitatively in a risk assessment.

Whenever the last stage in the 7-stage Armitage-Doll model is the only dose-related stage (i.e., k = 7, i = 7) this model (assuming a 10-year lag) predicts that the risk of chemical-induced mortality depends only on one's exposure exactly 10 years ago, with all other exposures presenting no risk whatsoever. This behavior, which I believe is biologically implausible, stems from a limitation in the multistage model formulation in that the time lag from cancer expression until death is assumed to be constant. OEHHA also implemented models that assume t hat the next to last stage is dose-related (i = 6). These models give most weight to exposure 10 years earlier, but also give some weight to exposures in earlier years. This seems to me to be a much more plausible assumption.

Because of the biological implausibility of models with i = 7, 1 did not investigate OEHHA's implementation of such models, but instead focused on their implementation of models with i = 6.

When fitting the Armitage-Doll model to the Garshick *et al.* (1988) data, just as when fitting other models, one must control for potential confounding effects of calendar year and age. The Armitage-Doll predicts that the background cancer hazard increases in proportion to lagged age raised to the power k-1. However, as illustrated by the differences between Figures 1 and 2, how one controls for age is very important when analyzing the Garshick *et al.* cohort, and controlling for age using a single continuous variable is generally not adequate. Changing from control of age using the method implemented in the Armitage-Doll method by OEHHA to controlling for age using category variables generally causes a large increase in the log-likelihood, sometimes accompanied by a large change in the DE exposure variable. Consequently, in order to ensure adequate control for age, it is more appropriate to implement only the relative risk predicted by the Armitage-Doll model, and to model the background risk non-parametrically. In contrast, OEHHA's implementation of the Armitage-Doll model included only a single continuous variable to control for age. Also, as noted earlier, OEHHA's implementation of the Armitage-Doll model did not control for calendar year at all. Because of these limitations, DE exposure appears to be confounded with age and calendar year in the Armitage-Doll models implemented by OEHHA.

Response: *The OEHHA staff are grateful to the commenter for pointing out the slight error in implementing the lag in the form of the model with the penultimate stage dose related. We apologize for any inconvenience that may have resulted from this error. The revised TSD corrects the error. In regard to the part of the comment stating that the assumption of a 10-year lag in the model with 7th stage the only dose related stage, predicts that the risk of chemically reduced mortality depends only on one's exposure exactly 10 years ago, the prediction intended in the TSD is for the average lag time to be 10 years. The actual lag should be considered a distribution about this average. The text has been revised to make it clear that all lag times should be considered distributions rather than single point values.*

Another part of the comment objects to OEHHA's implementation of the Armitage-Doll model because of the use of only a single continuous variable to control for age parametrically and the lack of control for age using a categorical variable and for calendar year. The control for external effects, as would be implemented by controlling for calendar year, is a useful idea. The current analyses found that control for external effects using age-at-start-of-study gave at least as good a fit as using calendar year and gave similar trends of hazard with exposure. The effect of either of these controls is to reduce the still significant slope, compared to not using either control, and to produce a less clear upward categorical trend of hazard with the weighted exposure. The change of deviance indicated that the addition of either of these controls, using the corrected calculations, was only marginally significant. In controlling for age, the comment's claim that it is more appropriate to implement the relative risk predicted by the Armitage-Doll model and to model the background risk nonparametrically is apparently based on the argument that the deviance is reduced with this approach. But that calculation apparently has the error in it. A counter argument is (1) that the use of the pure Armitage-Doll model, controlled only for calendar year or birth cohort, would appear to be more self-consistent and (2) that too much flexibility in the model, such as in using the categorical form of age to control the background

cancer rate, can cause undesirable fluctuations in the result while decreasing the deviance. So the current TSD does not go further than adding age-at-start-of-study as a control.

The model with 6th stage affected by diesel exhaust when controlled by age-at-start-of-study, even using the corrected calculation, still has a trend of hazard with exposure that displays little rise in the middle range and a decline at the highest exposure category. So it seems appropriate to consider a form of the multistage model that might produce this trend. The form we considered here used two suggestions from Cook et al. One is the equation that takes a limited susceptible population into account. The other is introducing a different lag into the background than into the diesel sensitive stage while still keeping the Armitage-Doll parametric form for background cancer rate. The background lag would account for the beginning of other exposure, such as cigarette smoke, at about age 20. The result of fitting this composite model to the data is the prediction of an overall hazard that is essentially level over midrange exposures and then declines at the highest exposure. The slope of hazard with exposure is not much changed from the simpler model. For this reason, it was not added to the report. The deviance is greatly reduced by this fully parametric form of background.

Comment 4: OEHHA's Analysis did not utilize the exposure data developed for this cohort. The exposure assessment conducted for this cohort (Woskie *et al.*, 1988a, 1988b; Hammond *et al.*, 1988) provided information on exposures of 13 groups of railroad workers. Exposures were quantified using several markers of DE exposure, including respirable particulate both adjusted and unadjusted for tobacco smoke, and amount of particulate extracted by a solvent both adjusted and unadjusted for tobacco smoke. In addition, significantly different exposures were measured on hot and cold days.

OEHHA's analysis was essentially independent of this extensive body of exposure information collected for this cohort. OEHHA's analysis assumed that clerks and signalmen were unexposed and all train riders (shopworkers were not included in OEHHA's analysis) were exposed to a constant level of 50 $\mu g/m^3$.. In their "roof" exposure pattern, this uniform concentration is assumed to increase linearly with decreasing calendar year to a three-fold higher level in 1959. However, Woskie *et al.* did not find that all train riders were exposed to the same levels. For example, they found differences between exposures of conductors/brakemen and engineers/firers and between freight engineers/firers and yard engineer/firers, which they considered to be plausible and consistent with work by other investigators.

The three-fold higher exposures in 1959 over exposures in 1980, which was assumed by OEHHA in their "roof" exposure pattern, are based on limited ad-hoc data in Woskie *et al.* (1988a) that suggests that earlier exposures to shopworkers could have been ten-fold higher. OEHHA arbitrarily reduced this ten-fold factor to three-fold for non-shopworkers.

The dose response trends and statistical p-values from OEHHA's analyses are independent of their assumption of an exposure level of 50 $\mu g/m^3$, since the only effect of a different assumption would be to change the exposure scale. Thus the dose response trends identified by OEHHA are not based on the exposure data developed for this cohort.

***Response**: In pilot calculations, the assumption of an exposure concentration of 50 mg/m^3 for all workers on trains produced substantially the same results as using the detailed measurements for each job group, and the assumption does simplify the lengthy analysis. The idea to use this single number arose from conversation with K. Hammond, one of the principal authors of the studies cited by the commenter above.*

The TSD's use of a 3-fold higher exposures in 1959 over the exposures in 1980 was a judgment made in relation to the data on the shop workers' 10-fold higher exposure, which would be expected to decline from 1959 to 1980 more steeply than in the workers on trains because of substantially improved ventilation in the shops, apparently not occurring in the trains.

The comment appears to assert that because the dose-response trends and statistical p-values from OEHHA's analysis are independent of the assumption of an exposure level of 50 mg/m^3, it follows that the dose response trends identified by OEHHA are not based on the exposure data developed for this cohort. But this does not follow. In any case, the OEHHA exposure data are clearly based on average exposure data. The comment suggests that including more job-specific concentrations would change the results. In our calculations, it did not. For this reason, in the 1997 report we chose a value of 5 µg/m^3 which was consistent with the uncertainty in the exposure estimate of the railroad workers overall.

Comment 5: Results based on the exposure data developed for this cohort. In contrast to the treatment of the exposure data by OEHHA in their current report, in our earlier report to U.S.EPA (Crump *et al.*, 1991) we did not assume that clerks and signalmen were completely unexposed to DE, but rather estimated the exposure of each job group based on the surrogates for DE exposure developed for that group. This approach was also taken in subsequent work (Crump, 1996b) and in previous analyses developed by OEHHA (Dawson, 1996a, 1996b). In our analyses of this type (Crump *et al.*, 1991; Crump, 1996a), we considered five measures of DE, RSP (respirable particulate adjusted for climate), UARP (RSP adjusted for environmental tobacco smoke), AARP (RSP adjusted for both tobacco smoke and climate), AEM (amount of RSP extractable by a solvent, adjusted for climate and tobacco smoke) and TEX (like AEM except unadjusted for tobacco smoke). When we applied these exposure measures in 80 analyses that involved both including and excluding shopworkers, both including and excluding last 4 years of follow-up, and controlling and not controlling for job category, we found that none of the 80 analyses found a significant positive association between a marker of exposure to DE and lung cancer. The majority (64/80) of the analyses obtained a negative correlation, 18 of which were statistically significant ($p < 0.05$). Sixteen of these significantly negative results were obtained in analyses that eliminated shopworkers, along with signalmen and hostlers.

We have also applied four of these exposure surrogates (AARP, UARP, AEM and TEX) in a battery of analyses similar to the 80 described above using the seven stage Armitage-Doll model. This battery of analyses examined all combinations of the following:

- four surrogate exposure measures (AARP, UARP, AEM and TEX);

- using entire cohort, eliminating shopworkers or eliminating shopworkers and last 4 years of follow-up;
- first stage or sixth stage assumed to be affected by exposure to DE;
- using estimated exposures or modifying them in accordance with the "Roof" exposure pattern used by OEHHA;
- 5 or 10-year lag.

These conditions define a total of 96 analyses. In each of these analyses, age and calendar year were controlled using category variables, and the Armitage-Doll exposure variable was categorized using ten categories with equal numbers of lung cancers in each group when the entire cohort was used for the analysis. Among these 96 analyses, six significant positive dose-response trends and six significant negative dose-response trends (two-sided p-value < 0.05) were identified. Among the 24 analyses that eliminated shopworkers and the last 4 years of follow-up (the types of analyses reported by OEHHA in Appendix E) there were two significant negative trends, but no significant positive trends. Thus, these analyses are consistent with other analyses we have reported in that they do not provide evidence of increasing risk of lung cancer in this cohort with increasing exposure to DE.

***Response**: None of the analyses cited in this comment subtracted the background particle measurements from either the exposed or unexposed groups. Woskie et al, (1988a, 1988b) reported concentration as respirable particles (RSP) corrected for environmental tobacco smoke (ETS). The RSP was adjusted by Woskie et al. to remove the fraction of cigarette smoke; the resulting concentration is referred to as ARP. Clerks and signal maintainers were classified as unexposed to diesel exhaust by Garshick et al. (1988). A level of ARP exposure was reported in the Woskie et al. studies for the unexposed clerks/signalmen. As indicated by Woskie et al. (1988a) the clerks/signalmen ARP exposure was in the background range of the national average. Thus, the ARP exposure of the exposed groups reflected diesel exhaust exposure plus the ambient background experienced by the clerks/signalmen. Furthermore, as indicated by Hammond et al. (1988a) the ARP exposures of the clerks "…almost certainly do not represent diesel exhaust from locomotives…" As pointed out by Garshick (1991) unless the clerks/signalmen group concentration is considered as nondiesel background, a cumulative exposure for the unexposed group will occur. This could result in a substantial number of unexposed individuals being intermixed with exposed individuals in the analysis. For example, assume a clerk's background exposure is misclassified as diesel exhaust exposure. Further assume that a freight conductor's combined background plus diesel exhaust exposure is considered to be all diesel exhaust. This would make the presumed diesel exhaust exposure for 20 years of work as a clerk virtually identical to 10 years of exposure as freight conductor. Thus, inclusion of the clerks/signalmen in the analysis in this manner could decrease the significance of the relative risk of the exposure groups. In Appendix D of the OEHHA report performed analyses adjusting the background concentration of the clerks/signalmen to zero. When this adjustment is made the results produce positive, statistically significant linear slopes. Without subtracting the background from the unexposed group the group cannot be considered unexposed.*

OEHHA did present results of preliminary work that followed the protocols of Crump et al. (1991) and his unpublished work through 1996, including not subtracting. The purpose was to replicate that work as closely as possible, correct the error and then check the corrected work for sensitivity to changing assumptions. The decision not to subtract background also allowed checking the hypothesis that ETS-adjusted RSP was related to lung cancer. OEHHA did not pursue that approach and instead subtracted background, as is necessary in order to appropriately test the hypothesis that diesel exhaust is related to lung cancer.

Comment 7: OEHHA suggests (page 7-19) that the negative slopes I obtained from my analyses were due to the use of an excessively coarse subdivision of time for allocating person-years. However, this is not the case. Negative associations were also observed when a finer subdivision of time is used. As I have indicated to OEHHA (Crump, 1996a), the 80 analyses dose-response described in the previous paragraph showed mainly negative correlations between DE exposure and lung cancer (64/80). None of the positive trends were statistically significant, although 18 of the negative trends were significant. All of these analyses, as well as all of those I have presented in this document, were based on subdividing time on a yearly basis.

The risk assessment calculations applied to both the case control and the cohort study involved multiplying lifetime risk of lung cancer by the relative risk computed for continuous exposure for 70 years. This is a relatively crude approach that effectively assumes that exposure at a given age can affect one's risk of cancer at an earlier age, which probably causes the method to inflate the risk somewhat. A lifetable approach to computing risk that computes the risk at each age interval based on one's previous exposure at that point in time would provide more accurate estimates.

Response*: The revised version of the TSD now uses the commenters suggestion in order to obtain more accurate estimates. We now use the lifetable approach for our calculations of relative risk.*

Comment 8: There is an error in OEHHA's use of the Garshick *et al.* (1988) data in its risk assessment (Table 7.10). Based on OEHHA's stated assumptions, in 1959 exposures should be three-fold the levels in 1980 and decrease linearly from 1959 to 1980. Further, in order to be compatible with the analysis by Garshick *et al.*, exposures during the most recent 5 years should be disregarded at each time point. Instead, OEHHA's calculations assumed that exposures had decreased to the 1980 levels by 1975, and neglected to incorporate the 5-year lag in its calculations.

Response: *The revised version of the TSD strives for consistent use of exposure patterns and in subtraction of lag. The revised TSD no longer contains the particular analysis that was in Table 7-10.*

Comment 9: There are several inconsistencies in the risk assessment approaches applied by OEHHA to the Garshick *et al.* case-control (Garshick *et al.*, 1987) and cohort (Garshick *et al.*, 1988) studies. OEHHA's adaptation of both data sets involved the background rate of lung cancer mortality, and OEHHA applied different values for this number in the two sets of calculations. OEHHA eliminated shopworkers in the risk assessment based on the cohort study, but shopworkers were included in their calculations based on the case control study. Deaths in persons

over the age of 65 were disregarded in OEHHA's use of the case control data, but included in their use of the cohort data. If OEHHA continues to rely upon both of these studies in its risk assessment, it should apply consistent assumptions and methods to each study. Moreover, in contrast to the extensive analyses conducted by OEHHA on the Garshick *et al.* (1988) cohort study, their treatment of the Garshick *et al.* (1987) case control study in their risk assessment is very cursory. It would be worthwhile also to look more carefully at the case control study.

***Response**: Based on the comment, OEHHA has worked towards reducing the inconsistencies between the analyses. The revised version of the TSD now follows the first two suggestions in the comment. That is, the TSD uses only the California background rates for calculating risk from relative hazard slope, and shop workers are now excluded from all analyses. No age exclusion is identified for the case-control result excluding shop workers. The TSD also uses all age groups in the cohort and case-control studies. It is not clear what, if any, problem might arise from using different criteria for age inclusion in the two studies, although complete consistency could potentially be useful in obtaining sharp answers to specific questions concerning comparisons. The revised version of the TSD considers the case-control system more extensively than the previous version of the TSD. Similar exposure scenarios are applied to both studies as well.*

Comment 10: In its document entitled "Public Comments and ARB/OEHHA Staff Responses to Part A and Part B of the Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant Report", OEHHA makes the following comment in numerous places (e.g., Page C-OEHHA-34): "Although Dr. Crump has corrected negative slopes that appeared in the report due to a programming error, he continues to conclude from his further analyses that 'the data in the Garshick *et al.* cohort do not show any convincing evidence of an association between risk of lung cancer and duration of exposure to diesel exhaust or quantitative measures of diesel exhaust.'" This statement may be misleading to persons who are not familiar with the context. I reported finding an error that affected a subset of earlier analyses and I provided a corrected set of analyses (Crump, 1996a). I didn't "correct negative slopes" per se, but rather I corrected all of the affected slopes. I have obtained negative slopes before the error was corrected and I continued to find negative slopes after it was corrected. In fact, after making the correction, 64 of the 80 slopes I calculated were negative and 18 of these negative slopes were statistically significant, whereas none of the 16 positive slopes were statistically significant. Thus, I continued to hold my opinion because there was nothing in my analyses that suggested that it needed to be changed.

***Response**: The commenter's suggestion of using the wording "corrected all affected slopes" does appear to be more accurate than the original statement in Part C of the 1997 TSD. In fact, even more accurately, all the slopes in Crump et al. (1991) were affected by the error. The commenter's report of the results is noted, and as above, it is important to point out that those results are based on calculations that used non-zero exposure concentrations for the unexposed group. When we made similar calculations assuming the unexposed workers were exposed to a zero concentration of diesel exhaust, the slopes were positive.*

Comments on behalf of Detroit Diesel, letter dated August 22, 1997 to Genevieve Shiroma from John Duerr

Comment 1 : The ARB/OEHHA draft report recommends that diesel exhaust be identified as a "toxic air contaminant" under the provisions of the California Health and Safety Code sections 39650-39674. If this recommendation is accepted, it will unjustifiably implicate diesel engines as a health concern in a way that will do irreparable harm to DDC, our industry, our customers who depend on diesel engines for their livelihood, and the people of California who benefit from the economic advantages that diesel engines provide. This is particularly true if the highly questionable cancer risks postulated by OEHHA remain as part of the report.

DDC opposes the draft report. Our opposition is not simply based on the economic harm that will accrue to our company and the people of California if the report is accepted, but is based on more fundamental issues.

First, we do not believe the risk estimates included in the draft report meet the regulatory requirements of being "based upon sound scientific knowledge, methods or practices" and being "consistent with current scientific data" (Sections 39661 and 39662 of the California Health and Safety Code). OEHHA has based their risk assessments on epidemiological studies in which actual diesel exposure levels are unknown and on rat studies which involved extremely high levels of exposure and mechanisms that do not occur in human beings. Several noted scientists including Dr. Eric Garschick and Dr. Joe Mauderly, the experts whose studies OEHHA relied on most heavily for their risk assessments, have stated that their studies cannot be reliably used for human risk assessment. Not only has OEHHA failed to develop quantitative risk estimates that are generally accepted in the scientific community, but it has not shown a mechanism or even established the existence of a causal link between ambient exposures to diesel exhaust and lung cancer in humans (DDC is a member of the Engine Manufacturers Association (EMA) and participated in the development of EMA's comment on the draft report. The EMA comments deal with these issues in more depth and provide more thorough rationale for the abbreviated statements made here).

Response: *OEHHA appreciates the comment's concerns regarding economic issues. OEHHA's job in the Toxic Air Contaminant listing process is to assess only the health effects of diesel exhaust.*

OEHHA acknowledges the comment's citation of the EMA submission as presenting more fully issues of concern to the commentator. OEHHA responds to many of this comment's issues in greater detail in the course of responding to the EMA comments.

OEHHA has sought to base its assessments upon sound scientific knowledge, methods or practices. With respect to the use of the animal data, the OEHHA 1997 draft document conducted a quantitative risk assessment using standard methodologies. The methods and results were therefore very similar to those presented by the USEPA in its 1994 draft and by the WHO in its 1996 document. As incorporated and discussed elsewhere in Part C, Dr. Mauderly has stated his view in commenting on our 1997 draft document that "it is not appropriate to use existing lung tumor data from rats to generate quantitative estimates of unit human lung cancer risks from

environmental exposures to diesel soot". It appears that Dr. Mauderly's views are not as broad as those indicated by the comment.

With respect to the comments of Dr. Mauderly on our use of the rat data, we shared them with the Scientific Review Panel on October 16, 1997. There was a thorough discussion of the issue of using the rat data for quantitative human risk assessment at the meeting. The sense of the panel was that rat data and calculations provide useful information and should be left in the document, however, since human epidemiologic evidence was available on which to base the human risk estimate, that the human data should be used to form the range of risks. Therefore, OEHHA now bases the range of unit risk estimates only on the epidemiological information.

With respect to the comments of Dr. Garshick, it is important to separate his comments regarding the cohort study from those of the case-control study. The comments regarding the concern for the dose-response relationship and consequently the cancer unit risk have been made regarding the cohort study. Regarding the concern that his railroad worker cohort epidemiologic data could not be reanalyzed to assign a unit risk with confidence, the draft presented a very broad range of unit risks and not a single value of unit risk to which such confidence would attach. Thus, we have taken Dr. Garshick's comments into consideration as we revised the document.

With respect to the quantitative assessment of risk based upon the occupational epidemiology studies, OEHHA's position differs from those of a number of individual scientists and authoritative agencies. However, in reaching their conclusions, they each relied to varying extents upon the results of the dose response analyses of the Garshick et al. 1988 cohort study data conducted by Dr. Crump. In those analyses, Dr. Crump did not obtain significantly positive dose response relationships between the incidence of lung cancer and diesel exhaust exposure. In its own later analyses using different methods, OEHHA obtained positive dose response relationships. This difference between some of the conclusions of OEHHA and these other authoritative bodies therefore depend in important part upon the comparative merits of the approaches of OEHHA and Dr. Crump. OEHHA, having characterized and reviewed the sources of this difference, finds its approach to be more appropriate. The differences are the subject of Appendices E and F of Part B, as well as Part C here and previously, and have been highlighted for independent review by the Scientific Review Panel.

The comment is correct in stating that OEHHA has not shown a mechanism or even established the existence of a causal link between ambient exposures to diesel exhaust and lung cancer in humans. Studies examining the association of long term <u>ambient</u> exposures to diesel exhaust on the incidence of lung cancer incidence have not been done. Therefore, OEHHA has principally relied upon the available occupational exposure studies to assess the potential cancer risk. The range of extrapolation from the occupational exposures to the ambient exposures of concern is not large. This fact adds confidence to the extrapolation of findings at occupational exposures to ambient levels of exposure. With respect to the possible mechanisms of carcinogenesis, OEHHA has reviewed them, including evidence bearing on the genotoxicity of diesel exhaust. The related evidence includes the presence of known genotoxins and carcinogens in diesel exhaust, the bioavailability of various diesel exhaust constituents, and the effects of diesel exhaust or its constituents in various in vitro and in vivo test systems for genotoxicity.

Comment 2: On an even more fundamental basis, the general objective of the draft report is inappropriate and inconsistent with the purpose and authority of the Health and Safety Code.

Section 39655 of the code defines "Toxic air contaminant" as *"an* air pollutant ... and throughout the relevant sections of the code a toxic air contaminant is referred to as *"a* substance" (emphasis added). (The code does not define "substance", however, it is defined in Webster's Ninth New Collegiate Dictionary as "matter of particular or definite chemical constitution.") We believe this language is clear and that the intent of the code is that toxic air contaminants should be specific chemical species. Until now, this principle has been followed as evidenced by the fact that since the code took effect in 1983, nearly 200 toxic air contaminants have been listed; and all are specific chemical compounds or very narrowly defined classes of chemical compounds (i.e. cadmium compounds or dibenzofurans).

In recommending that whole diesel exhaust be listed as a toxic air contaminant, the draft report proposes a significant departure from this intent and long standing precedent. Diesel exhaust is not a single substance. It is, as noted in the Executive Summary of the draft report, "a complex mixture of thousands of gases, vapors, and fine particles emitted by an internal combustion engine when it burns diesel fuel". Notably, the Executive Summary then goes on to state that, "The composition will vary depending on engine type, operating conditions, fuel, lubricating oil, and whether an emission control system is present". Thus, diesel exhaust not only is not a specific chemical substance, but rather is a complex mixture which is highly variable in its composition.

The discussion of the preceding paragraphs may appear to be related to fine semantic points which are in search of narrow technical grounds for rejecting the draft report. This is not the case. The report departs from precedent first by extending the meaning of "toxic air contaminant" to include highly complex mixtures and then by defining these mixtures not by their composition, but by their *source.* In effect, the draft report misapplies the code by recommending that a *class of sources* (diesel engines) rather than a specific substance be listed as a toxic air contaminant. As a result, the draft report has two fatal flaws. First, the "scientific" conclusions reached in the report are the result of unwarranted over-generalizations. And second, the draft report and the attendant review and approval process serve no legitimate regulatory or public purpose. The following paragraphs discuss these points more fully.

Response: *With respect to the listing of mixtures in general, past practice is consistent with the consideration of mixtures for listing. First, the great majority of the nearly 200 toxic air contaminants identified in California as Toxic Air Contaminants do not speak to the listing process here. Most of these substances were identified pursuant to Health and Safety Code Section 39656-7 which address listing as Toxic Air Contaminants those substances identified under Section 112 of the federal Clean Air Act. Second, under the Toxic Air Contaminant Program, mixtures may be found in Category II (environmental tobacco smoke, diesel exhaust). and Category III (carbon black extracts, creosotes, gasoline vapors).*

General Comments and Cover Letter, dated August 1, 1997, by Glenn F. Keller, Executive Director of the Engine Manufacturers Association, to Ms. Genevieve A. Shiroma

Comment: The Engine Manufacturers Association (EMA) hereby submits its written comments in response to the ARB/OEHHA draft report entitled, "Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant, May 1997" ("the Draft Report"). Specifically, our remarks are directed at "Part B" of the report authored by the Office of Environmental Health Hazard Assessment (OEHHA) which focuses on the health risk assessment for diesel exhaust. EMA's comments include research authored by Dr. Peter Valberg of Gradient Corporation and Dr. Tony Cox of Cox Associates - both independent experts retained by EMA to evaluate the Draft Report - and are provided as part of a constructive industry effort to ensure that the Draft Report reflects the most accurate and sound interpretation of the scientific data and reasoning that have been published to date. In that regard, it is important to note that the accompanying technical research reports have been endorsed and partially funded by the American Trucking Associations (ATA) and the National Mining Association (NMA).

EMA, ATA and NMA are most concerned that the toxic air contaminant (TAC) identification process with respect to diesel exhaust not advance to the stage of the Air Resources Board review without OEHHA first responding to the fundamental defects that we have outlined in our comments and that are inherent in the Draft Reports most basic premises. In fact, the Draft Report has manipulated, misrepresented or contradicted the findings of many of the health studies it cites, as indicated expressly by the principal authors of those studies at the July 1 workshop.

For the reasons stated herein, we believe that OEHHA should cease misrepresenting the current scientific understanding of the carcinogenic potential of diesel engine exhaust, and should accept the fact that the inherent limitations of the underlying data cannot justify the quantitative risk assessment that OEHHA has attempted to construct The potential public policy effects stemming from the Draft Report as written are such that the diesel engine industry will take all necessary measures to ensure that any final report actually utilizes and reflects "the best available scientific evidence" and "sound scientific knowledge" as mandated by California law.

***Response**: OEHHA acknowledges the EMA submissions and addresses the more detailed comments provided by Drs. Valberg, Cox and others below.*

Risk assessment is one of the more complex analytical functions undertaken by government, and there is much opportunity for reasonable minds to differ with respect to the judgments involved. With regard to diesel exhaust, where the health effects literature is particularly voluminous and complex, there is much opportunity for honest disagreement.

OEHHA has not "manipulated, misrepresented or contradicted" findings as suggested by this comment. OEHHA's analyses are based upon published findings in the scientific literature. Where OEHHA interpreted scientific studies, OEHHA used standard methods of scientific analysis. Where assumptions were required in the interpretation of experimental or epidemiological findings, OEHHA explicitly stated the assumption, the rationale underlying the

assumption, and the uncertainties introduced by use of the assumption. Our work is subject to replication and verification.

As a practical matter, where the subject matter is as complex as diesel exhaust quantitative risk assessment and the different entities' positions are so very closely nuanced at each level of analysis, it is not feasible to fully and reliably distinguish those positions in an Executive Summary. However, while necessarily abbreviated, the Executive Summary does not gloss over major differences in scientific opinion in the interpretation of the evidence. For instance, with respect to controversies in the use of the epidemiological data, the Executive Summary has a Section (1.4.3.1) titled, Differing Analyses of the Garshick et al. (1988) Cohort Data. This section is primarily devoted to describing the issues raised by Dr. Crump and refers the reader to those parts of the document treating the issue at length. The document states that "these differences in the analyses and interpretation of the Garshick cohort study underlie some of the differences amongst these authorities [U.S.EPA, HEI] in the overall quantitative assessment of the human risk based upon the epidemiological data. Furthermore, where multiple parties as here have relied upon the work of Dr. Crump to reach the same or similar conclusions, it clarifies the scope of the disagreement to give precedence and emphasis to the work of Dr. Crump. In any case, we have revisited the Executive Summary to address the issues raised by Mr. Greenbaum of the HEI.

Comments on behalf of the Engine Manufacturer's Association prepared by Timothy A. French of Neal, Gerber & Eisenberg dated August 22, 1997

Comment 1: The current Draft Report (OEHHA's second effort) falls well short of the mark. More specifically, in its second Draft Report, OEHHA has once again elected to rely primarily on Dr. Mauderly's study of rats and Dr. Garshick's study of railroad workers to construct a quantitative risk assessment. In so doing, OEHHA has ignored those researchers' otherwise clear cautions, and has taken positions deemed unjustified by every other national and international body that has considered this issue within the past several years. OEHHA has also ignored the current data indicating that a nonlinear dose-response function should be examined and utilized. Finally, OEHHA is pursuing a listing for "diesel exhaust," as opposed to any specific constituent(s) thereof, which as a practical matter (even overlooking the lack of scientific justification for such a listing) makes no sense whatsoever.

Consequently, the Draft Report must be revised substantially to reflect the significant scientific uncertainties that preclude the conclusions that OEHHA has attempted to justify. The principal uncertainties undermining the Draft Report are detailed in the attached reports from EMA's consulting experts -- Dr. Peter Valberg and Dr. Tony Cox. These expert reports, copies of which are appended hereto as Exhibits A and B, along with the other comments concerning the Draft Report that OEHHA has received from the leading researchers in this area (e.g. Drs. Mauderly, McClellan and Moolgavkar) confirm beyond any legitimate dispute that the overall findings and conclusions of the Draft Report are unsubstantiated, and by no means reflect the best available science. In fact, OEHHA has manipulated, misrepresented or contradicted the findings of many of the health studies cited in the Report, as indicated expressly by the principal authors of those studies at the July 1 workshop.

In sum, OEHHA should cease misrepresenting the current scientific understanding of the carcinogenic potential of diesel engine exhaust, and should accept the fact that the inherent limitations of the underlying data cannot justify the quantitative risk assessment that OEHHA has constructed. Since OEHHA has not done that, the Draft Report remains fundamentally flawed, as detailed below.

***Response**: OEHHA acknowledges the EMA submissions and addresses the more detailed comments provided by Drs. Valberg, Cox and others below.*

With respect to use of the rat studies in quantitative risk assessment, OEHHA disagrees with the comment that its approach would be considered unjustified by "every other national and international body." The U.S.EPA 1994 draft document used the animal cancer studies to develop a single point estimate of unit cancer risk. While the CASAC comments on the U.S.EPA work did state that "it appears these studies [the rat lung cancer studies] are not relevant for human risk assessments," the U.S.EPA has yet to respond to the CASAC recommendations. In its 1996 document, the WHO derived lung cancer risks estimates for diesel exhaust from the rat bioassay data. Unlike the U.S.EPA and WHO, OEHHA in its 1994 draft did not use the animal data to develop a single point estimate of cancer risk. In its 1997 draft, OEHHA continued to express a preference for the human data but included the animal data findings in estimating the

overall range of risks. In our 1997 draft, OEHHA, therefore, moved more closely to the WHO and U.S.EPA positions regarding use of the animal data. However, OEHHA felt that use of the animal data involved comparatively more uncertainty and therefore did not go as far in our use of the animal data as the WHO or the U.S.EPA who had relied upon the animal information to develop a preferred single point estimate of cancer risk. When Dr. Mauderly commented on our 1994 previous draft, he indicated agreement with our preference for use of the human data as affording the lesser degree of uncertainty. It would not have been clear from his comments then that he definitively opposed any use of the rat data for quantitative risk assessment. Dr. Mauderly recommended then that we exclude the rat squamous cell lung tumors from our analyses. We made those particular revisions as reflected in our 1997 draft. Since that time, the evidence favoring use of the human data has become much stronger as reflected in the character of Dr. Mauderly's more recent comments on the 1997 draft and the recent guidance we have now received from the independent Scientific Review Panel on this issue.

With respect to the use of the epidemiological data, Dr. Garshick's comments concern our use of his study, not the validity of his studies. Dr. Garshick has stated that "[I]t is not possible to use the human epidemiologic data that was reanalyzed to assign a unit risk with confidence due to the uncertainty of the exposure data." OEHHA acknowledges that, given limitations in the available exposure information, it is not possible to derive a <u>single</u> unit risk value with confidence. However, OEHHA developed a <u>range</u> of unit risk values based upon a wide range of plausible exposure patterns. Furthermore, OEHHA has now revised its analyses to include more recent information provided by the EMA with respect to diesel engine emissions and the potential magnitude of the past exposures of railroad workers.

OEHHA did differ from the U.S.EPA and the WHO in relying upon the human data to estimate a range of cancer risks. Similar to those agencies, we did not use the information to derive a single preferred point estimate of risk. Generally, however, OEHHA has been more favorable to the use of the epidemiological information than other agencies. This difference between OEHHA and these other authoritative bodies depends in important part upon the comparative merits of the approaches of OEHHA and Dr. Crump to reanalyzing the Garshick et al. (1988) cohort study. OEHHA, having characterized and reviewed the sources of the differences between its analyses and those of Dr. Crump, finds its approach to be more appropriate. The basis for this position are the subject of Appendix F of Part B, as well as Part C here and previously, and have been highlighted for independent review by the Scientific Review Panel.

As the comment points out, the U.S.EPA has stated, primarily upon the basis of the limited exposure information available and the reanalyses of Dr. Crump, that the human data are inadequate for quantitative risk assessment. However, the U.S.EPA does include the cancer risks based upon the epidemiological data in its document. Apparently, the U.S.EPA found the epidemiological data inadequate in the context of deriving a preferred single point estimate for diesel exhaust's carcinogenic potency. The U.S.EPA therefore derived a point estimate from the animal cancer data. In discussions of the OEHHA analyses with U.S.EPA staff, we have found them to be generally supportive of including epidemiological data in framing a range of risks. At the July 1, 1997 workshop Dr. Koppikar indicated that the U.S.EPA would be using epidemiological data from the Garshick et al. (1987) case-control study in the next quantitative

risk assessment it presents to the Clean Air Scientific Advisory Committee. Dr. Koppikar also stated that the resulting risk values were "pretty consistent and very similar with what Cal-EPA has presented here today." (Transcript of the Public Workshop for the Diesel Exhaust Identification Report, p.75).

Risk assessment is one of the more complex analytical functions undertaken by government, and there is much opportunity for reasonable minds to differ with respect to the judgments involved. With regard to diesel exhaust, where the health effects literature is particularly voluminous and complex, there is much opportunity for honest disagreement.

OEHHA has not "manipulated, misrepresented or contradicted" findings as suggested by this comment. OEHHA's analyses are based upon published findings in the scientific literature. Where OEHHA interpreted scientific studies, OEHHA used standard methods of scientific analysis. Where assumptions were required in the interpretation of experimental or epidemiological findings, OEHHA explicitly stated the assumption, the rationale underlying the assumption, and the uncertainties introduced by use of the assumption. Our work is subject to replication and verification.

As a practical matter, where the subject matter is as complex as diesel exhaust quantitative risk assessment and the different entities' positions are so very closely nuanced at each level of analysis, it is not feasible to fully and reliably distinguish those positions in an Executive Summary. However, while necessarily abbreviated, the Executive Summary does not gloss over major differences in scientific opinion in the interpretation of the evidence. For instance, with respect to controversies in the use of the epidemiological data, the Executive Summary has a Section (1.4.3.1) titled, Analyses of the Garshick et al. (1988) Cohort Data. This section is primarily devoted to describing the issues raised by Dr. Crump and refers the reader to those parts of the document treating the issue at length. The document states that "these differences in the analyses and interpretation of the Garshick cohort study underlie some of the differences amongst these authorities [U.S.EPA, HEI] in the overall quantitative assessment of the human risk based upon the epidemiological data. Furthermore, where multiple parties as here have relied upon the work of Dr. Crump to reach the same or similar conclusions, it clarifies the scope of the disagreement to give precedence and emphasis to the work of Dr. Crump. In any case, we have revisited the Executive Summary to address the issues raised by Mr. Greenbaum of the HEI. OEHHA has also addressed these comments in its response to Mr. Keller of the Engine Manufacturers Association who incorporated them into his comment letter.

Comment 2: The draft OEHHA report is scientifically inadequate. The ultimate conclusion of the draft OEHHA report is that ambient day-to-day exposures to "diesel exhaust" at concentrations of 2.2 $\mu g/m^3$ **w**ill cause up to approximately 2,143 lung cancer deaths per year in California [150,000 deaths ÷ 70 years]. "Sound scientific knowledge" -- the mandatory benchmark of the TAC review process -- does not support this purported conclusion. Indeed, if OEHHA were correct, it would mean that diesel exhaust, presumably diesel exhaust particulate matter, would kill approximately 16,400 people each year in the U.S. [2,143 x (260MM ÷ 34MM)]. But that is more than the total of 15,000 premature deaths that U.S. EPA has attributed to <u>all</u> particulate matter (not just diesel exhaust particulate) in this country. OEHHA's conclusions therefore belie common sense as well

as EPA's much-publicized studies concerning the health effects allegedly associated with particulate matter.

OEHHA's conclusions, then, clearly do not reflect the best available scientific evidence relating to this issue. In fact, as the following discussion demonstrates, many of the premises that OEHHA proffers as support for its conclusions are not scientifically defensible. This renders the Draft Report, as a whole, scientifically inadequate.

Response: *OEHHA's 1997 draft document did not indicate that diesel exhaust "would kill approximately 16,400 people each year in the U.S." as the comment states. Nor, had OEHHA determined that diesel exhaust would cause 2,143 lung cancers per year in California. We estimated that up to 10 to 2,000 Californians exposed to ambient levels of diesel exhaust over their lifetime might develop lung cancer per year. The 200-fold range of estimated life-time cancer risk provided by OEHHA captures much of the uncertainty involved in those estimations. The comment's value pertains to the top of this range. The NAAQS for fine particulates were not based upon cancer. Rather, the NAAQS largely addresses the association of short-term exposures to fine particulates with increased morbidity and mortality. In setting the NAAQS, therefore, the U.S.EPA did not reach the risk assessment issues addressed by OEHHA.*

Comment 3: The Animal/Rat Studies. Over the past five years, the conclusions to be drawn from the earlier inhalation studies of rats have changed dramatically. What was known in the 1980's was that lung tumors could be induced in rats if you exposed them for nearly their entire lifetimes to exceedingly high levels of concentrated diesel exhaust (2,000-10,000 $\mu g/m^3$ v. ambient concentrations of 2.2 $\mu g/m^3$). What has been learned since then is that:

- Whole diesel exhaust is not genotoxic in laboratory tests.
- To be biologically effective, the organic fraction of diesel exhaust must be extracted with strong solvents and then concentrated.
- Even if bioavailable, the total quantity of the organic fraction of diesel particulate is in all likelihood too small to have any effect.
- The organic fraction of diesel exhaust is not necessary for tumor induction in rats.

Consequently, the tumor response in rats is now believed to be initiated by a physiological response to particulate matter. But (as depicted on the chart on p. 3A) the response can be duplicated for many types of inert particles - not just diesel particulate -- and requires lung "overload." Thus, the growing scientific consensus is that the observed tumor response is a species-specific response (significant lung inflammation and cell proliferation) unique to the rat. The rat, then, has been found to be an "outlier." In addition, the data demonstrate that <u>there is a threshold</u> below which no response is triggered.

Faced with these findings, the conclusions of the leading research organizations and experts (in addition to those set forth in the appended expert reports) are most instructive:

A. <u>Health Effects Institute Special Report, Diesel Exhaust: A Critical Analysis of</u> <u>Emissions, Exposure, and Health Effects</u> (HEI April, 1995) (hereinafter, "HEI Diesel Report")

- "The lung tumors observed in rats exposed to high concentrations of diesel emissions may be due to a species-specific response to inhaled particulate matter rather than to a carcinogenic mechanism that also occurs in humans." (p.7)

- "[T]he limited data for coal miners suggest that even when particle clearance mechanisms are overwhelmed and the lungs contain heavy particle burdens, cancer does not necessarily develop." (p. 50)

- "Most of the U.S. population is exposed to relatively low, long-term average atmospheric concentrations of diesel particulate matter (1 to 10 μ/m^3), and for this population the relevance of the rat bioassay data to estimate human lung cancer risk is questionable." (p.50)

B. Correspondence from CASAC Chairman, George Wolff, to Carol M. Browner, Administrator US EPA, dated August 3, 1995

> "The cancer-causing mechanism in the rat may be unique to the rat and does not appear to occur in other species, including humans. The mechanism in rats is apparently related to particulate overload followed by a sequence of events beginning with inflammation and ending in tumorigenesis. These events are conditional upon particle overload which also occurs in rats exposed to high concentrations of inert dust as well. Consequently, it appears that these studies are -not relevant for human risk assessments."

C. Correspondence) from Joe L. Mauderly to US EPA Science Advisory Board, dated May 8-1995

> "The animal carcinogenicity data available for exposure-response modeling come from only one species, the rat, with no comparable response in two other species [mice and hamsters]. Moreover, the weight of evidence (that is growing monthly) strongly suggests that the rat is an outlier in its typical neoplastic response to chronic lung irritation. We have no strong evidence that human lungs behave similarly, and quite suggestive evidence that they do not."
>
> "Tumors occur in rats only under conditions which are not expected to occur in any substantial number of humans, and certainly not as a result of environmental exposure."
>
> "Current thinking is that the rat neoplastic response should not be used for estimating lung cancer risks from exposures two orders of magnitude, or greater, below those of the rats."

That OEHHA would ignore these concerns, especially those of Dr. Mauderly, raises very troubling questions, and provides clear evidence that OEHHA's report does not reflect sound scientific knowledge or the best available evidence. If the rat is not even predictive of other rodents, how can OEHHA be so convinced that the rat is predictive -- quantifiably predictive -- of human responses?

***Response**: The evidence over the past decade does not support the comment's conclusion that whole diesel exhaust is not genotoxic in laboratory tests. OEHHA surveyed diesel exhaust-, diesel exhaust particulate-, and diesel exhaust extract-induced genotoxicity in bacteria, yeast, Drosophila, rodents, non-human primates and humans. Further review may be found in IARC (1989), HEI (1995) and WHO (1996). Much of the information regarding genotoxicity has been obtained using diesel exhaust particles or extracts of diesel exhaust particles. Diesel exhaust particles or their extracts are mutagenic in bacteria (Salmonella typhimurium and E. coli) and in several mammalian cell systems (CHO, V79, BALB/c3T3, L5718Y mouse lymphoma, human lymphoblasts). Diesel exhaust particles or their extracts induce chromosome aberrations, aneuploidy, and sister chromatid exchange in rodent and human cells in culture. Diesel exhaust particles and their extracts are also capable of inducing cell transformation. Diesel exhaust particles or their extracts can also produce superoxide and peroxide radicals and inhibit the antioxidant enzymes responsible for radical scavenging. Exposure to diesel exhaust particulate matter can cause unscheduled DNA synthesis in vitro in mammalian cells. DNA adducts have been isolated from calf thymus DNA in vitro and mouse lung DNA following intratracheal instillation. Some information regarding genotoxicity also has been obtained directly from diesel exhaust exposures. Inhalation exposure to diesel exhaust results in DNA adduct formation in rodents and monkeys. Increased levels of human peripheral blood cell DNA adducts are associated with occupational exposure to diesel exhaust. The genotoxic effects of diesel exhaust may be involved in the initiation of pulmonary carcinogenesis in humans.*

The comment also questions the bioavailablity of the genotoxins in diesel exhaust. Several lines of evidence suggest bioavailability. First, the in vitro genotoxic activity of diesel exhaust particulates dispersed in pulmonary surfactant exhibited similar activity to particulates extracted with dichloromethane. Second, inhalation exposure of rats and monkeys to diesel exhaust results in DNA adduct formation and in vitro exposure of rat tissues to diesel exhaust induces unscheduled DNA synthesis. DNA adducts have been associated with occupational exposure to diesel exhaust. Fourth, urinary metabolites of PAHs have been found following exposure of rats to diesel exhaust. Preliminary evidence indicates the same may be true for humans. Consequently, it appears that organic chemicals adsorbed onto the particles, particularly the genotoxic components, are likely to be bioavailable in humans.

OEHHA disagrees with the comment's position that "[e]ven if bioavailable, the total quantity of the organic fraction of diesel particulate is in all likelihood too small to have any effect." Given the lack of information as to the chemical species which are predominantly responsible for the carcinogenic effect of diesel exhaust, and the lack of knowledge as to their individual carcinogenic potencies and their potential for synergistic interactions, it is not possible to state with any confidence whether or not the total quantity of the organic fraction of diesel exhaust is in all likelihood too small to have an effect.

With respect to the predictive value of the rat data for humans, OEHHA responds separately to related comments from Dr. Mauderly and others below. With respect to the related comments of Dr. Mauderly and others, we shared them with the Scientific Review Panel on October 16, 1997. There was a thorough discussion of the issue of using the rat data for quantitative human risk assessment at the meeting. The sense of the panel was that rat data and calculations provide useful information and should be left in the document, however, since human epidemiologic evidence was available on which to base the human risk estimate, that the human data should be used to form the range of risks. Therefore, OEHHA now bases the range of unit risk estimates only on the epidemiological information. However, it was also clear that the information presented to justify the existence of a threshold response in the rat data were not considered convincing by the Panel.

Comment 4: The Epidemiologic Studies. The data from the relevant epidemiologic studies are no better. The results of the epidemiologic studies are generally consistent in showing a "weak association" (HEI Diesel Report, p. 6) - <u>not causation</u> - between exposure to diesel exhaust in occupational settings and lung cancer. But the increase in relative risk (1.2 to 1.5) was small <u>(see</u> summary chart on p. 4A) and many of the measurements involved were imprecise. As a result, many of the studies are not statistically significant.

In addition, and as detailed in the appended expert reports, many of the studies that OEHHA relies on did not control adequately for confounding factors such as smoking, environmental tobacco smoke, nondiesel particulate matter, asbestos exposure, socioeconomic factors, diet, or exposures to other air pollutants. Even more significantly, the key epidemiologic studies of Garshick, *et al*. lack <u>any</u> actual exposure data; we do not know the actual exposure levels involved or how the composition of diesel exhaust then at issue compares to today's exhaust. Thus, as HEI has noted,

"The absence of exposure measurements in the study populations is the main methodologic problem limiting interpretation of the epidemiologic data and its use in quantitative risk assessments." (HEI Diesel Report, p. 28.)

While past epidemiologic studies are fundamentally flawed, it is also important to bear in mind that the composition of diesel fuels and the combustion process that creates exhaust have changed dramatically since the 1960's. These changes (as depicted on the following chart, p. 5A) include:

- reductions in particulate emissions
- HC + CO emissions reduced to 10% of current standards
- 75% reductions in NOx emissions
- development and use of low sulfur/low aromatics fuels
- development and implementation of advanced diesel engine designs
    - high pressure fuel injection
    - computerized timing
    - turbocharging and charge air cooling
    - improved oil control
    - reshaped combustion chambers

Consequently, what came out of one particular type of diesel engine - locomotives - in the 60's and 70's is simply not the equivalent, either in quantity or composition, of diesel engine emissions today. Given these profound shortcomings, the leading experts continue to caution against relying on the existing epidemiologic data to make any specific or quantitative conclusions:

A. HEI Diesel Report

- "The results of most of the studies were not statistically significant." (p.27)

- "[T]he lack of definitive exposure data for the occupationally exposed study populations precludes using the available epidemiologic data to develop quantitative estimates of cancer risk." (p.8)

B. US EPA: Health Assessment Document for Diesel Emissions (September 1994)

- "Human data are preferable for developing risk estimates. However, ... use of the human data are, considered in this case to be inadequate for this purpose. First of all, the relative risk ratios for the human epidemiology studies are generally only slighter greater than one. Small errors in the adjustment for possible confounding factors, especially smoking, could result in a large percentage change in relative risk."

- "Finally, an attempt was made to use the Garshick, *et al.* (1987) railroad worker study to develop a unit risk assessment. However, attempts to relate increasing duration or intensity of exposure to increasing response rates were unsuccessful." (p.12-6)

C. World Health Organization: IPCS, Environmental Health Criteria #171, Diesel Fuel and Exhaust Emissions (1996) (hereinafter, 'WHO Diesel Report")

- "[H]istorical measurements of exposure to diesel exhaust are unreliable and exist only for current workers in two industries. A quantitative risk assessment cannot be conducted on the basis of epidemiological data in which job title was used as a surrogate of exposure .... Consequently, there are no human data suitable for estimating unit risk." (p. 254)

Thus, epidemiologic studies are not recommended for risk assessment by EPA, WHO or HEI. Even, Dr. Eric Garshick -- the researcher upon whom OEHHA principally stakes its claim -- has, confirmed to CASAC (see Exhibit C hereto) that his studies do not constitute a sound scientific basis for any quantitative risk assessment.

Correspondence from Eric Garshick to US EPA Science Advisory Board, dated May 30, 1995

> "We agree with EPA that a major limitation in the use of this data set and others to conduct a risk assessment is the crudeness of the exposure data and the inability to determine how significantly exposures changed (decreased) over time."
>
> "[I]t is not possible to use the human epidemiologic data that was reanalyzed to assign a unit risk with confidence due to the uncertainty of the exposure data."

More importantly, subsequent analyses of the Garshick data have established that the railroad workers study does not confirm a positive dose-response relationship between increased exposure to diesel exhaust and an increased relative risk for lung cancer. See K. Crump, Statistical Exposure-Response Analysis of a Retrospective Cohort Study of Lung Cancer Mortality in U.S. Railroad Workers Exposed to Diesel Exhaust, Journal of Occupational and Environmental Medicine (submitted 1997). While OEHHA attempts to justify its wholesale rejection of Dr. Crump's analyses on the basis of a supposed cross-tabulation "error" (see Correspondence from Peter D. Venturini to Glenn Keller, 6/13/97), OEHHA knows full well that the purported "error" has been addressed and that the results of Dr. Crump's analyses still hold true. Thus, as Dr. Crump demonstrated at the July 1 workshop: (i) lung cancer mortality was not significantly elevated among railroad shopworkers in comparison to clerks and signalmen, despite the fact that shopworkers likely had the most intense exposures of any group; (ii) the relative risk of lung cancer tended to decrease with increasing duration of exposure within exposed railroad workers; and (iii) there is no convincing evidence for an effect of diesel exhaust exposure upon lung cancer in the railroad workers cohort.

In response to these otherwise clear limitations (limitations which Dr. Garshick himself has recognized), and as detailed in the Valberg Report, OEHHA has made completely arbitrary adjustments to its reanalyses of the Garshick *et al.* data in a transparent effort to manufacture a dose-response relationship. This manipulation of the data includes: (i) requantification of historic exposure levels to minimize differences in estimated exposures experienced by shopworkers, on the one hand, and all other exposed workers, on the other hand; (ii) outright exclusion of the shopworkers from the reanalyses, even though they were the workers estimated to have received the largest exposures to diesel exhaust; and (iii) rejection of analyses of the estimated exposures in terms of total lung-deposited amounts of diesel exhaust, apparently because this metric offered the most dramatic illustrations of the absence of any increasing dose-response relationship. OEHHA does not and cannot adequately justify any of these post hoc "adjustments" of the data.

In sum, and as further explained in the attached expert reports, OEHHA's risk assessment pretends that very little is known about the specific causal mechanism of cancer induction in rats, while much is known about diesel exhaust epidemiology. Neither position is accurate. In fact, as evidenced by Dr. Mauderly's comments, a great deal is known about rat lung carcinogenesis in response to over-burdening by diesel exhaust, but this knowledge has not been utilized properly in OEHHA's risk models or estimates. On the other hand, and despite OEHHA's "likely explanations" to the contrary, the epidemiological evidence has failed to establish a causal link between diesel exhaust and development of human lung cancers, and OEHHA's causal interpretation of a relation between diesel exhaust and human lung cancer is unsupported by any statistical tests for causation. In fact,

OEHHA did not even run any causality tests. Not surprisingly, then, OEHHA's purported conclusions fly in the face of the principal finding of the Health Effects Institute:

> The average levels of diesel exhaust found in most occupational settings, which are below 1 00 $\mu g/m^3$, would not likely be a cancer hazard for these workers, nor would ambient levels (1 to 10 $\mu g/m^3$) present a cancer risk for the general population...[While] one cannot exclude the possibility that [there is] a mechanism involving direct action between the chemical mutagens in diesel exhaust and DNA...the available epidemiologic and animal data are insufficient to support this hypothesis or to be used in quantitative risk assessments. (HEI Diesel Report, pp. 1-2.)

OEHHA's Draft Report is therefore scientifically inadequate, based as it is on studies that even their authors claim are wholly unsuited to quantitative risk assessments. Indeed, the epidemiologic studies in question were not designed for risk-assessment purposes. OEHHA therefore should accept those limitations and acknowledge that it is not possible at this time to quantify a hypothetical risk associated with exposure to diesel exhaust.

***Response**: The comment reiterates many of the comments submitted by other persons on behalf of the Engine Manufacturer's Association. OEHHA also separately responds to those more detailed comments in Part C.*

With respect the qualitative relationship between lung cancer and diesel exhaust exposure, the comment's quotation of the HEI "The results of most of the studies were not statistically significant." (p.27) does not adequately convey the meaning of the HEI. A fuller quote would be "However, as illustrated in Figures 7 and 8, the increase in relative risk of lung cancer was generally small and many of the measurements were imprecise; thus, most of the studies were not statistically significant. However, the results of the more robustly designed studies were statistically significant, which increases confidence in interpreting the positive lung cancer data in occupational cohorts. Moreover, in some studies, the largest relative risks were seen in categories expected to have the greatest cumulative exposure to diesel exhaust..." Garshick et al. 1987, 1988.

The comment is correct in stating that the HEI found evidence for a weak association. Again, however, the comment does not fully capture HEI's position with respect to causation. With respect to this issue, the HEI Paper Summary (p.284) states: "Epidemiologic data are strongest for lung cancer. More studies have been conducted for lung cancer than for other diseases, and more detailed information has been collected on exposure, particularly in several recent studies of specific occupational groups. The available evidence suggests that occupational exposure to diesel exhaust from diverse sources increases the rate of lung cancer by 20% to 40% in exposed workers generally and to a greater extent among workers with prolonged or intense exposure. These results are not readily explicable by confounding due to cigarette smoking or other known sources of bias. In a general sense, the elevated rates of lung cancer among exposed workers are consistent with experimental data reviewed elsewhere in this report. These data indicate that whole diesel exhaust and selected constituents are mutagenic and produce cancer in several rodent species."

With respect to causation, the HEI Executive Summary states (p. 6), "The epidemiologic data are consistent in showing weak associations between exposure to diesel exhaust and lung cancer. The available evidence suggests that long-term exposure to diesel exhaust in a variety of circumstances is associated with a 1.2 to 1.5-fold increase in the relative risk of lung cancer compared with workers classified as unexposed.

Despite the concern that confounding by cigarette smoke might explain the observed risk elevations, most studies that controlled for smoking found that the association between increased risk of lung cancer and exposure to diesel emissions persisted after such controls were applied, although in some cases, the excess risk was lower, Only a few epidemiologic studies considered other potential confounders such as non-diesel particles, environmental tobacco smoke, asbestos exposure, diet, and socioeconomic factors. At present, there is insufficient evidence to conclude whether confounding by these factors influenced the results."

With respect to causation, the 1994 U.S.EPA draft document similarly states '"Collectively, the epidemiology studies show evidence of an association between inhalation of diesel exhaust and lung cancer in humans. Although, the evidence for carcinogenicity in humans was in most cases positive, it is judged to be limited according to the EPA's weight of evidence guidelines, because the observed increases in risk were quite low and the influence of confounding factors could not be completely accounted for." While the WHO in 1996 found "The relative risks for lung cancer as a result of exposure to diesel exhaust are generally low, and risks of this magnitude are more susceptible to chance and to the effects of unmeasured confounding factors and imprecision for adjusting for known confounding factors. As discussed above, the elevated risk for lung cancer in the four most informative studies is unlikely to be due to confounding by cigarette smoking and is probably due to exposure to diesel exhaust. Other studies, although limited primarily by the exposure ascertainment, support this assessment."

As noted by the comment, with respect to the quantitative assessment of risk based upon the occupational epidemiology studies, OEHHA's position differs from those of a number of individual scientists, including the current position of Dr. McClellan, and also from the previous positions of several authoritative bodies such as the HEI and the WHO generally have found that the inadequacy of exposure information limited the ability to conduct a quantitative risk assessment. However, in reaching their conclusions, they each relied to varying extents upon the results of the dose response analyses of the Garshick et al. 1988 cohort study data conducted by Dr. Crump. In those analyses, Dr. Crump did not obtain significantly positive dose response relationships between the incidence of lung cancer and diesel exhaust exposure. In its own later analyses using different methods, OEHHA obtained positive dose response relationships. This difference between some of the conclusions of OEHHA and these other authoritative bodies therefore depend in important part upon the comparative merits of the approaches of OEHHA and Dr. Crump. OEHHA, having characterized and reviewed the sources of this difference, finds its approach to be more appropriate. The differences are the subject of Appendix E and F of Part B, as well as Part C here and previously, and have been highlighted for independent review by the Scientific Review Panel.

As the comment points out, the U.S.EPA, HEI and others have stated, primarily upon the basis of the limited exposure information available and the reanalyses of Dr. Crump, that the human data are inadequate for quantitative risk assessment. However, the U.S.EPA does include the cancer risks based upon the epidemiological data in its document. Apparently, the U.S.EPA found the epidemiological data inadequate in the context of deriving a preferred single point estimate for diesel exhaust's carcinogenic potency. The U.S.EPA therefore derived a point estimate from the animal cancer data. OEHHA assumed a range of reasonable exposure conditions in order to derive a range of risks from the human data using a variety of dose response models. OEHHA did not develop a single point estimate. OEHHA feels the presentation of a range of risks is a reasonable approach which adequately addresses the uncertainty in the exposure information.

OEHHA has not, as alleged by the comment, made "completely arbitrary adjustments to its reanalyses of the Garshick et al. 1988 data in a transparent effort to manufacture a dose-response relationship". With respect to dose response assessment, elimination of the shop worker data does not, in fact, substantially alter the conclusions of the OEHHA dose response analysis. Any conclusion that the "shopworker data were excluded because the data do not support the thesis posed by the assessment, that diesel exhaust is carcinogenic." is therefore without basis. With respect to the reconstruction of the past exposures of interest, OEHHA has principally modeled past exposures in three different ways in order to characterize the uncertainty due to the limited exposure information available. One of these exposure patterns, the "ramp" pattern, was developed by Dr. Crump. Another, the "block" pattern, was previously used by Dr. Crump and Dr. Garshick. The third, the "roof" pattern was developed by OEHHA and the rationale for its development is described in chapter 7. The "roof" pattern developed by OEHHA estimated that worker exposures were greater in the past. This pattern therefore gave substantially lower estimates of the cancer risk than the other two exposure patterns . Any allegation, therefore, that the OEHHA roof pattern represents a "post hoc' adjustments of the data to exaggerate risk is also without basis. The EMA has subsequently provided OEHHA with information (Memo dated January 23, 1998 from Glenn Keller to George Alexeeff Re: Diesel Exposure Factor for Railroad Study Analysis) that suggests both that OEHHA's roof pattern is preferable to the ramp pattern of exposure and that the roof peak was too low. The EMA information indicates that railroad engines produced 6.8 to 11.5 times higher smoke an/or particulates than in 1974. OEHHA has incorporated this newly provided information by adding analyses which increase the peak of the roof pattern from 3-fold to 10-fold above the measured 1982 baseline values.

Contrary to any inference from the comment that OEHHA did not conduct a "statistical test" for causality, statistical tests reveal associations not causation. There is no per se statistical test for causation. However, OEHHA did assess the evidence for causality (Section 6.2.4 Causal Inference for Diesel Exhaust Exposure and Lung Cancer) using standard criteria. These criteria included 1) the consistency of the findings; 2) the strength of the associations, 3) the possibility that the findings were due to bias, 4) the probability that the findings were due to chance, 5) evidence of exposure response relationships, 6) temporality of the associations, and 7) biological plausibility of the associations. The great majority of the epidemiological studies find an association. The small magnitude of the relative risk increases the potential for confounding. However, the number and diversity of the occupations studied, and the various analyses of sources of confounding (e.g. smoking, ETS exposure, recall bias, informational bias) do not indicate that

confounding or chance accounts for the observed results. While limited exposure intensity information was available, based upon duration of exposure, there was evidence of an exposure response trend. While biological plausibility is not required for causal inference, there is biological evidence to support the association: 1) diesel exhaust contains many mutagens, 2) diesel exhaust causes lung cancer in animal studies, 3) diesel exhaust contains many substances which occur in other complex mixtures which are respiratory carcinogens in the human, and 4) diesel exhaust contains known and probable human carcinogens.

The comment quotes the HEI .."The average levels of diesel exhaust found in most occupational settings, which are below 100 mg/m^3, would not likely be a cancer hazard for these workers, nor would ambient levels (1 to 10 mg/m^3) present a cancer risk for the general population." The HEI text preceding this statement indicates that the statement is conditional on the validity of certain major assumptions: "This suggests that there may be a threshold for particle-induced tumorigenesis because inflammation and cell proliferation are thought to have important roles in the development of rat lung tumors. If so, and if the mechanism of rat lung carcinogenesis is relevant to humans, *[emphasis added], then the levels of diesel exhaust particulate matter found in some occupations (greater than 1,000 µ/m3) might be a cancer hazard for the relatively small numbers of workers exposed to these levels, and there might some reason for concern for those exposed to one order of magnitude lower. The average levels of diesel exhaust found in most occupational settings, which are below 100 mg/m^3, would not likely be a cancer hazard for these workers, nor would ambient levels (1 to 10 mg/m^3) present a cancer risk for the general population."*

OEHHA's presentation of a range of risks reflects the limitations in the available information. The meaning of the comment statement that "the epidemiological studies were not designed for risk assessment" is unclear. The studies were designed to examine the association between diesel exhaust exposures and lung cancer so as to help to identify any possible lung cancer hazard. While the studies lack exposure information and therefore do not by themselves support a quantitative risk assessment, OEHHA has relied upon other published reports of exposures in the relevant industries to estimate a reasonable range of exposures so as to derive a range of cancer unit risks.

Comment 5: Non-Linear Dose-Response. The Valberg and Cox Reports amply describe the scientific inadequacies in OEHHA's efforts to justify a linear dose-response relationship between exposure to diesel exhaust and lung cancer. Briefly, since OEHHA is attempting to estimate risks purportedly associated with diesel exhaust exposures much lower than those at issue in either the occupational epidemiologic studies or in the rat bioassay studies, OEHHA must establish a linear relationship to extrapolate down to low ambient doses (i.e. 2.2 µg/m^3 v. 2000 µg/m^3). But OEHHA cannot support its assertion of a linear relationship, which is typically associated with genotoxic carcinogens.

First, OEHHA cannot establish that the absorbed hydrocarbons (approximately 0.0007%, by weight, of "diesel exhaust"; see WHO Diesel Report, p. 101) are bioavailable or bioactive. Indeed, if human lungs are not under "overload" conditions as a result of diesel exhaust exposures (as OEHHA suggests), and macrophages are not impaired in their ability to take and remove particles,

and organic material is not released from the particles by lung surface fluids, then it is difficult to imagine how lung epithelial cells are at risk of exposure to mutagenic organic compounds. Second, experimental data demonstrate a threshold for responses that are mechanistically related to particle-induced tumorigenesis in the rat model. Even Dr. Garshick's analysis suggests that virtually all of the elevated lung cancer risk is associated with occupational exposures exceeding 20 years. And, as noted in the Cox Report, OEHHA's own analysis of the Garshick *et al.* data indicates that a threshold model is much more plausible than a linear low-dose model. More specifically, Figure 7-3 of the Draft Report shows that relative risks do not increase for the three lowest cumulative exposures, but increase dramatically for the fourth. This data pattern is fully consistent with and supportive of a threshold model. Finally, a mechanism of action for the proposed carcinogenicity of diesel exhaust in humans has not been identified. In rats, as noted above, it appears that the diesel exhaust-induced tumorigenesis is mediated by non-genotoxic mechanisms that exhibit a threshold.

The evidence for genotoxic mechanisms for diesel exhaust is thus entirely speculative and cannot be used to justify a linear dose-response model. This is more than a little significant inasmuch as non-linearity of the dose-response relationship would negate OEHHA's attempts to construct a quantitative risk estimate for low exposures resulting from ambient air.

***Response**: First, OEHHA is not attempting to extrapolate risks to much lower levels. The range of extrapolation from the occupational studies to the environmental exposures is modest, from 7 to 30-fold, depending on the measure. Second, with respect to the interpretation of Figure 7.3, it does not provide evidence for a threshold in response. The three lowest cumulative exposures cited in Figure 7-3 are statistically increased above background. This pattern is not consistent with a threshold. In any case, as the exposed workers had similar exposure levels, the cumulative exposure measure largely reflects the duration of exposure, not the intensity of the exposure. Therefore, Figure 7.3 can not provide evidence of an exposure threshold. Third, the evidence for the genotoxicity of diesel exhaust is not speculative. It is based upon in vitro and in vivo animal and human studies. Studies show the formation of DNA adducts in workers exposed to diesel exhaust. Fourth, OEHHA agrees that the mechanism of diesel exhaust lung carcinogenesis in humans, however, is not yet established.*

Comment 6: Listing "Diesel Exhaust" Is Nonsensical. In addition to being flawed in its details, the Draft Report is fundamentally unsound in its general objective. The Draft Report seeks a TAC listing for "diesel exhaust," not any particular component of diesel exhaust. But diesel exhaust, by weight, is 75.2% nitrogen, 15.0% oxygen, 7.1% carbon dioxide, and 2.6% water vapor. (See WHO Diesel Report, p. 101.) Thus, it is beyond dispute that 99.9% of "diesel exhaust" is not toxic. A listing for "diesel exhaust" as a whole therefore makes no sense.

Moreover, it remains the case that for as long as there are diesel engines in operation, there will be "diesel exhaust." Even if diesel technology were to advance (which it may) to the point where diesel engine emissions consist solely of nitrogen, oxygen, carbon dioxide and water vapor, those entirely non-toxic emissions would still constitute "diesel exhaust," and so would still be a TAC under the rubric espoused by ARB/OEHHA in the Draft Report. The proposal to list "diesel exhaust" is therefore nonsensical for this reason as well.

Finally, how are regulators or engine manufacturers supposed to respond to a TAC listing for "diesel exhaust?" Such a listing is simply far too broad. The constituents of "diesel exhaust" vary depending upon the engine, the fuel type and the operating conditions at issue. Consequently, "diesel exhaust" is in many respects a continuously evolving and complex mixture of substances, many of which (e.g. particulate matter) are already subject to stringent and effective emission control programs. Consequently, any effective regulatory program must be -- and is -- based upon the identification of one or more specific constituents of "diesel exhaust," not the whole mixture. ARB itself has recognized this necessary regulatory strategy through its negotiation of the Statement of Principles (SOP) that will govern the control of emissions from diesel engines through the year 2004 and beyond. In this case, then, if no specific constituent of diesel exhaust is identified as the supposed toxic agent, what emission constituents should manufacturers and regulators endeavor to reduce beyond the SOP standards? Or is the object of the pending TAC proposal simply the elimination of diesel engines altogether? If not, then the Draft Report must be revised to list the supposed agent that is allegedly responsible for the health effects that OEHHA has hypothesized. In the absence of such a specific listing, OEHHA's proposal will be fundamentally unsound, contrary to existing emission control strategies and, in effect, nonsensical.

For all of the foregoing reasons, and as further explicated in the attached expert reports, the Draft Report is scientifically inadequate. OEHHA therefore should heed the cautions of the leading researchers and concede that, at this juncture, the many significant scientific uncertainties preclude the development of any valid quantitative risk assessment.

Response: *The ARB is to respond to this inquiry.*

Comment 7: The Draft OEHHA Report Is Inaccurate If Not Misleading. In addition to being scientifically inadequate, the Draft OEHHA Report is inaccurate if not affirmatively misleading. EMA will present just a sampling of OEHHA's misrepresentations of the relevant data.

Sample #1 -- At pages ES-13 and 1-5 of its report, OEHHA claims that "HEI found that the epidemiological data are consistent in showing associations between exposures to diesel exhaust and lung cancer."

What OEHHA conveniently fails to note, however, is that HEI actually concluded that "the lack of definitive exposure data for the occupationally exposed study populations precludes using the available epidemiologic data to develop quantitative estimates of cancer risk." (HEI Diesel Report, p. 8.) Misrepresenting HEI's conclusions underscores the lack of sound scientific reasoning in the Report.

Sample #2 -- At page 1-7 of its report, OEHHA discusses risk assessments based on occupational studies and claims that "U.S. EPA cited these same...values as being practical in assessing human risks involving exposures in the range of study observations."

What OEHHA omits from its Report is that in its 1994 report U.S. EPA unambiguously concluded that the human data are inadequate for quantitative risk assessments. Once again,

misrepresentation is not sound science. OEHHA also fails to mention the recent conclusion of WHO/IPCS. That body also stated in unequivocal terms that: -

A quantitative risk assessment cannot be conducted on the basis of epidemiological data in which job title was used as a surrogate of exposure .... Consequently, there are no human data suitable for estimating unit risk. (WHO Diesel Report, p. 254.)

Sample #3 -- On page 2-1 of its report, OEHHA touts the two-day workshop convened in San Francisco in January, 1996 to discuss the use of epidemiologic data for quantitative risk assessments, and asserts that the report "has been updated and revised to reflect the benefit of those discussions."

What the report utterly fails to note, however, is that the one clear conclusion from the San Francisco workshop was that the critical issue of whether Dr. Garshick's studies actually demonstrate a dose-response trend had to be resolved. Indeed, on January 30, 1996, George Alexeeff publicly declared that the Crump/Dawson debate "must be settled." But that debate has not been settled, and OEHHA has refused repeated requests to participate in a forum of the leading bio-statisticians to resolve the debate one way or the other. This clear failure by OEHHA to implement the one overriding consensus of the 1996 workshop should be duly noted in the body of the Report, not consigned to inaccurate footnote discussion in Appendices E and F.

Sample #4 - OEHHA repeatedly stresses the significance of its "meta-analysis" of 31 epidemiological studies.

What OEHHA does not note in the body of its report, however, is that it excluded 16 of the 47 relevant studies from its meta-analysis. At the very least, OEHHA should discuss fully the impact of this potential selection bias as well as the attendant publication bias. OEHHA also deliberately fails to mention that the two leading meta-analyses of the relevant studies--the meta-analyses conducted by Drs. Stober and Abel in 1996, and by Drs. Muscat and Wynder in 1995--each concluded that the epidemiological evidence is insufficient to establish diesel engine exhaust as a human lung carcinogen. (See Valberg Report, p. 23.). OEHHA's apparent effort simply to ignore contrary findings is once again not sound science and instead demonstrates bias in this case.

Sample #5 -- OEHHA asserts at pages 1-9 and 6-50 of its report that the relevant epidemiologic studies support a "causal association" between diesel exhaust exposure and human lung cancer.

This is not so. No epidemiological study has ever been published that establishes a causal link between diesel exhaust exposure and human lung cancer. Only studies of statistical association have been undertaken thus far, and these studies are inconclusive since they have not been designed or analyzed to preclude false positives due to multiple comparisons and modeling errors. This very immature state of the scientific knowledge is a critical factor for any meta-analysis to uncover and explain. Yet it passes without mention in OEHHA's highly selective discussion.

***Response**: Risk assessment is one of the more complex analytical functions undertaken by government, and there is much opportunity for reasonable minds to differ with respect to the*

judgments involved. With regard to diesel exhaust, where the health effects literature is particularly voluminous and complex, there is much opportunity for honest disagreement.

The OEHHA document is neither inaccurate nor misleading as suggested by this comment. OEHHA's analyses are based upon published findings in the scientific literature. Where OEHHA interpreted scientific studies, OEHHA used standard methods of scientific analysis. Where assumptions were required in the interpretation of experimental or epidemiological findings, OEHHA explicitly stated the assumption, the rationale underlying the assumption, and the uncertainties introduced by use of the assumption. Our work is subject to replication and verification.

With particular respect to sample concern #1, our statement that the "HEI found that the epidemiological data are consistent in showing associations between exposures to diesel exhaust and lung cancer." is found in the part of the Executive Summary (Section 1.3, Carcinogenicity) which addresses hazard identification. With respect to the comment's concerns regarding any selective citation of the HEI findings, the allegedly 'missing' quote can be found in an appropriate place, Executive Summary Section 1.5 which addresses sources of uncertainty in the dose response assessment. This section states that "The HEI (1994) cited the general lack of exposure information as limiting the ability to perform a reliable quantitative risk assessment using any of the human data."

The apparent contradictions described in sample concern #2 relate to the complexity of the subject matter at hand and are reconcilable within the context of the U.S.EPA and OEHHA documents. As the comment points out, the U.S.EPA document does state that the human data are inadequate for quantitative risk assessment. However, the U.S.EPA does include the cancer risks based upon the epidemiological data in its documents. Apparently, the U.S.EPA found the epidemiological data inadequate in the context of deriving a preferred single point estimate for diesel exhaust's carcinogenic potency. The U.S.EPA therefore derived a point estimate from the animal cancer data.

The full text of the sentence from the OEHHA Executive Summary quoted by the comment is "In their 1994 draft document, the U.S.EPA cited these same McClellan et al. (1986) UCL values as being practical in assessing human risks involving exposures in the range of study observations." In this quote, OEHHA does not cite the U.S.EPA has supporting use of the McClellan data for deriving a single point estimate. OEHHA did not develop a single point estimate.

Rather, OEHHA cited the U.S.EPA position as being consistent with use of the McClellan data in framing the <u>range</u> of cancer risks. This interpretation is consistent with the U.S.EPA presentation of the McClellan results in its Executive Summary and the subsequent statement found on page 1-26 of the U.S.EPA's 1994 draft document Executive Summary which states "In view of the uncertainty inherent in these types of calculations, the human and animal estimates should be viewed as complementary. For a bounding estimate intended to determine whether an exposure level as the potential to pose a hazard to human health, the published human estimates [the Harris, McClellan studies] may be practical for exposure levels in the range of observations in these studies. On the other hand projection of the public health impact of an exposure level may

benefit from using estimates derived from animal experiments, because of the closely controlled conditions and their precisely measured levels, absence of many confounding factors, and narrow confidence limits around the tumor incidence rates. A unit risk estimate of 3.4 x 10^{-5}/ μ/m^3 for continuous life-time exposure, which is the geometric mean of the upper bound estimates calculated from the three rat experiments is therefore recommended. The proper use and understanding of these risk estimates is discussed in Chapter 12."

Furthermore, OEHHA took care in characterizing the U.S.EPA work. OEHHA requested the U.S.EPA review our draft language prior to its public release. They did not object to this characterization of their work. In discussions of the OEHHA analyses with U.S.EPA staff, we have found them to be generally supportive of including epidemiological data in framing a range of risks. At the July 1, 1997 workshop Dr. Koppikar indicated that the U.S.EPA would be using epidemiological data from the Garshick et al. (1987) case-control study in the next quantitative risk assessment it presents to the Clean Air Scientific Advisory Committee. Dr. Koppikar also stated that the resulting risk values were "pretty consistent and very similar with what Cal-EPA has presented here today." (Transcript of the Public Workshop for the Diesel Exhaust Identification Report, p.75.)

With respect to sample concern #3, OEHHA has not glossed over the controversy in the analyses and interpretation of the work by Dr. Garshick. OEHHA and the ARB co-sponsored the workshop to facilitate the scientific discussion of the epidemiological data. Since the workshop, OEHHA has worked to further clarify the substantive scientific issues which remained murky after the workshop. For instance, the Armitage-Doll analyses of the Garshick (1988) cohort study presented in Appendix E were conducted in response to recommendations received at the workshop. To better understand the bases for the differing analyses by Dr. Crump and OEHHA of the Garshick et al. (1988) study, OEHHA shared its analyses with Dr. Crump so he could replicate our findings. Dr. Crump also shared his information with OEHHA so we could replicate his findings. OEHHA then worked to clarify the bases of the scientific disagreements. Appendix E presented our findings. Appendix F, which clarifies the bases for the scientific differences, was reviewed by Dr. Crump prior to its release. Where Dr. Crump made suggestions for improving the accuracy of our description of his work in Appendix F, we incorporated them. In our revised document we have provided another Appendix which discusses the differing conclusions regarding the cohort data set.

It is not clear what is being referred to regarding the statement that the dispute "must be settled" and that OEHHA refused repeated requests to participate in a forum. It has always been our intent to bring closure if possible. It has also been apparent that the conclusion may result in a clarification of the points in dispute without securing a mutual agreement as to the merits of the differing positions. Our new Appendix F attempts to further clarify the differing assumptions and their effects on the analyses.

Finally, OEHHA did not foreclose further consideration of the issue with the release of its March 1997 document. OEHHA released that document as a draft for public comment. At the October, 1997 public meeting of the independent Scientific Review Panel, OEHHA presented the scientific issues and specifically requested guidance on the matter. Pursuant to California Health and

Safety Code Section 39661, it is the independent Scientific Review Panel which determines whether or not the OEHHA health effects report is based upon sound science. OEHHA has now made its revisions in response to those public comments and Scientific Review Panel guidance. This resulting third document is now being released again with opportunity for public comment. Clearly, OEHHA and the ARB have provided an open public process with opportunity for extensive public comment on all the scientific issues related to its work.

Contrary to any inference from the comment's characterization that these issues were "consigned to inaccurate footnote discussion in Appendices E and F" to make them seem less important, these issues are given prominent attention in our document. The Executive Summary devotes a section to the subject (Section 1.4.3.1 Differing Analyses of the Garshick et al. (1988) cohort). This section refers the reader to the appendices for the detailed discussion. Furthermore, Chapter 7 (Section 7.3.2.4 Comparison To The Reanalyses Using The Individual Data) focuses on these issues and refers the reader to an Appendices E and F for more detailed information.

Nor, in any case, can the Appendices be reasonably construed as an informational burial grounds. Our chapters not only discuss and rely upon the information in the appendices, they refer the reader to the appendices for more information. Clearly, the appendices are presented as an integral part of the document. Any perusal of the Table of Contents would show that OEHHA has tended to put information there because it is somewhat complete on its own or sufficiently lengthy to compromise the efficient organization of other information in the existing document. Ironically, because the issues raised by this comment merited and received such a detailed and lengthy consideration, they were placed in our Appendices. This OEHHA approach was not unusual. In their 1994 draft document, the U.S.EPA (sponsor of the early work by Dr. Crump) similarly placed Dr. Crump's analyses of the Garshick et al. (1988) cohort study into appendices.

With respect to sample concern #4, the meta-analysis is presented in full in Appendix D. Chapter 6.2 discusses each of the studies included in the meta-analysis and 13 of the 16 excluded studies. This chapter (Section 6.2.2, Meta-analysis on the Relationship Between Occupational Exposure to Diesel Exhaust and Lung Cancer) presents the salient results of the meta-analysis and refers the reader to Appendix D for the details. The meta-analysis itself was according to standard design. The study exclusion criteria are presented in Appendix D and the excluded studies are listed, with the applicable reasons for exclusion in Table D-1. The potential for selection and publication bias are discussed in Appendix D as well.

With respect to the absence of any discussion of recent literature reviews by Drs. Muscat and Wynder in 1995 and by Drs. Stober and Abel in 1996, OEHHA was aware of their work and has past benefit of earlier comments from these individuals on our work. However, OEHHA generally compared and contrasted its review only to those of other authoritative bodies. Nevertheless, elsewhere in Part C, OEHHA again responds to the separate comments offered by Drs. Stober and Muscat. In contrast to the assessments of these individual authors, an expert panel assembled by the World Health Organization also reviewed the evidence available, and concluded that the evidence was consistent with a causal association. The WHO concluded that "The relative risks for lung cancer as a result of exposure to diesel exhaust are generally low, and risks of this magnitude are more susceptible to chance and to the effects of unmeasured confounding factors

and imprecision for adjusting for known confounding factors. As discussed above, the elevated risk for lung cancer in the four most informative studies is unlikely to be due to confounding by cigarette smoking and is probably due to exposure to diesel exhaust. Other studies, although limited primarily by the exposure ascertainment, support this assessment." International Programme on Chemical Safety, Diesel Fuel and Exhaust Emissions. Environmental Health Criteria 171. World Health Organization, Geneva 1996. p.252. Ultimately, however, it is not the combined weight of the authoritative bodies which persuades, it is the combined weight of the available evidence.

With respect to sample concern #5, OEHHA did assert on pages 1-9 of the 1997 draft of its report that "A reasonable and likely explanation for the increased rates of lung cancer observed in the epidemiological studies is a causal association of diesel exhaust exposure with lung cancer."

The comment correctly implies that statistics alone cannot serve as the basis for causal inference. Other factors that should be considered are discussed at length in section 6.2.4 of the March 1997 version of the document, and include the consistency of the findings, the possibility that findings are due to bias or chance, the evidence for exposure-response relationships, the temporal sequence and biological plausibility of the associations.

OEHHA assessed causal inference using standard criteria. These criteria included 1) the consistency of the findings; 2) the strength of the associations, 3) the possibility that the findings were due to bias, 4) the probability that the findings were due to chance, 5) evidence of exposure response relationships, 6) temporality of the associations, and 7) biological plausibility of the associations. The great majority of the epidemiological studies find an association. The small magnitude of the relative risk increases the potential for confounding. However, the number and diversity of the occupations studied, and the various analyses of sources of confounding (e.g. smoking, ETS exposure, recall bias, informational bias) do not indicate that confounding or chance accounts for the observed results. While limited exposure intensity information was available, based upon duration of exposure, there was evidence of an exposure response trend. While biological plausibility is not required for causal inference, there is biological evidence to support the association: 1) diesel exhaust contains many mutagens, 2) diesel exhaust causes lung cancer in animal studies, 3) diesel exhaust contains many substances which occur in other complex mixtures which are respiratory carcinogens in the human, and 4) diesel exhaust contains known and probable human carcinogens.

Page 6-50 of the 1997 draft summarized the evidence giving rise to this interpretation: "The epidemiological studies concerning lung cancer risk and exposure to diesel exhaust provide evidence consistent with a causal relationship. The many associations found between lung cancer and diesel exposure are unlikely to be due to chance. Also, with the possible exception of the studies that did not take smoking into account, the findings reviewed above are unlikely to be due to confounding or bias. The results of various studies are consistent in the direction of an effect and are even somewhat similar in magnitude of effect. For example, all the studies of diesel railroad workers with adequate latency and more than 50 cases show evidence of an effect. Although the strength of the associations is weak, with low relative risk estimates being reported, several studies show clear exposure-response relationships. The temporal relationship between

exposures and lung cancer is consistent with a causal relationship. Finally, it is biologically plausible that exposure to diesel exhaust would increase the risk of lung cancer. Therefore, a reasonable and very likely explanation for the increased risks of lung cancer observed in the epidemiological studies is a causal association between diesel exhaust exposure and lung cancer."

Comment 8: The Draft Report Is Invalid. The implementing California statutes (Cal. Health and Safety Code §§39650-39661) require, in part, that the OEHHA Draft Report:

1. "utilize the best available scientific evidence gathered from the public, private industry, the scientific community, and federal, state and local agencies;"

2. "consider all available scientific data, including, but not limited to, relevant data provided by . . . international and federal health agencies, private industry, academic researchers, and public health and environmental organizations;" and

3. be based upon "sound scientific knowledge, methods and practices."

The OEHHA Report fails to clear this threshold by a wide margin. The Draft Report's conclusion that ambient concentrations (2.2 $\mu g/m^3$) of diesel exhaust will kill more than 2100 people each year in California is directly contrary to the conclusions of the Health Effects Institute and the American Congress of Governmental Industrial Hygienists. (ACGIH has proposed a TLV of 150 $\mu g/m^3$). The Draft Report's reliance on the Garshick studies for quantitative risk analysis is directly contrary to the conclusions of U.S. EPA, WHO/IPCS, HEI and Dr. Garshick himself. The Draft Report's reliance on Dr. Mauderly's rat studies for human risk analysis is contrary to the conclusions reached by HEI, CASAC and, most notably, Dr. Mauderly himself. And the Draft Report articulates a unit risk (2×10^{-3}**)** that is a full two orders of magnitude greater than that articulated by U.S. EPA and WHO/IPCS (3.4×10^{-5}). In sum, it appears that bias, not sound science, lies at the base of the Draft Report.

Response*: OEHHA has conducted a comprehensive review of the literature which "utilizes the best available scientific evidence gathered from the public, private industry, the scientific community, and federal, state and local agencies. In evaluating this complex information, OEHHA has endeavored to consider all available scientific data, including, but not limited to, relevant data provided by international and federal health agencies, private industry, academic researchers, and public health and environmental organizations." Repeatedly, OEHHA has sought broad public review, and comment on its work to assure the quality of its effort. In releasing three versions of the draft for public comment, OEHHA has twice revised its draft. To facilitate public input, review and comment, OEHHA has held numerous workshops to discuss its effort. Furthermore, OEHHA has brought controversial issues to the attention of the independent Scientific Review Panel and requested its guidance. OEHHA responds to each substantive comment received from the public in Part C. These analyses and responses speak for themselves as to whether OEHHA positions are based upon "sound scientific knowledge, methods and practices."*

OEHHA's draft report did not conclude, as the comment states, "that ambient concentrations (2.2 mg/m^3) of diesel exhaust will kill more than 2100 people each year in California." OEHHA has not determined that diesel exhaust would cause 2,143 lung cancers per year in California. We estimated that up to 10 to 2,000 Californians exposed to ambient levels of diesel exhaust over their lifetime could develop lung cancer per year. The 200-fold range of estimated life-time cancer risk provided by OEHHA captures much of the uncertainty involved in those estimations. The comment's value pertains to the top of this range.

With respect to the Threshold Limit Value established by the American Conference of Governmental Industrial Hygienists (ACGIH), the ACGIH is a non-governmental body which develops exposure guidelines, which are not regulatory standards, to limit occupational exposures, not environmental exposures. The ACGIH holds that its limits are not to be used for any other purpose including the "evaluation or control of community air pollution nuisances".

OEHHA responds separately to related comments from the HEI, Dr. Mauderly, and others below. With respect to the related comments of Dr. Mauderly, Dr. Garshick and others, we shared them with the Scientific Review Panel on October 16, 1997. There was a thorough discussion of the issue of using the rat data for quantitative human risk assessment at the meeting. The sense of the panel was that rat data and calculations provide useful information and should be left in the document, however, since human epidemiologic evidence was available on which to base the human risk estimate, that the human data should be used to form the range of risks. Therefore, OEHHA now bases the range of unit risk estimates only on the epidemiological information. However, it was also clear that the information presented to justify the existence of a threshold response in the rat data were not considered convincing by the Panel.

With respect to Dr. Garshick's concerns regarding our use of his study, Dr. Garshick's comments and the OEHHA responses to them are more fully discussed elsewhere in Part C. Dr. Garshick has not retracted his studies. His comments reflect the absence of exposure measurements in his study. They also relate to the uncertainties in the exposure history reconstructions and modeling assumptions involved in performing a quantitative risk assessment using his data. OEHHA acknowledges that uncertainties as to the study population's exposure histories and modeling approaches must reduce the degree of confidence in the results. OEHHA has presented a broad range of unit risk values which capture the uncertainties involved in the exposure data and modeling assumptions.

Comment 9: The Draft Report also is inconsistent with the conclusions recently articulated in the Report (May 24, 1996) of the Risk Assessment Advisory Committee (RAAC). These conclusions, contained in the RAAC Report, entitled A Review of the California Environmental Protection Agency's Risk Assessment Practices, Policies and Guidelines (hereinafter, the "RAAC Report'), and which must be implemented pursuant to Executive Order W-137-96, include the following:

1. OEHHA should assure consistency with U.S. EPA and other agencies (RAAC Report, pp. ES-6, ES-7, 3-4, 4-7, 4-1 0);

2. A formalized program for independent external peer review should be developed (RAAC Report, pp. ES-6, ES-8, ES-9, ES-1 4, 2-4, 2-24, 2-26, 2-27, 3-4, 3-9, 4-6);

3. The Agency needs "further resources" (i.e. lacks expertise) in human health effects and epidemiology. ["epidemiology is not well represented in Cal/EPA"] (RAAC Report, pp. ES-5, ES-12, 2-9);

4. The use of large uncertainty factors when the underlying data are poor should be avoided (RAAC Report, ES-15, 4-6, 4-8);

5. The uncertainties in models, data sets, and parameters and their relative contributions to total uncertainty in a risk assessment should be reported in written risk assessment documents, and when different models may be employed in a risk analysis, perhaps leading to different conclusions, parameter uncertainty should be analyzed at a similar level of detail for all the models (RAAC Report, p. 7-10); and

6. The Agency's "risk characterization practices fall somewhat short of what the profession now considers generally feasible" (RAAC Report, p. ES-16).

The OEHHA Draft Report runs afoul of each of these mandates and as a result falls short of the RAAC's recommendations which are in the process of being codified. Here, too, the OEHHA document misses the mark.

***Response**: With respect to the first issue concerning OEHHA's consistency with U.S. EPA and other agencies, in discussions of the OEHHA analyses with U.S.EPA staff, we have found them to be generally supportive of including epidemiological data in framing a range of risks. At the July 1, 1997 workshop Dr. Koppikar indicated that the U.S.EPA would be using epidemiological data from the Garshick et al. (1987) case-control study in the next quantitative risk assessment it presents to the Clean Air Scientific Advisory Committee. Dr. Koppikar also stated that the resulting risk values were "pretty consistent and very similar with what Cal-EPA has presented here today." (Transcript of the Public Workshop for the Diesel Exhaust Identification Report, p.75). OEHHA's proposed use of the animal data to which the comment elsewhere objects was substantially based upon the U.S.EPA approach. OEHHA notes that to the extent to which the comment elsewhere recommends that OEHHA not use the animal data in quantitative risk assessment, the comment itself advises OEHHA to depart from the U.S.EPA. Consistent with RAAC recommendations, where OEHHA substantially differs with the U.S.EPA we provide the scientific bases for the difference.*

With respect to the second concern that "a formalized program for independent external peer review should be developed", the independent Scientific Review Panel which is charged with reviewing Toxic Air Contaminant program documents (e.g. diesel exhaust) is the model upon which the RAAC recommendations for all of OEHHA was based.

With respect to the comment concerning Agency resources in epidemiology, OEHHA notes that the RAAC addressed the number of staff epidemiologists. It did not fault the expertise of the

individual staff epidemiologists. The Air Toxicology and Epidemiology Section has an epidemiology component on its staff. Furthermore to enhance the diesel exhaust effort, OEHHA contracted for additional epidemiological support from the University of California.

With respect to the comment regarding use of large uncertainty factors when the underlying data are poor, the OEHHA cancer quantitative risk assessment did not use the uncertainty factor approach. With respect to the non-cancer risk quantitative risk assessment, the underlying data were not judged to be poor, the uncertainty factors were not large, and the recommended value was identical to the U.S.EPA RfC.

OEHHA feels that the document, through the information in its Executive Summary, chapters, and appendices, meets the RAAC recommendation that "The uncertainties in models, data sets, and parameters and their relative contributions to total uncertainty in a risk assessment should be reported in written risk assessment documents, and when different models may be employed in a risk analysis, perhaps leading to different conclusions, parameter uncertainty should be analyzed at a similar level of detail for all the models."

With regard to the comment's quote of the RAAC report regarding risk characterization, the quoted recommendation does not address the validity of OEHHA's work. It addresses the presentation and scope of that work. The full quote of the RAAC report executive summary (ES-16) states "The procedures and practices of Cal/EPA in risk characterization are quite comparable to those of US EPA, with some differences generally attributable to differences in the laws implemented by the two agencies. Both agencies' risk characterization practices fall somewhat short of what the profession now considers generally feasible; however, the Committee believes that Cal/EPA is moving forward to improve its practices. The Committee recommends that Cal/EPA improve the characterization of uncertainty and variability in its risk assessments and the communication of this information to risk managers and the public.

There should be considerably more communication between the risk assessor and risk manager. Risk assessors should better understand the needs of risk managers in terms of expression of uncertainty and variability -- what the risk managers need and why they need it, and how it can be provided. Further the extent and depth of the analysis should be responsive to the needs of the decision-maker. In some cases a problem can be "overanalyzed", with an unnecessary expenditure of scarce resources. The depth of the risk assessment should be tailored toward the decision it is intended to support..."

OEHHA does not agree that the diesel exhaust document has run afoul of the RAAC recommendations. Rather, it has been developed with the benefit of those recommendations.

Comment 10: Finally, the Draft Report is inconsistent with recent decisions within the Ninth Circuit Court of Appeals. For example, in Daubert v. Merrell Dow Pharmaceuticals, Inc., 43 F.3d 1311, 1318 (9th Cir. 1995), the court found that one means of showing that a report is based on scientifically valid principles is to demonstrate "that the research and analysis supporting the proffered conclusions have been subjected to normal scientific scrutiny through peer review and publication." The OEHHA Draft Report cannot satisfy this criterion. Indeed, as EMA has noted

separately to ARB, the unreasonable timing and overly short review period provided in advance of the July 1 workshop seem calculated to evade peer review, not foster it as required by RAAC.

The Ninth Circuit has also concluded in Daubert at 1321, that "for an epidemiological study to show causation . . . the relative risk . . . arising from the epidemiological data will, at a minimum, have to exceed 2." See also Hall v. Baxter Health Care Corp., 947 F. Supp. 1387, 1403 (D. Or. 1996), where the district court stated that:

> The threshold for concluding that an agent was more likely the cause of a disease than not is relative risk greater than 2.0. Recall that a relative risk of 1.0 means that the agent has no effect on the incidence of disease. When the relative risk reaches 2.0, the agent is responsible for an equal number of cases of disease as all other background causes. Thus a relative risk of 2.0 implies a 50% likelihood that an exposed individual's disease was caused by the agent.

As supposedly confirmed by OEHHA's own "meta-analysis" the relative risk at issue here is approximately 1.43, well short of the minimum requisite 2.0 level of causation. Consequently, OEHHA's assertion that there is a causal association between diesel exhaust exposure and lung cancer will not withstand judicial scrutiny. It also will not withstand scrutiny under the standards articulated by the scientific community. For example, Dr. Frank Speizer, a co-author of the Garshick *et al.* (1988) cohort study, has emphasized the importance of finding a relative risk of 2.0 or more when investigating respiratory cancers. Writing in 1986, he states:

> "Because of the overwhelming effect of cigarette smoking, population-based studies that report on environmental effects, particularly at relatively low levels of excess risk (RR greater than 1.0 but less than 2.0), and that do not attempt to take cigarette smoking into account, must be considered seriously flawed. These studies, therefore, can contribute very little to our understanding of risk factors for respiratory cancer" (Speizer, 1986, Environmental Health Perspectives, 70:9-15, p. 9).

Courts within the Ninth Circuit also have found that in order for animal studies to be sufficient to prove causation in humans, "there must be good grounds to extrapolate from animals to humans, just as the methodology of the studies must constitute good grounds to reach conclusions about the animals themselves." ' Hall v. Baxter Health Care Corp., 947 F. Supp. at 1397. Moreover, "extrapolations of animal studies to human beings are generally not considered reliable in the absence of a scientific explanation of why such extrapolation is warranted." Id. at 1410. Here, as evidenced by the more recent findings and conclusions of Dr. Mauderly, as well as those of CASAC, there are no good grounds to extrapolate from the relevant animal studies to humans. Consequently, the Draft OEHHA Report will fail on this basis as well.

***Response**: The Daubert and Hall courts were concerned with causation pursuant to tort claims where an individual plaintiff has the burden of proving more likely than not that a particular defendant actually caused their alleged injury. Those courts generally held that Relative Risks exceeding 2.0 were needed to establish more likely than not causation with respect to an individual plaintiff's injury.*

In contrast to the subject matter of Daubert and Hall, the OEHHA document does not address civil liabilities for personal injury. OEHHA is concerned with the prevention of disease in the general population. OEHHA's risk assessment mandates are found in Section 39650 et seq. of the Health and Safety Code. Section 39650(e) provides that "while absolute and undisputed scientific evidence may not be available to determine the exact nature and extent of the risk from toxic air contaminants, it is necessary to take action to protect public health."

The OEHHA risk assessment guidelines and similar U.S.EPA guideline documents do not recognize the Relative Risk value of 2.0 to be a threshold of concern. Even weak associations can be important public health concerns. Relative risk (RR) is defined as "The ratio of the risk of disease or death among the exposed to the risk among the unexposed." (Last 1988). Neither Monson (1990) nor the other texts consulted by OEHHA staff make a categorical assertion that weak associations cannot represent causal relationships. The principal concern with weak associations is that they may be due to uncontrolled confounding. This does not mean that uncontrolled confounding cannot also affect the interpretation of studies in which large relative risks are reported; rather that a confounder that could provide the entire explanation for a strong observation should be easier to identify, measure, and control. Regardless of the magnitude of relative risk, however, a variety of biases (e.g., observation bias, selection bias) may affect the interpretation of the observed associations.

Furthermore, if the implications of the commentator's statement were true, that is, if so-called "weak" or "very weak" associations could not serve as a basis for causal inference, then much of what passes for clinical knowledge in medicine today is without epistemological foundation. Weak associations are found in many areas of medicine and, in conjunction with other aspects of causal inference, constitute the basis for a variety of clinical and public health interventions to prevent or ameliorate disease. To place the commentator's "rule of thumb" in context, the following table represents estimates of relative risk of death from cardiovascular disease due to cigarette smoking in several prospective epidemiological studies:

Study on cigarette smoking	*Estimate of RR of death from cardiovascular disease*
British doctors	*1.6*
Males in 25 states	
ages 45-64	*2.08*
ages 65-79	*1.36*
U.S. Veterans	*1.74*
Japanese study	*1.96*
Canadian veterans	*1.6*
Males in nine states	*1.70*
Swedish males	*1.7*
Swedish females	*1.3*
California occupations	*2.0*

Source: U.S. Department of Health and Human Services (1989)

According to the comment's relative risk criterion, most of these reported associations between cigarette smoking and heart disease would be classified as either "weak" or "very weak". Yet active cigarette smoking is widely recognized as one of the principal causes of heart disease; several years ago, the Centers for Disease Control estimated that 156,000 deaths/year were due to heart disease caused by cigarette smoking.

Similarly, several large national reviews of the evidence have concluded that environmental tobacco smoke (ETS) exposure causes lung cancer in nonsmokers, even though most pooled RRs are in the range of about 1.2 to 1.9 (National Research Council, 1986; U.S. Department of Health and Human Services 1986; National Institutes of Health 1993).

Here, even if the relative risks are below 2.0, the 1997 draft OEHHA risk assessment estimated that diesel exhaust exposure may cause up to 10 to 2000 cases of lung cancer in California each year.

Comment 11: Instead of pursuing a course of action that clearly is not based on sound science and that seems destined to lead to protracted and contested proceedings, OEHHA should actually utilize the best available scientific evidence and conclude that, at this time, a reliable quantitative risk assessment for the constituent(s) of diesel exhaust is simply not possible. That being said, EMA remains willing to participate in and sponsor appropriate studies to advance our knowledge of the subject to the point where quantitative risk assessments could be feasible. We already have undertaken efforts with HEI to pursue that objective. We encourage OEHHA also to participate in these efforts and make the necessary revisions to the Draft Report such that scarce resources can be allocated to further research instead of further confrontations.

Response: *OEHHA has endeavored to base its risk assessment on sound science. Here, OEHHA believes the development of a range of unit risks rather than a single preferred cancer unit risk value adequately captures the uncertainty in the estimates. In any case, the OEHHA work will be subject to independent scientific review by the Scientific Review Panel established for the Toxic Air Contaminant Program by statute.*

Comments on behalf of the Engine Manufacturer's Association prepared by Peter Valberg of Gradient Corporation, dated August 15, 1997

Comment on Human Evidence of Carcinogenicity

Comment 1: OEHHA Risk Estimate Calculated Using the Garshick *et al* (1987) Case-Control Data. OEHHA (1997) uses the results from Garshick *et al.'s* case-control study (1987) of railroad workers along with computations carried out by McClellan *et al.* (1989) to calculate a unit risk estimate (incremental risk per average lifetime exposure to diesel exhaust in $\mu g/m^3$).

Garshick *et al.* (1987) matched each of 1,256 lung cancer deaths identified from 15,059 death records for former railroad workers in the United States with two control cases. The control cases were drawn from the same set of 15,059 death records, and were matched with cases having nearly the same birth and death dates. The cause of death for each control was listed as a *"specified natural cause with no mention of cancer on the death certificate"* (pp. 7-15 Part B, OEHHA). Logistical regression that controlled for the effect of age, asbestos exposure, and smoking history suggested that the risk of lung-cancer increased by 1.648% for each year of exposure to diesel exhaust (as assumed from years of employment).

McClellan computed the incremental risk of lung cancer per μg-year of exposure to diesel exhaust by assuming that workplace diesel exhaust concentrations were 125 $\mu g/m^3$ or 500 $\mu g/m^3$, and by correcting for exposure frequency in an occupational setting (40 hours per week, 5 days per week, 48 weeks per year). McClellan used the Garshick *et al.* (1987) result of 1.648% increased risk per year to compute the incremental risk of lung cancer per lifetime average exposure to 1 $\mu g/m^3$ diesel exhaust. Assuming that occupational diesel exhaust concentrations were 125 $\mu g/m^3$, the incremental unit risk came out to 1.16×10^{-3}; assuming that occupational diesel exhaust concentrations were 500 $\mu g/m^3$, the incremental unit risk came out to 2.90×10^{-4}.

There are three problems with OEHHA's interpretation of the results from this study. First, the study does not provide a quantitative estimate of risk because exposure was not measured. Second, the results from the Garshick *et al.* 1987 study do not support the downward extrapolation of risks to exposure levels typical of non-occupational exposures. Third, Dr. McClellan has declared to CARB that his quantitative estimates are no longer valid.

***Response**: Responses to the first two of the three problems mentioned in this comment are given in connection with the more detailed comments which follow. Responses to Dr. McClellan are given in connection with his letter. Brief responses are provided here. The data on relative risk were used in conjunction with the information from McClellan et al. (1989) to conduct the risk assessment. The Garshick et al. (1987) study was designed to examine the association between diesel exhaust exposure and lung cancer. While the Garshick et al. (1987) study did not have sufficient data to conduct quantitative risk assessment, OEHHA relied upon other published reports of exposures in the relevant industries to estimate a reasonable range of exposures so as to derive a range of cancer unit risk. As discussed below, the criticism regarding extrapolation of the case-control study are not justified. Finally, Dr. McClellan asserts that "a rigorous reanalysis of the Garshick et al. data raises serious concerns about the validity of the conclusions*

drawn by Garshick et al. The detailed reanalysis was conducted by Crump and associates (1991)." There are two serious problems with this reasoning. The first is that the conclusions from the reanalysis of the cohort study has without further support a connection to the other study. The other is that a problem with misclassification invalidated the second half of the Crump et al. report, and the TSD has shown that the finding in the first half is not representative of the range of plausible findings, which show a positive relation of risk to exposure. Furthermore, we are now using data from Woskie et al (1988a,b) to derive exposure estimates for the case-control study.

Comment 2: Garshick *et at* (1987) <u>Does Not</u> Provide a Basis for Quantitative Risk Assessment. OEHHA's analysis of Garshick *et al.* (1987) computes risk as a function of employment duration in the railroad industry. That is, the computation does not assess the relationship between risk and diesel exhaust concentration, a value that is assumed to be constant across all workers during their employment in this industry. In Section 2.1.2, we demonstrate that changes in diesel exhaust concentration appear to be unrelated to changes in risk. The relative risk identified may therefore represent the risk associated with some factor associated with duration of employment in the railroad industry other than exposure to diesel exhaust. For example, levels of environmental tobacco smoke (ETS) in the railroad industry may have been elevated relative to other occupations. Also, the dichotomous classification of smokers vs. non-smokers is likely to allow considerable residual confounding from smoking.

In short the analysis of Garshick *et al.* (1987) data does not show a relationship between exposure to diesel exhaust and the risk of lung cancer, but instead shows a relationship between lung cancer risk and employment duration (and with whatever factors "employment duration" may be correlated). The failure of the Garshick *et al.* (1988) cohort data (discussed in Section 2.1.2) to establish an association between diesel exhaust concentration and lung cancer risk supports the hypothesis that some other factor is responsible for the relative risk quantified on the basis of the Garshick *et al.* (1987) case-control study.

At the very least, because the diesel exhaust exposure concentrations assumed by OEHHA in its derivation of a unit risk are completely hypothetical, the unit risk results cannot be regarded as quantitative estimates. Even if exposure to diesel exhaust were associated with some level of increased risk of lung cancer, OEHHA's analysis of the Garshick *et al.* (1987) case-control data is not a valid estimate of this risk's magnitude.

Response: *The comment mentions environmental tobacco smoke (ETS) as an example of an effect possibly accounting for the higher risk among the workers exposed to diesel exhaust. However, Woskie et al. (1988a) measured higher ETS levels among clerks than among workers exposed to diesel exhaust. The comment also mentions that the dichotomous classification of smokers vs. non-smokers is likely to allow considerable residual confounding from smoking. However, Garshick et al. used "lifetime smoking (pack-years)" to do the adjustment.*

The TSD used and still uses a relationship between lung cancer risk and employment duration in exposed jobs as a basic step in developing a quantitative risk assessment from the case-control study of Garshick et al. (1987a). In the version that the comment addresses, the TSD then simply used the constant exposure assumptions and the analysis and the results of McClellan. The

present version considers an extended range of exposure assumptions, including time-varying concentration in new calculations to approximate a range of unit risks. These are not "hypothetical" but are based upon Woskie et al (1988 a,b) and other available data.

Comment 3: OEHHA's Analysis of Garshick *et al* (1987) Does not Support Downward Extrapolation of Risks to Levels Below Occupational Exposures. Setting aside the problems due to a lack of actual measurements of diesel exhaust concentration, the Garshick *et al.* (1987) case-control study does not support OEHHA's attempt to extrapolate risks from occupational levels to much lower non-occupational levels. Garshick *et al.'s* analysis suggests that virtually all of the elevated lung cancer risk is associated with occupational exposures exceeding 20 years. Garshick *et al.* (1987) notes on p. 1244 that *"subjects with ³ 20 yr. of diesel exhaust exposure had the highest, significantly elevated odds ratio (OR = 1. 64, 95% CI = 1. 18, 2.29). Subjects with 5 to 19 diesel-years had an odds ratio of 1. 02 (95% CI = 0. 72, 1.45)."* That is, the risk of lung cancer among workers exposed to diesel exhaust between 5 and 19 years was essentially unchanged, compared to the risk of lung cancer in workers with between 0 and 4 years of diesel exposure.

Let us assume that:

- The relevant measure of dose is the product of incremental diesel concentration above background and duration of exposure;

- The incremental diesel exposure among exposed workers can be estimated as 43 $\mu g/m^3$ (see Section 2.1.2 of our report or p. 7-18 in OEHHA, 1997, Part B);

- The average exposure duration among workers with between 5 and 19 years of exposure is 12 years (the average of 5 and 19 years), while the average exposure duration among workers with between 0 and 4 years of exposure is 2 years (the average of 0 and 4).

From these assumptions, it follows that the incremental exposure to diesel exhaust among workers with between 5 and 19 years of occupational experience in the railroad industry (in μg-years/m^3) is (12 years - 2 years) x 43 $\mu g/m^3$, or 430 μg-years/m^3.

OEHHA notes that *"the average annual ambient concentration of diesel exhaust to which Californians are exposed is 2.2 $\mu g/m^3$...,* (OEHHA, 1997, Part B, pp. 7-28). Multiplying this value by an average life span of 70 years yields 154 μg-years/m^3, far less than the 430 μg-years/m^3 that Garshick *et al.* (1987) indicates is **without elevated risk.** Hence, even if the Garshick *et al.* (1987) results are accepted at face value, they do not support OEHHA's inference that exposure to diesel exhaust is responsible for lung cancer among non-occupationally exposed individuals living in California.

***Response**: Garshick et al. (1987) presented the results of a two-step categorical analysis of the trend of risk with duration of exposure, as quoted in the comment. Contrary to the comment Garshick et al. did not indicate that the average exposure for the lower of the two categories was "without elevated risk." On the contrary, the confidence interval quoted in the comment implies*

that the lower category would have a 2.5% chance of an excess relative risk of 45%, or greater, on average. This is a substantial population risk. Furthermore, the categorical analysis over the whole range of exposure durations gave an average increment of risk of 1.648% for each year of exposure duration. This value does not indicate that duration exposures in the lower category are without elevated risk.

Comment 4: Comments on OEHHA Risk Estimate Calculated Using the Garshick *et al*. (1988) Cohort Data Analysis of the 1994 California EPA Risk Assessment. The 1994 OEHHA risk assessment divides two groups of workers (exposed and unexposed) into four exposure categories based on the length of time they worked in the railroad industry after 1959. OEHHA relies on Garshick *et al.'s* reasoning that exposure to diesel exhaust **prior to 1959** was the same among all exposed workers in the study cohort. Hence, any differences in exposure among members of this cohort reflect differences in work tenure **following 1959.**

For each exposure category, Garshick *et al.* reports the relative risk of lung cancer compared to the corresponding cohort of unexposed workers. For example, the first exposure category includes workers who remained in the railroad industry for up to four years following 1959 (i.e., workers who left the industry between 1959 and 1963). Garshick *et al.* calculated the relative risk of exposure for these workers by comparing them to **unexposed** workers who remained in the railroad industry for up to four years following 1959. The remaining three exposure categories include workers with between 5 and 9 years of railroad employment following 1959, workers with between 10 and 14 years of railroad employment following 1959, and workers with between 15 and 17 years of railroad employment following 1959.

Additionally, Garshick *et al.* reported relative risks using two different definitions of the "exposed" worker cohort. The first set of relative risk values reflects lung cancer prevalence among what Garshick *et al***.** refers to as all "exposed" workers. This group includes workers in various job categories thought to expose subjects to diesel exhaust, **including shopworkers.** The second set of reported relative risk values reflects lung cancer prevalence among all exposed workers, **excluding shopworkers.** Both sets of relative risk values appear in Table 2.1

Table 2.1. Relative Risk of Lung Cancer for Exposed Workers as Reported by Garshick *et al.* (1988)

	Relative Risk of Lung Cancer Compared to "Unexposed"' Railroad Workers with the Same Job Tenure	
Years Worked in Railroad Industry After 1959	**Exposed Workers, Including Shopworkers'**	**Exposed Workers Excluding Shopworkers'**
1 to 4	1.20	1.34
5 to 9	1.24	1.33
10 to 14	1.32	1.33
15 to 17	1.72	1.82

OEHHA used data from Woskie *et al.* (1988a,b), along with anecdotal information about changes in exposure conditions in the railroad industry, to quantify atmospheric diesel concentrations over time. Woskie *et al.* (1988a) report atmospheric diesel concentrations measured in 1983 for various railroad job categories. OEHHA (1994) assumed that historical concentrations equaled contemporary concentrations [i.e., 1983 concentrations reported by Woskie *et al.* (1988a)] multiplied by an exposure factor. Specifically, **OEHHA's 1994 analysis assumed that:**

- **For shopworkers,** the exposure factor was 0.0 in 1945 and increased linearly to a peak value of 15.0 in 1959. OEHHA inferred the factor of 15.0 from historical measurements of nitrogen dioxide in railroad shops. (To allow for assumed, but not established heterogeneity in shopworker exposure, only one-half were assumed to be exposed at the measured levels.)

- **For other exposed workers,** the exposure factor was 0.0 in 1945, and peaked at 2.0 in 1959.

- **For all exposed workers** (including exposed shopworkers), the exposure factor was 2.0 between 1960 and 1970 and then dropped to 1.0 thereafter.

Figure 7-2 in OEHHA (1994) illustrates the exposure factor functions for shopworkers and other exposed workers, while Figure 7-2 in OEHHA (1997) illustrates the exposure factor function for exposed workers excluding shopworkers (the 1997 document includes only one exposure factor function since, as explained in Section 2.2.1c, OEHHA (1997) excludes shopworkers from consideration). Section 2.1.2b (below) describes in detail the relationship between contemporary (1983) diesel concentrations and historical concentrations assumed in OEHHA's 1997 analysis. In short, OEHHA assumed in 1997 that the ratio of historical concentrations to 1983 concentrations among other exposed workers is somewhat larger than the ratio assumed in OEHHA's 1994 analysis. In contrast, OEHHA assumed in 1997 that the ratio of historical concentrations to 1983 concentrations among exposed shopworkers is substantially smaller than the ratio assumed in OEHHA's 1994 analysis. As we explain in Section 2.1.2b, the 1997 historical exposure

assumptions, which lack foundation, eliminate the influence of including or excluding shopworkers from the analysis of the dose-response relationship.

OEHHA's 1994 analysis calculated two measures of incremental exposure for the exposed workers relative to the unexposed workers. The first measure was μg-years/m^3, which was computed by multiplying the area under the exposure factor function by the 1983 atmospheric diesel concentration reported by Woskie *et al.* (1988a). Application of appropriate multiplicative factors to account for occupational exposure frequency and duration yielded a lifetime average exposure in μg/m^3. The second measure is calculated using the first measure, but applies a physiological model to determine the total incremental deposition of diesel exhaust in the lungs for shopworkers and for other exposed workers. OEHHA's 1994 analysis computed exposure for the cohort consisting of all exposed workers and all shopworkers making the assumption that half of the shopworkers were exposed to elevated diesel exhaust concentrations, while the other half of the shopworkers were unexposed.

Plotting the relative risk values reported by Garshick *et al.* against the corresponding exposure measures computed by OEHHA yields two sets of dose response curves (one for exposure measured in terms of lifetime average μg/m^3, and one in terms of lifetime diesel exhaust deposition in the lung (mg)). Our critique of OEHHA's 1994 analysis was straightforward. Because Garshick *et al.* reports two sets of relative risks - one for exposed workers excluding shopworkers, and the other for all exposed workers and all shopworkers - we computed lifetime exposure for these two groups.[1] These exposure levels appear in Tables 2.2a (measured in terms of average lifetime exposure concentration in μg/m^3) and 2.2b (measured in terms of total lifetime lung burden in mg).[2] We then compared the dose-response relationships for these two groups, showing that the dose-response relationship is not monotonic.

1

2

[1] We note here that Gradient's 1994 comments incorrectly excluded unexposed shopworkers from the calculation of exposure for the second of these two cohorts. Nonetheless, our findings remain the same, as described in this section.

[2] Note that, like OEHHA (1994), the calculations in Tables 2.2a and 2.2b are based on the assumption that only 50% of the 12,092 shopworkers are exposed (i.e., 6,046 shopworkers are exposed).

Table 2.2.a. Average Incremental Lifetime Exposure to Diesel for the Garshick *et al.* (1988) Cohort: Average Concentration During Employment in the Railroad Industry Converted to Lifetime Average Concentration (μg/m^3)

	Exposed Shopworkers N=6,046 Con>Background: 102 Tg/m^3 Lifetime Avg. Exposure (Tg/m^3)			Unexposed Shop-workers N=6,046 Con>Background: 0 Tg/m^3 Lifetime Avg. Exposure (Tg/m^3)	Other Exposed Workers N=6,046 Con>Background: 43 Tg/m^3 Lifetime Avg. Exposure (Tg/m^3			All Exposed Workers N=6,046	
Exposure Category	AUC[b]	Un-adjusted[c]	Adjusted[c]	Adjusted	AUC[b]	Un-adjusted[c]	Adjusted[c]	Un-adjusted[c]	Adjusted[c]
1-4 yr.	115.5	168.3	55.5	0	18	11.0	3.6	32.0	10.7
5-9 yr.	124.5	181.4	59.8	0	27	16.6	5.5	37.9	12.6
10-14 yr.	133.5	194.5	64.1	0	36	22.1	7.3	43.6	14.5
15-17 yr.	137.5	200.3	66.00	0	40	24.6	8.1	46.1	15.4

Notes:

a Garshick *et al. calculate lung cancer relative risk for the following subsets of his cohort. individuals who continued railroad work for 1 to 4 years after 1959, individuals who continued railroad work for 5 to 9 years following 1959, individuals who continued railroad work for 10 to 14 years following 1959, and individuals who continued railroad worker 15 to 17 years following 1959. OEHHA (1994) approximated exposure or each of these groups by using the midpoint of each of these ranges For example, OEHHA assumed that individuals who continued railroad work for 10 to 14 years past 1959 worked in the railroad industry from no later than 1945 through 1971.*

b *"AUC " is the "Area under the Curve "for the exposure factor (see Figure 7-2 in OEHHA, 1994) between 1945 and retirement. The AUC values in this table appear in columns 2 and 3 of Table 7-8 in OEHHA (1994). The exposure factor function used in the 1997 analysis differs substantially (see Figure 7-2 in OEHHA, 1997, Part B).*

c *Unadjusted lifetime average exposure is computed by multiplying the AUC value by the contemporary incremental exposure above background (which yields mg-years/m^3) and then dividing this result by the average duration of a lifetime (assumed to be 70 years). The adjusted lifetime exposure (mg/m^3) is the unadjusted value multiplied by the product of: 1/2 ((proportion of inhalation that takes place at work), 5/7 (number of days at work each week), and 48/52 (number of weeks at work each year) to calculate the adjusted exposure value. The overall product of these factors is 0 33. Hence the unadjusted lifetime average exposure for exposed shopworkers of 168.3 mg/m^3 is 115.5 (AUC) x 102 mg/m^3 (contemporary incremental exposure above background) ¸- 70 years (lifetime duration). The adjusted lifetime average exposure of 55.5 mg/m^3 is equal to the unadjusted exposure of 168.3 mg/m^3 multiplied by 0.33.*

d *Although the Woskie et al. data do not suggest that the shopworkers are a heterogeneous population in so far as diesel exhaust exposures, OEHHA assumed in 1994 that half the shopworkers were "unexposed " The lifetime average exposure for* ***unexposed*** *shopworkers is zero since the incremental diesel exhaust concentration above background for these workers is zero. In their 1997 analysis, OEHHA omits the shopworkers entirely, even though workers for whom exposure was greatest would be expected to provide the greatest "signal " for lung cancer risk*

e *The lifetime average exposure for exposed workers and all shopworkers is the average exposure for exposed shopworkers, other exposed workers, and unexposed shop workers, weighted by the number of individuals in each category.*

Table 2.2.b. Average Incremental Lifetime Exposure to Diesel for the Garshick *et al.* (1988) Cohort:Incremental Lung Burden During Employment in the Railroad Industry Converted to Average Annual Lung Burden[a]

Exposure Category	Exposed shopworkers N=6,046 Incremental Deposition Above Background (mg)[b]	Unexposed Shopworkers N=6,046 Incremental Deposition Above Background (mg)[c]	Other Exposed Workers N=29,290 Incremental Deposition Above Background (mg)[d]	All Exposed Workers and Unexposed Shopworkers N=41,382 Incremental Deposition Above Background (mg)[e]
1-4 yr.	1416.3	0	6.2	65.2
5-9 yr.	655.2	0	10.6	103.2
10-14 yr.	892.7	0	15.0	141.0
15-17 yr.	1052.0	0	17.5	166.1

Notes

a *Total* particle *deposition in* the lungs *of unexposed* workers is assumed *to be 3.5* mg *(1-4* yr.); *4.4 mg (5-9 yr.);* 5.5 *mg (I* 0- *1* 4 *yr.); and* 7.2 *mg* (15-1 7 *yr.). Incremental exposures equal* total deposition minus these *background* total *deposition values. These values were* computed *by dividing values in the fourth column of Table 7-9* in *OEHHA (I 994)* (mg**yr for unexposed shopworkers) by the lifetime duration of 70 years.*

b *Computed by dividing the* values *in* column *2 of Table 7-9 in OEHHA (1994) by 70 and subtracting the background exposure* values *detailed in footnote (a).*

c *Incremental exposure* above background for unexposed shopworkers is zero.

d *Computed* by dividing *the values* in *column* 3 of Table 7-9 *in* OEHHA (1994) by 70 and subtracting the *background exposure values detailed* in footnote *(a).*

e *Incremental lifetime average annual* deposition *above* background for all exposed workers and all *shopworkers is the average of the* incremental *deposition for exposed shopworkers, unexposed shopworkers, and other* exposed *workers, weighted* by the *number* of individuals in *each* of *these* groups (i.e., weighted by *N).*

In the present report, we make the same point as in the 1994 Gradient report, but in a slightly different manner. Specifically, using the relative risk values published by Garshick *et al.* (1988) and the exposure levels computed in Tables 2.2a and 2.2b for the two exposed worker groups, we created four dose-response lines - one for each exposure category (railroad employment up to 4 years following 1959, railroad employment 5 to 9 years following 1959, and so forth). The relative risks for the two sets of exposed workers in each exposure category tended to be similar (compare column entries in Table 2.1 above). However, the exposure levels differed substantially (see Tables 2.2a and 2.2b). Specifically, average exposure among exposed workers **excluding shopworkers** was always substantially less than average exposure among exposed workers **including shopworkers.** Hence, the dose-response lines are, in all four cases, nearly horizontal, or even slightly downward sloping. Figure 2-1 illustrates the four dose-response lines where exposure is measured in terms of average lifetime exposure concentration in μg/m3; Figure 2-2 illustrates the four dose-response lines where exposure is measured in terms of lifetime lung

burden. In any case, the Garshick *et al.* results, together with OEHHA's 1994 reconstruction of historical exposure, show no evidence of an association between higher diesel exhaust concentrations and an increased risk of lung cancer.

As noted at the beginning of this section, OEHHA's 1997 analysis of the Garshick *et al.* cohort data differs in some respects from its 1994 analysis. In an effort to eliminate the "paradox" revealed by the preceding analysis, OEHHA has:

- Re-quantified historic exposure levels so that the exposure levels in the cohort of all exposed workers and all shopworkers do not differ substantially from the exposure levels in the cohort of exposed workers excluding shopworkers. This change has the effect of making the exposure levels for the two groups in each exposure category the same, thus eliminating the horizontal dose-response lines.

- Omitted the shopworkers from the analysis, hence eliminating the right endpoint of each of the dose-response lines in Figures 2-1 and 2-2.

- Omitted analysis of the exposure in terms of total lung-deposited diesel exhaust - the exposure measure that offered the most dramatic illustration of the absence of an upward-sloping, dose-response relationship (see Figure 2-2).

Sections 2.1.2b, 2.1.2c, and 2.1.2d challenge the validity of these changes to OEHHA's analysis.

Response*: The small decrease in relative risk (relative hazard) for exclusion of shop workers appears to be far from statistically significant, and the comment furnishes no statistical analysis to support the claim. Furthermore, as pointed out in the comment, the proportion of unexposed shop workers is not known and could effectively be greater than 50% thus tending to lower exposure of shop workers below that of workers riding on trains. Such uncertainties are among the main reasons for excluding the shop workers from all analyses in the current version of the TSD.*

Comment 5: OEHHA's New Estimate of Historic Exposures Lacks Adequate Foundation. We identify two problems with OEHHA's new estimates of historic exposure to diesel exhaust among railroad workers. First, compared to its 1994 analysis, OEHHA has, with no explanation, increased the peak exposure (achieved in 1959) among the cohort of exposed railroad workers excluding shopworkers by 50%. Second, OEHHA has implicitly decreased the assumed peak exposure level for exposed shopworkers, again without adequate explanation or justification (other than the fact that the results better **fit** the desired conclusion). The effect of these changes is to make the calculated average historic exposure to diesel exhaust among exposed railroad workers insensitive to the inclusion or exclusion of shopworkers from this cohort thus eliminating the paradox we described in our Section 2.1.2a.

OEHHA explains the basis for its 1994 estimate of historic diesel exposures in Section 7.3.5 of OEHHA (1994). Of particular interest are the second and third paragraphs in that section. Here, OEHHA infers historic diesel concentrations based *on "limited historical data on nitrogen*

dioxide levels in railway repair shops. " OEHHA continues, *"These data showed that in the era designated 'pre-ventilation'' (1950-1959) the average level (based on 22 measurements at four locations) was 10 times the average measurement for 1983 (based on 238 measurements)..."* Using this information, OEHHA concludes that for shopworkers, *"The exposure factor for 1945-1960 was taken to increase linearly from zero to a peak value of 15x in 1960"* (pp. 7-14). For other exposed workers, OEHHA assumes that diesel exposure increases from 0 in 1945 to twice current levels in 1960 (see pp. 7-15 in OEHHA, 1994).

For its **revised** analysis, OEHHA takes up the issue of historic exposures in Section 7.3.2.2.2 of OEHHA (I 997). The extent to which historic diesel concentrations exceed contemporary levels is addressed in the final paragraph of that section. Here, OEHHA states that *"For train workers, the exposed group in the present calculation, the linear rise from 0 in 1945 is assumed to peak at 3 times the 1983 level at the beginning of 1959..."* (pp. 7-17). OEHHA explains that *"The exposure factor of 3 is a scaled down version of the factor of 10 that Woskie et al. (1988b) reported for exposure of railroad shopworkers to nitrogen dioxide. "* We note that **OEHHA does not explain why its 1997 analysis assumes a peak exposure concentration for exposed railroad workers excluding shopworkers that exceeds contemporary levels by a factor of 3, while its 1994 analysis assumes peak levels exceeded historic levels by a factor of** 2. Moreover, the *':factor of* 10" reported by Woskie *et al.* (1988b) specifically applied to **repair shops.** OEHHA does not justify why a "scaled down" version of the Woskie *et al.* (1988b) factor of 10 should be used.

Even more dramatic is OEHHA's most recent revision of assumptions regarding exposure of shopworkers. In the fifth item in its discussion of sources of uncertainty in its 1997 analysis (Section 7.3.3), OEHHA (1997) asserts that its exclusion of shopworkers from its analysis does not substantially affect its results. Specifically, OEHHA states that, "If... *the proportion of unexposed shopworkers is set at 0.5 ... then the overall risk for the shopworkers would be about the same as for the train workers. With these assumptions, this finding of Garshick et al. on the effect of excluding shopworkers would be within random variation... "* (OEHHA, 1997, Part B, pp. 7-21).

Assuming that half the shopworkers are exposed and the other half are unexposed, OEHHA's assertion that the overall risk for shopworkers is the same as for other exposed workers can hold only if one assumes that the area under the "exposure factor" curve multiplied by the incremental contemporary exposure for exposed shopworkers (102 $\mu m/m^3$) is only twice the corresponding value for other exposed workers (for whom the incremental contemporary exposure is 43 $\mu m/m^3$). Because the ratio of 102 $\mu m/m^3$ to 43 $\mu m/m^3$ is 2.37, this condition implies that the area under the exposure factor curve for shopworkers must be less than the corresponding area for other exposed workers. Specifically, the area under the exposure factor curve for exposed shopworkers must be only 84.3% of the area under the exposure factor curve for other exposed workers (2-0 - = 2.37). This condition is in direct contradiction to the only empirical evidence available (the nitrogen dioxide measurements cited in OEHHA's 1994 document).

OEHHA provides no evidence or reason to explain why the ratio of the areas under the exposure factor curves for shopworkers vs. other exposed workers is now less than unity, compared to

OEHHA's 1994 estimated ratio, which greatly exceeded unity[3]. ([3] The ratio of the area under the exposure factor curve for exposed shopworkers to the area under the exposure factor curve for other exposed workers in OEHHA's 1994 analysis ranges from a maximum value of 7.5 (for workers terminating employment in 1960) to a minimum value of 3.17 for workers terminating employment in 1980 (see Figure 7-2 in OEHHA, 1994). Assuming that 50% of shopworkers are unexposed, the corresponding ratio comparing the AUC for all shopworkers to the AUC for other exposed workers would range from a maximum of 3.75 (50% x 7.5) to a minimum of *1.58 (50%* x 3.17).) The absence of such an explanation renders OEHHA's 1997 analysis of the Garshick *et al.* cohort data invalid.

***Response**: In regard to the assertion that "OEHHA has, with no explanation, increased the peak exposure," the choice of the roof exposure pattern in the 1997 draft TSD was the result of thinking afresh the nature of the likely peak exposure of workers on trains in 1959. At the January 1996 Scientific Workshop there was more support than not for that roof pattern, with the 3-fold peak. In the current draft TSD more exposure patterns have been added to show the range of results in view of the uncertainties. The current draft TSD discusses evidence for a decline of exposure concentration from the 1960's to the 1980's, supporting some kind of peaked-roof pattern, for example one with a peak exposure factor of 3.*

In regard to the assertion that "OEHHA has implicitly decreased the assumed peak exposure level for exposed shop workers," the 1997 draft TSD did not include shop workers in any the calculations done in the document. Only the results of McClellan for the Garshick et al. (1987) case-control study, which were used in the document, included shop workers. Thus, there appears to be no basis for the claim in the first paragraph of the comment. The new draft TSD does not use the McClellan results but instead uses results from the Garshick et al. (1987) case-control study, which exclude shop workers.

One of the difficulties in including the shop workers is illuminated by the issue raised in the last three paragraphs of the comment. The calculation there compares areas under the exposure factor curve for shop workers to that for other exposed workers. That calculation finds that the ratio of the two areas is 0.84, apparently using the assumption that half of the shop workers were unexposed and half were exposed to the tested level. The comment asserts that the ratio implied by the exposure factors used for train workers and shop workers in the 1994 TSD far exceeds unity, again on the assumption that half the shop workers were unexposed and half were exposed to the tested level. Footnote 3 calculates that ratio to be between 1.58 and 3.75. The comment then concludes that the lack of an explanation of this difference "renders OEHHA's 1997 analysis of the Garshick et al. cohort data invalid." How this conclusion follows is not explained. The calculation in the comment does, however, raise a question about TSD's suggested explanation for the Garshick et al. (1988) finding that if shop workers are excluded from the analyses, then relative risks are virtually unchanged. The explanation quoted in the comment has been changed to suggest that one factor may be that a large proportion of the shop workers, perhaps more than half, may not have been exposed to the measured levels of diesel exhaust. This is to address the concern that the exposed shop workers effective cumulative exposure was probably more than twice the train workers effective cumulative exposure. As indicated by Garshick (1991), the problem with the published exposures is that the shopworkers

who worked in the diesel repair shops shared job codes with workers in nondiesel shops where there was no diesel exhaust. This point was emphasized by several investigators at the January, 1996 Scientific Workshop. Unless specific information becomes available on the proportion of shopworkers exposed, it is prudent to follow the suggestions made at the Scientific Workshop to exclude shopworkers from the analysis.

Comment 6: <u>OEHHA's</u> Omission of the Shopworkers from the Analysis Lacks Adequate Foundation. Section 2.1.2a of our report notes OEHHA's omission of the shopworkers from consideration in its 1997 analysis of the Garshick *et al.* cohort data. In Section 7.3.3 of its 1997 analysis (fifth item), OEHHA states that

> *"...it was prudent to exclude [shopworkers] from some of the analyses. This assumption was made because of the great heterogeneity of exposures in the very broad classification of shopworker. Many of the shopworkers were in shops near the engines [sic] sheds, which had the very high exposures when engines were running without modern ventilation systems. Other shopworkers were in facilities not subject to diesel exposures... There does not seem to be any useful information on the proportion of shopworkers in the unexposed or lesser exposed shops (pp. 7-21, emphasis in original)."*

The singling out of shopworkers by OEHHA (1997) as a group whose historical exposure is unacceptably uncertain lacks adequate explanation. Certainly, historical exposure among all groups of workers, including, for example, engineers, brakemen, and so forth, is uncertain. OEHHA does not explain what level of uncertainty is acceptable (hence warranting the inclusion of exposed workers other than shopworkers), and what level of uncertainty is uncertain (hence warranting the exclusion of the shopworkers). Without such criteria, the exclusion must be considered to be arbitrary.

Contemporary data in Woskie *et al.* (1988b) indicate that shopworker exposure to diesel exhaust is relatively **homogenous.** Woskie *et al.* (1988b) divided the railroad worker cohort into five career exposure groups (clerk, signal maintainer, engineer/firer, braker/conductor, and shop), each of which consisted of one or more job groups. For example, "shop" consisted of three job groups - electricians, machinists, and supervisor/laborer/other shopworkers. The Tukey-Kramer multiple comparison test found that although there were no significant differences in exposure within groups ($p < 0.05$), *"A single factor ANOVA model using career group as the explanatory variable for the log of the ARP [Adjusted Respirible Particulate] concentration found career group was a [sic] significant ($p = 0.00001$) in explaining the variations in ARP exposures ($R2 = 0.25$)"* (p. 398). There is one exception: the freight conductor/hostler comparison within the braker/conductor career group yielded statistically significant differences.) In fact, an examination of the shop exposure data in Table I of Woskie *et al.* (1988b) reveals that the coefficient of variation (sample mean -standard error of the mean) for each job group is on the order of 5% to 15%. Moreover, these results are based on a substantial number of observations - 176 shopworkers in all. Mean exposure concentrations among shopworker job groups (125 to 157 μg/m3) substantially exceed levels among the next most highly exposed group, brakers/conductors (83 to 95 μg/m3). Woskie *et al.* (1988b) conclude that, *"Although we*

suspect that there are differences within the career groups, they are small enough that more sampling would be needed to resolve them" (p. 401).

Even if it is assumed that exposure among shopworkers were heterogeneous, and that there is no readily available data quantifying the proportion of shopworkers who were "exposed" vs. "unexposed," OEHHA's decision to drop this group from the analysis remains unwarranted. As shown in Section 2.1.2b of our report, even if it is assumed that half the shopworkers were unexposed (an *"unbiased estimate in the absence of information"* (p. 7-21), according to OEHHA, Part B, 1997), the inclusion of the shopworkers in the dose-response analysis has a substantial effect on the results.

***Response**: Contrary to the language of the comment, the TSD did not omit the shop workers, it excluded them, as suggested at the January 1996 Scientific Workshop. The primary ground given for exclusion was (Section 7.3.2) and is (Section 7.3.4.2) the large uncertainty about the proportion of shop workers exposed to the measured exposure concentration. This point was less clear in the passage quoted from Section 7.3.3, and item 5 therein (now Section 7.3.5) has been revised to clarify that point. The comment notes that the Woskie et al. (1988) data "indicate that shop worker exposure to diesel exhaust is relatively homogeneous." Those measurements were made in engine repair shops. They do not include the unexposed shop workers in different enclosures. The inclusion of shop workers in substantial proportion makes the exposure highly heterogeneous.*

Comment 7: OEHHA's response to this issue (OEHHA, 1997, Part C), as it was raised in our public comments in 1994, does not adequately address this issue. OEHHA responds to this issue on p. C-OEHHA-2 (in response to comment 1) and on p. C-OEHHA- 1 5 (in response to comment 19). The first response (p. C-OEHHA-2) makes several claims:

"As pointed out in the TSD [technical support document (OEHHA, Part B, 1997)], a substantial proportion of the shopworkers was not exposed *to diesel exhaust"* (emphasis in original). The OEHHA (1997) report asserts that shopworker exposure was heterogeneous. As far as we can tell, however, the report never even claims that a "substantial proportion of these workers were unexposed. At the very least, no documentation or analysis is presented supporting this assertion.

***Response:** Garshick (1991) had previously called attention to the effect of dilution of the effect of the shop workers because of those without true exposure. This is now stated in the revised draft at Section 7.3.4.2.*

Comment 8: *"The revised TSD shows that, due to their very heterogeneous exposure, the overall risk of shopworkers is most likely to be about the same as for train workers. "* As discussed in Section 2.1.2b, shopworker exposure is very likely to substantially exceed exposure among other exposed workers (and hence, risk among exposed shopworkers should greatly exceed other exposed workers if diesel exhaust is a carcinogen). In fact, as pointed out in Section 2.1.2b, historical exposure among shopworkers (exposed and unexposed) equals historical exposure among other exposed train workers only if the ratio of historical exposures to contemporary exposures among exposed shopworkers is less than the corresponding ratio for

other exposed train workers. Such a claim is inconsistent with empirical evidence that has been gathered on this issue (specifically, the nitrogen dioxide measurements).

Response: *The revised TSD no longer makes this assertion.*

Comment 9: *"On this basis, the very small reduction of risk with shopworkers removed, as reported in Garshick et al. (1988), is well within the random variation of the data. "* As just demonstrated, there is no basis for this comment.

Response: *The revised TSD no longer makes this assertion.*

Comment 10: *"OEHHA staff were unable to replicate the three-fold and eight-fold increases in estimated risk suggested by the comment. "* Tables 2.1, 2.2a, and 2.2b detail our derivation of a dose-response relationship for exposure measured both in terms of $\mu g/m^3$ and in terms of lung-deposited diesel exhaust. Figures 2-1 and 2-2 demonstrate that the dose-response relationship does not show **a** positive association between diesel exposure and the risk of lung cancer. It is this finding that invalidates OEHHA's conclusions.

Response: *The comment still does not explain the 3-fold and 8-fold assertions.*

Comment 11: OEHHA's second set of responses (p. C-OEHHA- 15) to this issue is likewise inadequate. OEHHA claims: *"As pointed out in the TSD, a substantial proportion of the shopworkers was not exposed to diesel exhaust."* We addressed this point above.

Response: *Garshick (1991) had previously called attention to the effect of dilution of the effect of the shop workers because of those without true exposure. This is now stated in the revised draft at Section 7.3.4.2.*

Comment 12: *"That portion of shopworkers who were very highly exposed undoubtedly did experience a disproportionately high lung burden. However, quantifying that excess is highly problematic and the new version of the TSD (Section 7.3) does not attempt to estimate lung burden in humans. "* Whether or not this is true does not address the comment to which this text responds.

Response: *Comment noted.*

Comment 13: *"OEHHA staff were unable to replicate the three-fold and eight-fold increases in estimated limit risk suggested by the comment.* "We addressed this point above.

Response: *The comment still does not explain the 3-fold and 8-fold assertions.*

Comment 14: *"There is not enough information on the shopworkers to support a conviction of the analysis of other data from the Garshick et al. (1988) cohort. "We* disagree. As discussed earlier, only by making a number of highly implausible assumptions - assumptions that contradict

the only available empirical data (the nitrogen dioxide measurements) - is it possible to dismiss the absence of dose-response relationship made apparent the inclusion of the shopworker data.

Response: *The comment does not establish that the inclusion of the shop workers in a full analysis results in an absence of dose-response relationship. However, our analyses also include a 10-fold assumption which is consistent with the nitrogen dioxide measurements.*

Comment 15*: "New analyses in Appendix E and in Section 7.3 do not include shopworkers in principal results because of the highly variable exposure in that group. "* Again, even if exposure were highly variable, a broad range of assumptions regarding its nature yields a dose-response curve inconsistent with OEHHA's hypothesis that diesel exhaust causes lung cancer in humans.

Response: *The response quoted refers primarily to the uncertainty in the proportion of shop workers. In any case the previous response also applies here.*

Comment 16: Given the importance of the shopworker data, OEHHA's justification for the proposed regulation should demonstrate that historic exposures among this group (including both exposed and unexposed shopworkers) were low enough to conclude that inclusion of these workers in the analysis does not have a statistically significant effect on the calculated dose-response relationship. Short of a quantitative analysis demonstrating this point, it must be concluded that OEHHA's omission of the shopworkers from its analysis is a *post hoc* adjustment. Such an adjustment eliminates important information mitigating against the claim that the Garshick *et al*. cohort data demonstrate a positive dose-response relationship between diesel exhaust exposure and lung cancer.

Response*: Contrary to the language of the comment, the TSD did not omit the shop workers, it excluded them, as suggested at the January 1996 Scientific Workshop. The primary ground given for exclusion was (Section 7.3.2) and is (Section 7.3.4.2) the large uncertainty about the proportion of shop workers exposed to the measured exposure concentration. This point was less clear in the passage quoted from Section 7.3.3, and item 5 therein (now Section 7.3.5) has been revised to clarify that point. The comment notes that the Woskie et al. (1988) data "indicate that shop worker exposure to diesel exhaust is relatively homogeneous." Those measurements were made in engine repair shops. They do not include the unexposed shop workers in different enclosures. The inclusion of shop workers in substantial proportion makes the exposure highly heterogeneous. As indicated by Garshick (1991), the problem with the published exposures is that the shopworkers who worked in the diesel repair shops shared job codes with workers in nondiesel shops where there was not diesel exhaust. Apparent exposure as a shopworker based on job code was then diluted with workers with the same job code, but without true exposure, making it less likely to see an effect in the shopworkers group. This point was emphasized by several investigators at the January 1996 Scientific Workshop.*

Comment 17: <u>OEHHA's</u> Omission of the Cumulative Lung Burden Exposure Lacks Explanation. OEHHA's omission from its 1997 analysis of cumulative lung burden as a measure of exposure lacks adequate explanation, especially given the prominent role that measure of

exposure played in OEHHA's 1994 analysis. As far as we can tell, OEHHA offers only two explanations for this change. In the last paragraph of Section 7.2.3 (OEHHA, Part B, 1997, p. 7-2), OEHHA states that *"lung burden was not used to calculate risk estimates from human study data because the human exposures used in the calculations are not considered great enough to result in sufficient human lung burden to require use of a lung burden model "* If this is the case, then OEHHA should state what levels of exposure are great enough to warrant use of this model. Outside reviews could then better understand the rationale for this change.

OEHHA's second explanation for omission of the lung burden model, which appears in its response to comments document (OEHHA, 1997, Part C), states in response to comment 19 (p. 2-OEHHA-15) that *"quantifying* ...[lu*ng burden] is highly problematic and the new version of the TSD .(Section 7.3) does not attempt to estimate lung burden in humans. "* However, OEHHA does not explain what factors complicate use of the model now, in contrast to its previous application in OEHHA's 1994 report.

As OEHHA's 1994 analysis presented lung burden as a plausible measure of exposure, and, given that this measure of exposure provides an even more striking illustration of the non-monotonicity of the dose-response data based on the Garshick *et al* cohort data (Garshick *et al.*, 1988), OEHHA should provide a coherent rationale for now dropping this measure.

***Response**: The lack of need for the complicated analysis remains the reason for not using it.*

Comment 18: OEHHA's **Analysis of Garshick *et al* (1987) does not Provide an Adequate Basis for Dismissing Smoking as a Confounder.** Although the Garshick *et al* (1988) cohort study did not control for the effect of smoking, OEHHA (1997) claims that this omission is not important because the Garshick *et al.* (1987) case-control study established that failure to control for cigarette smoking does not substantially affect estimates of relative risk. We believe that the failure to control for smoking invalidates inferences drawn from either study.

It is well established that smoking elevates the relative risk of lung cancer. The landmark 1964 Report of the Surgeon General determined that the risk of lung cancer among male smokers exceeded the corresponding risk among non-smokers by factor of approximately 10 (Brandt, 1990). Because smoking has such a substantial influence on the risk of lung cancer, Dr. Frank Speizer, a co-author of the Garshick *et al.* (1988) cohort study, has emphasized the importance of controlling for this factor when investigating other potential risk factors for this disease. Writing in 1986, he states,

> *"Because of the overwhelming effect of cigarette smoking, population-based studies that report on environmental effects, particularly at relatively low levels of excess risk (RR greater than 1.0 but less than 2.0), and that do not attempt to take cigarette smoking into account, must be considered seriously flawed. These studies, therefore, can contribute very little to our understanding of risk factors for respiratory cancer"*
> (Speizer, 1986, p. 9).

By this reasoning, the Garshick *et al.* (1988) cohort study is flawed because it does not control for cigarette smoking and because the relative risks reported are less than 2.0.

OEHHA (1997) attempts to rely on the Garshick *et al.* (1987) case-control study results to justify the validity of the Garshick *et al.* (1988) cohort results, even in the absence of controlling for smoking. However, the results from the case-control study indicate that its statistical control for smoking were inadequate. The Garshick *et al.* (1987) case-control study reports both the crude relative risk of lung cancer among workers with over 20 years of occupational exposure to diesel exhaust (relative risk = 1.39), and the corresponding risk ratio value that has been adjusted for exposure to asbestos and for lifetime smoking, measured in pack-years (relative risk = 1.41). That is, adjusting for smoking and exposure to asbestos had no effect on the relative risk associated with long term diesel exposure.

The minimal difference between crude and adjusted relative risk values would not be unexpected if the prevalence of smoking among both the cases and controls were the same. In fact, members of the case cohort smoked more than members of the control cohort. As detailed in Table 3 of Garshick *et al.* (1987), 80% of the cases smoked, while only 73% of the controls smoked. Among cases 32% were classified as having no more than 50 pack-years of smoking, while 43% were classified as having more than 50 pack-years of smoking. The corresponding figures for the controls were 38% and 30%[5]. (The remaining cases and controls were either non-smokers (9% for cases, 11% for controls) or could not be classified (23 % for cases and 22% for controls).)

As there are more smokers among the cases, failure to control for smoking would be expected to result in a **higher** relative risk than after adjustment, but the reported result goes in the opposite direction. If, as Garshick *et al.* (1987) claim, long term exposure to diesel exhaust is associated with an increased risk of lung cancer, then failure to control for smoking should have increased the estimated "crude" risk ratio for this parameter. Because it did not, there is reason to believe that either long term diesel exposure is not associated with the risk of lung cancer, or Garshick *et al.* (1987) did not adequately control for smoking. In any case, the finding by Garshick *et al.* (1987) that the unadjusted crude risk ratio and the adjusted risk ratio for diesel exposure were nearly identical does not justify OEHHA's lack of concern regarding statistical control of smoking in its 1997 analysis of the Garshick *et al.* (1988) cohort study.

Stober and Abel (1996) outline additional problems with relying on the Garshick *et al.* (1987) case-control study results to address the lack of control for smoking in the Garshick *et al.* (1988) cohort study. They state (p. S-38),

> *"The cohort study by Garshick et al. (1988) has received particular attention because the authors maintained that they had dealt with the question of the effect of the subjects' smoking habits. However, the investigation on smoking habits to which they refer (Garshick et al. 1987b) is very unsatisfactory. Firstly, the proportion of smokers found in this study was unusually high (³ 80%); secondly, it related only to employees more than 50 years old,- and thirdly, the proportion of smokers was derived only from 50 asbestos-exposed railroad workers and their 192 controls, with just a simple differentiation being made between smokers and nonsmokers. In any case, the information on smoking is by no*

means adequate so that, judged by the quality criteria previously established by one of its own co-authors (Speizer, 1986), the study must be regarded as seriously flawed "

***Response**: The Garshick cohort study (1988) does not control for smoking, as is the case for most occupational cohort mortality studies. This does not invalidate the results of the investigation, however. The entire cohort included only blue-collar workers, which diminishes the likelihood of a systematic difference in smoking prevalence between those members of the cohort who were exposed to diesel exhaust versus those who were not. This in turn would make it unlikely that cigarette smoking would significantly confound the results. This notion is corroborated by the results of the case-control study, in which the relative risk (RR) for diesel-associated lung cancer changed little with adjustment for smoking.*

The commenter suggests, however, that the adjustments for smoking in the Garshick et al. (1987) case-control study were insufficient to have confidence in the estimated RR of lung cancer attributed to diesel exhaust exposure. The commenter indicates that the differences in smoking prevalence between the cases and the controls in that study should have resulted in a decline in the RR after adjustment for smoking. Instead, Garshick et al. found little difference between the crude and adjusted RRs. There are several plausible explanations for these results. First, in the Garshick case-control study, information on smoking was obtained from next-of-kin, and was therefore highly likely to have been subject to misclassification. Minor differences in misclassification rates of smoking status or pack-years between the cases and the controls could easily explain why the smoking-adjusted RR did not change much. Moreover, in Table 3 of the Garshick paper referred to by the commenter, there were substantial numbers of subjects for whom there were data missing on smoking status (17% of the cases and 16% of the controls) and on pack-years (23% of the cases and 22% of the controls). Garshick et al. (1987) included a category for missing smoking data in their regression model. The regression coeffcients listed in Garshick's Table 4 indicated that those in the "pack-years missing" category included a number of heavy smokers. Although the distribution of these heavy smokers between cases and controls is not available, it is likely that their inclusion in the analysis could explain the ostensible discrepancy noted by the commenter. In addition, information about the correlations between asbestos exposure and cigarette smoking among the cases and the controls is not available. Given the well-known multiplicative interaction between smoking and asbestos exposure in the induction of lung cancer, a small difference in the proportions of asbestos-exposed smokers among the controls versus the cases could also explain the Garshick results. In additional analyses using only those workers for whom the investigators had detailed smoking data (n = 758), the ORs for 20 yr of diesel exposure ranged from 1.50-1.53, adjusted for asbestos exposure and several specifications of cigarette smoking history. These models included pack-years as a single continuous variable, as two independent variables (cigarettes per day and years of smoking), or as a categorical variable classified in terms of the number of years the study subject had stopped smoking prior to death. Thus, OEHHA staff disagree with the commenter's assertions, and would suggest that the control for smoking in the Garshick case-control study, while not ideal, is adequate to infer that confounding by cigarette smoking is unlikely to explain the finding of an increased RR of lung cancer associated with diesel exhaust exposure.

Comment 19: OEHHA's Meta Analysis Does not Establish a Relationship Between Diesel Exposure and Lung Cancer. Although the meta-analysis conducted by OEHHA includes some 31 studies and 40 estimates of the magnitude of the relative risk of lung cancer associated with exposure to diesel exposure, there are several problems with the analysis that cast doubt on the inference drawn by OEHHA. Perhaps the most important problem compromising OEHHA's meta-analysis is evidence of publication bias combined with the use of statistical methodology that provides more weight to small, imprecise studies. This problem is discussed in Section 2.2.l. Second, without adequate justification, OEHHA dismisses a number of studies yielding low relative risk estimates based on the assertion that the risk estimates from these studies were depressed by the healthy worker effect. We discuss this problem in Section 2.2.2. Finally, in Section 2.2.3, we point out that the meta-analysis fails to provide an estimate of risk associated with changes in atmospheric diesel exhaust concentrations.

Before proceeding, we note several **other reviews** of the epidemiological literature investigating the association between diesel exhaust and lung cancer. Upon review of studies published through June of 1993, Cohen and Higgins (1995) concluded that *"exposure to diesel exhaust in a variety of occupational circumstances is associated with small to moderate relative increases in lung cancer occurrence and/or mortality"* (p. 269). However, even Cohen and Higgins find the epidemiological data inadequate to support a quantitative estimate of risk, stating, *"Although these data provide relative rankings of exposure, the absence of concurrent exposure information is the key factor that limits interpreting the epidemiologic findings and using them to make quantitative estimates of cancer risks"* (p. 6). Other reviewers found that the epidemiological literature does not support even a qualitative association between diesel exhaust exposure and lung cancer. An extensive review of this literature by Stober and Abel (1996) concluded that *"there is no causal relationship between diesel exhaust inhalation and lung cancer"* (p. S-41). They continue, stating that, *"At present.. it can he subsumed from the cohort studies that no definite increase in lung cancer risk from diesel emissions has so far been demonstrated epidemiologically. And there is certainly not any good evidence of a dose-effect relationship "* (p. S-41). Muscat and Wynder reviewed 14 case-control or cohort studies. They state that, *"Using common criteria for determining causal associations, the epidemiologic evidence is insufficient to establish diesel engine exhaust as a human lung carcinogen "* (Muscat and Wynder, 1995, p. 812).

***Response**: The commenter is correct in stating that the meta-analysis does not establish a causal association between diesel exhaust exposure and lung cancer. Causal inference in chronic disease epidemiology involves an assessment of statistical associations, but also requires an evaluation of a variety of other factors as well, including (among others) the consistency of the findings among multiple studies, whether the findings are likely to be due to bias or chance, biological plausibility, and the existence of exposure-response relationships. These and other considerations are discussed at length in section 6.2.4, "Causal inference for diesel exhaust exposure and lung cancer."*

The methodological concerns about the meta-analysis are also addressed below in responses to more detailed comments submitted by the same commenter. Briefly, the OEHHA report acknowledges the existence of some evidence of publication bias, but notes that "Although

publication bias cannot be ruled out, the inclusion of numerous studies of varying sample sizes and statistically insignificant findings makes it unlikely that the results can be completely explained by publication bias." The current version of the report includes an additional sensitivity analysis of potential publication bias, which indicates that the tertile of studies with the smallest standard errors showed a higher pooled risk estimate than the other two tertiles. Although this result suggests the presence of some potential publication bias, it should be noted that the studies in the tertile with the smallest standard errors were almost exclusively cohort studies that did not adjust for smoking and which also had a clear healthy worker effect, suggesting that other significant biases are likely to have played a role in creating an appearance of publication bias. The commenter refers to the "use of statistical methodology that provides more weight to small, imprecise studies" as exaggerating the effect of any publication bias. Whether the use of a random- versus a fixed-effects model exaggerates the potential effect of publication bias is irrelevant, as these estimates tend to converge with diminishing statistical evidence of among-study heterogeneity. Where there is significant evidence of heterogeneity, it is inappropriate to derive pooled estimates of effect. While the pooled risk estimates for all studies combined and for several subgroups showed evidence of serious heterogeneity, this was not the case for several of the subsets (e.g., case-control studies, studies that adjusted for smoking and other potential confounders, studies of truck drivers and of other professional drivers). In these instances, the pooled estimates derived from both random- and fixed-effects models were virtually identical.

OEHHA staff have not "dismissed" or "discounted" the pooled estimate for studies because they were depressed by the healthy worker effect (HWE). OEHHA staff conducted analyses to explore heterogeneity among the cohort studies. The pooled estimates for the subset of cohort studies without clear evidence of a HWE showed little evidence of heterogeneity relative to all cohort studies combined, while those estimates derived from the studies manifesting a HWE still retained substantial heterogeneity. Consequently, the document expresses reservations about the validity of the pooled estimate for these studies because of the problem of heterogeneity.

Finally, as is often the case in occupational and environmental epidemiology, detailed exposure measurements (e.g., workplace or atmospheric concentrations of diesel exhaust or its constituents) for the subjects under study are the exception rather than the rule. For this reason, OEHHA staff had to rely on surrogate measures of exposure as assessed primarily by job titles and duration of employment.

The comment also cites recent literature reviews (Muscat and Wynder, 1995; and Stober and Abel, 1996) in which the authors did not conclude that the existing epidemiological data support the notion of a causal relationship between occupational diesel exhaust exposure and lung cancer. These review articles were reviewed and considered in preparation of the OEHHA document. However, OEHHA staff disagree with the conclusions reached by those authors. With regard to the review by Cohen and Higgins, their qualitative finding of a likely causal relationship for "increases in lung cancer occurrence and/or mortality" in the epidemiology chapter of the HEI review is distinct from the conclusion in the Executive Summary to the report, which reflects conclusions of the HEI Diesel Working Group regarding "quantitative estimates of cancer risks". That is, Cohen's and Higgins' analysis can be distinguished from HEI's

proposed application of their analysis. The OEHHA document is consistent with the findings of Cohen and Higgins with regard to identification of elevated lung cancer rates among diesel exhaust-exposed workers. Another review not cited by the commenter is that of the expert panel assembled by the World Health Organization, whose members concluded that the evidence was consistent with a causal association. The WHO concluded that "The relative risks for lung cancer as a result of exposure to diesel exhaust are generally low, and risks of this magnitude are more susceptible to chance and to the effects of unmeasured confounding factors and imprecision for adjusting for known confounding factors. As discussed above, the elevated risk for lung cancer in the four most informative studies is unlikely to be due to confounding by cigarette smoking and is probably due to exposure to diesel exhaust. Other studies, although limited primarily by the exposure ascertainment, support this assessment." International Programme on Chemical Safety, Diesel Fuel and Exhaust Emissions. Environmental Health Criteria 171. World Health Organization, Geneva 1996. p.252. More recently, Dr. Allan Smith and colleagues at the University of California, Berkeley, conducted a meta-analysis of the literature independent of the effort undertaken by OEHHA staff, reaching conclusions generally similar to those found by OEHHA staff (Reference: Bhatia R, Lopipero P, Smith AH. Diesel exhaust exposure and lung cancer. Epidemiology 1998;9:84-91). In addition, Dr. Smith and his co-workers found the evidence supportive of a causal relationship between increased risks for lung cancer and exposure to diesel exhaust. Ultimately, however, it is the combined weight of the available evidence that persuades, it is not the combined weight of others' reviews of the evidence.

Comment 20: Likely Publication Bias Invalidates OEHHA's Meta-Analysis. Publication bias invalidates OEHHA's meta-analysis. Curiously, OEHHA claims -- with no supporting quantitative analysis -- that the (funnel) graphs provide no evidence of publication bias." The commenter provides a tabular interpretation of the two OEHHA funnel plots in Appendix D, indicating that the central estimates of relative risk are generally larger for the smaller studies in both the case-control and cohort graphs.

The commenter notes that the random effects model "yields that higher pooled estimates of relative risk" [and] places greater weight on smaller studies - the very studies that are more likely to be subject to publication bias..."

***Response**: Publication bias refers to the increased likelihood of publication of statistically significant results compared to nonsignificant or null results, which may potentially distort a pooled risk estimate. Publication bias is generally attributed to journal editorial policies that prefer "positive" results, so that small, statistically insignificant studies are less likely than large, statistically insignificant studies to be published or even submitted for publication. One way to assess graphically whether publication bias is likely to have affected the results of a meta-analysis is to construct funnel plots of the logarithms of the relative risk (log RRs) versus sample size. If there is no publication bias, the plot should resemble an inverted funnel with the apex located approximately over the mean log RR. Publication bias is discussed on pp. D-7, D-10 and D-11 and addressed graphically by Figures D-6 and D-7 of the March 1997 draft.*

Contrary to the commenter's assertion, the OEHHA report does not state that "there is no evidence of publication bias." The relevant text in the OEHHA draft appears below (emphasis added for these responses):

> In order to examine whether the data set used in this meta-analysis could have been affected by publication bias, we created two funnel plots, one for cohort and one for case-control studies.
>
> The funnel plots (Figures D.6 and D.7) revealed no systematic relationship between study size and magnitude of risk indicative of publication bias. The plots for both the cohort and case-control studies appear to show the lower left portion of the plots, representing small, null studies, as modestly sparser than the lower right**.** However, in both sets of study designs, estimates from the smaller studies spanned the range of relative risks.
>
> Concern about publication bias is more acute in random-effects than fixed-effects models, as the former tend to weight studies more evenly. By adding the among-study variances to the within-study variances to derive weights for estimating the pooled RRs, estimates derived from random-effects models may be more sensitive to the effects of large risk estimates derived from small studies. However, many of the studies in this meta-analysis focused on several potential chemical exposures, adverse health outcomes, or occupations with variable diesel-exhaust exposure patterns, so that the diesel-exhaust-to-lung cancer relationship represented only one aspect of these studies. Furthermore, a plot of the RRs versus sample size for each study design revealed no systematic relationship between study size and magnitude of risk, though there is a lower density of studies in the lower left portion of Figures D.6 and D.7, indicating fewer small, statistically insignificant studies**.** In the cohort studies, both the lowest and highest risk estimates occurred in studies with fewer than 700 exposed subjects. For the case-control studies, estimates from the smaller studies involving fewer than 50 exposed cases and controls were spread across the range of RRs. Although publication bias cannot be ruled out, the inclusion of numerous studies of varying sample sizes and statistically insignificant findings makes it unlikely that the results can be completely explained by publication bias**.**

In other words, the OEHHA text does not rule out the possibility of publication bias among the studies purporting to examine the relationship between occupational exposure to diesel exhaust and lung cancer. Whether the use of a random- versus a fixed-effects model exaggerates the potential effect of publication bias is irrelevant, as these estimates tend to converge with diminishing statistical evidence of among-study heterogeneity. Where there is significant evidence of heterogeneity, it is inappropriate to derive pooled estimates of effect. While the pooled risk estimates for all studies combined and for several subgroups showed evidence of serious heterogeneity, this was not the case for several of the subsets (e.g., case-control studies, studies that adjusted for smoking and other potential confounders, studies of truck drivers and of other professional drivers. In these instances, the pooled estimates derived from both random- and fixed-effects models were virtually identical.

Having stated this, in the revised version of the OEHHA meta-analysis, we have modified how publication bias is addressed. First, we have included one funnel plot combining the results for case-control and cohort studies by plotting the inverse of the standard error versus the logarithm of the individual study estimates of relative risk. As Pettiti (1994) notes, such a plot is more informative than having separate graphs for case-control and cohort studies. This plot, as was true of the separate plots by study design, also indicates a relatively less dense distribution of individual study estimates in the lower left corner of the graph. Second, we have addressed the issue of publication bias with a separate sensitivity analysis, examining the effect of pooling progressively larger studies (i.e., with smaller standard errors). Finally, there is a methodological paradox in the commenter's implications about the effects of publication bias. If one assumes that there is no relationship between diesel exhaust exposure and lung cancer, then finding increased risks among some smaller studies must be a result only of random error, even if these were otherwise well conducted studies. The implication is that there should be a roughly equivalent number of null or negative studies that, but for the past editorial policies favoring publication of "positive" studies. However, there are many reasons why specific studies do not get published. Inadequate sample size in a given study decreases the power to detect an association between an exposure and a health effect: this represents a serious design flaw. Thus, in a sense, it is not clear that underpowered studies, particularly those that are not published, should be accorded the same weight as studies that have passed the scrutiny of peer review.

Comment 21: OEHHA's Dismissal of Studies Because of the Alleged Healthy Worker Effect Lacks Foundation. OEHHA (1997) defines the "healthy worker effect" (HWE).as the *"manifestation of selection bias related to hiring and retention of workers who are typically healthier than the general population, resulting in spuriously lower risk estimates for a variety of illnesses, including those potentially related to occupational exposures"* (p. D-8). OEHHA reports that the pooled relative-risk estimate calculated using the fixed-effect model was, as expected, smaller among studies OEHHA labeled as exhibiting the HWE (0.99) than it was among those studies OEHHA labeled as **not** exhibiting this effect (1.49).

Although the HWE may be important in the context of some illnesses, it is unlikely to affect studies of illnesses with long latency periods-like lung cancer. Simply put, the "less healthy" workers-*i.e.*, those that develop cancer-do not develop the illness until late in life and hence will not be excluded from the study cohort by hiring and retention practices. In fact, if employers or employees can predict long-term cancer risks, their diagnostic acumen should be the subject of intense study.

Because it is unlikely that the HEW substantially affects the analysis of the relationship between exposure to diesel exhaust and the development of lung cancer, OEHHA has not demonstrated that those studies it believes is affected by this phenomenon should be discounted. Hence, the pooled estimate of 0.99, referred to above, cannot be easily dismissed, contradicting OEHHA's conclusions.

***Response**: The basis for this comment is unclear to OEHHA staff. The HWE is a composite of several processes resulting in selection bias, including a "healthy hire effect", a healthy worker survivor effect, and the general decline in health occurring with time since hire (Arrighi 1994).*

However, OEHHA staff have not "dismissed" or "discounted" the pooled estimate for such studies because they are likely to produce biased results. Rather, OEHHA staff subset the cohort studies by whether there was clear evidence of a HWE a priori as one approach for exploring heterogeneity among the cohort studies. The pooled estimates for the subset of cohort studies without clear evidence of a HWE showed little evidence of heterogeneity relative to all cohort studies combined, while those estimates derived from the studies manifesting a HWE still retained substantial heterogeneity. In this context, the latter estimates demonstrate that the presence or absence of a clear HWE has some explanatory power with respect to the sources of heterogeneity among the cohort studies.

The March 1997 draft states that the pooled estimates for the HWE group of studies are of "doubtful statistical validity", but this is due to the presence of substantial residual heterogeneity, not to other dimensions of the HWE. We indicated in several places in the meta-analysis that the presence of significant heterogeneity undermines the validity of pooled estimates. Several methodological papers on meta-analysis make the same point (Greenland 199?, Pettiti 1994). The concern about the lack of validity of pooled estimates in the presence of heterogeneity affects many of the other pooled estimates, including those for all studies combined, studies that were not adjusted for smoking, mechanics, heavy equipment operators and dock workers, all North American studies, and all European studies, among others. Thus, contrary to the commenter's assertion, OEHHA staff did not "dismiss" studies because of the "alleged healthy worker effect." Rather, the document expresses a reservation about the validity of the pooled estimate for these studies because of the problem of heterogeneity.

Comment 22: <u>OEHHA's</u> Meta-Analysis Fails to Address the Relationship Between Diesel Exhaust Concentration and the Development of Lung Cancer. OEHHA suggests that because the vast majority of studies used duration of employment to quantify diesel exposure that exposure assessment was adequate. This measure does not distinguish among workers exposed to different concentrations of diesel exhaust. It also fails to rule out some other factor that is coincident with time. Problems related to ignoring differences in exposure concentrations were discussed at length in section 2.1.2b in the context of the Garshick *et al.* (1988) cohort study of U.S. railroad workers. In that case, there was an increased risk of lung cancer associated with time employed in the railroad industry. However, risks did **not** increase as a function of time-averaged exposure concentration. Similar phenomena may have affected the studies included in OEHHA's meta-analysis.

***Response**: As is often the case in occupational and environmental epidemiology, detailed exposure measurements for the subjects under study are the exception rather than the rule. Therefore, in these types of studies, indirect measures or surrogates of exposure are used. As noted in responses to comments by the California Trucking Association, the absence of direct exposure measurements does not invalidate a meta-analysis. Rather, combining studies together that have different measures of exposure can contribute to heterogeneity (i.e., substantial differences among the results of the individual studies that preclude combining them into a summary estimate of effect).*

The OEHHA meta-analysis was based on occupational epidemiological studies involving diesel exhaust in which there were no concurrent industrial hygiene measurements. Thus, the commenter is correct in stating that the meta-analysis does not directly examine the relationship between diesel exhaust concentration and lung cancer: given the lack of industrial hygiene data, this is clearly not possible. OEHHA staff therefore had to rely on surrogate measures of exposure as assessed primarily by job titles and duration of employment. In addition, as noted in section 6 (pp. 6-47-49), several studies in different industries showed evidence of exposure-response relationships, using the surrogates of duration of employment, "possibly vs. probably" exposed categories, "machine time" and several other diesel exhaust exposure variables. Although indirect, these studies do provide some evidence of an exposure-response relationship. In addition, the pooled estimates for the truck driver subgroup in the meta-analysis, showed modest evidence of an exposure-response relationship, where exposure was measured as duration of employment.

Comments on Role of a Genotoxic Mechanism of Action

Comment 23: Extraction of Adsorbed Organic Compounds. In section 5.1.2.6 *Extraction Under Physiological Conditions,* OEHHA discusses those studies investigating the ability of physiological media to remove particle-bound organic material. The Office notes the **failure** of simulated body fluids to remove bound mutagens (Brookes *et al.,* 1981; King *et al.,* 1981; Siak *et al.,* 1981). In other studies, the addition of protein (Clark and Vigil, 1980) or macrophages (King *et al.,* 1983) **decreased** the mutagenic potential of the organic fraction extracted by solvents. OEHHA then discusses the work by Wallace *et al.* (1987) and Keene *et al.* (1991), who demonstrated an increase in mutagenicity after incubation with a phospholipid emulsion. Although Part B acknowledges that the methodology used by Wallace, Keene, and coworkers differs from other similar types of studies, OEHHA emphasizes repeatedly the importance of the Wallace, Keene, and coworker studies, when the Office addresses public comments about bioavailability.

The interpretation of the results from Wallace *et al.* (1987) and Keene *et al.* (1991) is problematic. First, Wallace *et al.,* (1987) used scraped aged, accumulated soot from an exhaust pipe as their source of diesel particulate, which is different from particulate in fresh, airborne diesel exhaust. Second, both Wallace *et al.* (1987) and Keene *et al.* (1991) prepared samples and reported mutagenicity in a peculiar manner. The suspended particles were subjected to sonication and agitation, the effects of which on particle size and surface properties are unknown. In addition, sonication and agitation do not simulate physiological processes. Third, in both investigations, after incubation with the emulsion, the investigators separated the particles from the media and observed that the mutagenicity resided with the particulate fraction and not the filtered supernatant. That is, the emulsion was **not** effective in extracting the organic material off the diesel particles. The authors suggested that the phospholipid emulsion acted to "*solubilize* " the adsorbed components. Because the bioactivity resided in the particle fraction, it is unclear what the authors meant when they used the term "*solubilize* ". The relevance of their test system to the *in vivo* situation remains to be explored and validated, and it is inappropriate for OEHHA to rely prematurely on these studies. For example, if the lungs are not under overload conditions (as OEHHA suggests), and macrophages are not impaired in their ability to take and remove

particles, and organic material is not released from the particles by lung surface fluids, then it is difficult to imagine how lung epithelial cells are at risk of exposure to mutagenic organic compounds.

Response: *Section 5.1.2.6 describes attempts to determine if data from in vitro genotoxicity tests concerning bioavailability of the genotoxic component of diesel exhaust can be generated which would aid in determining if in vivo genotoxicity occurs as a result of exposure to diesel exhaust. Several investigators (Brookes et al., 1981; King et al., 1981; Siak et al., 1981; King et al., 1983) found that extraction of diesel exhaust particulate matter with simulated physiological fluids such as saline, bovine serum albumin, dipalmitoyl lecithin and fetal calf serum resulted in little or no mutagenic activity being present in the extract supernatant after filtration. However, it should be noted that King et al. (1981) also found that excitation and emission fluorescence spectroscopy data indicated that incubation of diesel exhaust particulate matter with both serum and lung cytosol extracted a substantial portion (79 - 85%) of the solvent-extractable mutagens. Although the serum-associated mutagens did not induce significant mutagenicity in Salmonella, incubation of the serum with protease increased the mutagenic activity of the serum, suggesting that the serum-extracted mutagens were bound to proteins and therefore unavailable to bind to Salmonella DNA under the assay conditions used by the authors. Sun et al. (1988) stated that the studies by Brooks et al. (1981) and King et al. (1981, 1983) "suggest that particle-associated organics become "bioavailable" to respiratory tract cells, allowing metabolic processes to occur".*

Additionally, direct exposure of Salmonella to a diesel exhaust stream resulted in mutation induction (Courtois et al., 1993). Finally, diesel exhaust particulate matter suspended in dipalmitoyl lecithin, a major component of pulmonary surfactant, also induced mutations in both Salmonella and mammalian cells (Wallace et al., 1987; Keene et al., 1991; Gu et al., 1992). These studies indicate that solubilization of the genotoxic component of diesel exhaust particulate matter is not required for that component to exert a genotoxic effect in in vitro test systems, and suggests the same for in vivo genotoxicity. In order to clarify the intent of this section, the title of Section 5.1.2.6 has been changed to BIOAVAILABILITY UNDER PHYSIOLOGICAL CONDITIONS.

OEHHA notes that Wallace et al. scraped their soot from the exhaust pipe, thus allowing a different concentration of organic matter on the particles from that of the earlier samples obtained directly from the exhaust stream. However, we also note that a follow-up study by the same group (Keane et al., 1991) demonstrated similar results with either exhaust pipe soot or particles obtained directly from the exhaust stream. Also, the application of sonication and agitation to particulate suspensions used in in vitro genotoxicity tests is a commonly accepted practice in order to achieve a uniform particle suspension.

Comment 24: Bioavailability and Metabolic Activation. In Part C, OEHHA (p. 6, Part C) refers repeatedly to

> *"...data indicating that both animal and human occupational exposure to diesel exhaust has been shown to result in the production of urinary metabolites of PAHs and nitroPAHs, indicating bioavailability of those compounds* "(Chapter 3).

Turning to section 3.4 of Chapter 3, Part B, OEHHA briefly and uncritically presents results from the studies by Kanoh *et al. (1*993) and by Scheeper *et al. (1*994).

Kanoh *et al.* (1993) conducted a short-term rat study to assess the use of urinary 1-hydroxypyrene as a marker of PAH exposure. Rats were exposed to 4.2 mg/m^3 diesel exhaust for 7 hr/day, 5 day/week, for 8 weeks. Urine samples were collected, and 1-hydroxypyrene was measured 2, 4, and 8 weeks after the exposure ended. The pyrene content in the diesel particulate was 36.0 ng/mg. The authors reported an increase in urinary 1-hydroxypyrene levels, peaking at 4 weeks post-exposure. However, the concentration of pyrene contained in the rodents' food was 9.0 ng/mg, and the authors did not properly account for the relative contribution of inhaled and ingested pyrene in the diesel-exposed and sham-exposed animals. First, the authors calculated that the daily dose of pyrene inhaled was 24.77 ng and ingestion was 135 ng. However, for the calculation of inhalation, the authors used airborne concentration of diesel particulate and not the deposition fraction. Therefore, pyrene values for inhalation should be 12% to 20% of 24.77 ng, that is, only 3 to 5 ng. Second, the authors implied that the two groups of rats consumed the same amount of food, but it does not appear that the authors measured food consumption. Mauderly *et al.* (1994) reported lower body weights in rats exposed to 2.2 or 6.0 mg/m^3 diesel exhaust than in sham-exposed rats. In Kanoh's short-term study, it is conceivable that food consumption could have increased in a compensatory manner after particle exposures ended. Without actual measures of food consumption, the authors cannot assume that the particle-exposed rats ingested the same amount of pyrene-containing food as the control rats. Because of the overestimation of inhaled pyrene and possible underestimation ingested pyrene, we disagree that exposure to diesel exhaust was a significant factor in the reported differences. Even if food consumption did not increase, and even if all the pyrene adsorbed to diesel particles were bioavailable, diesel exhaust-derived pyrene only accounted for about 2-3 % of the daily pyrene dose.

Scheeper *et al.* (1994) measured the concentration of urinary 1-aminopyrene in 3 diesel train-engine mechanics and 2 office clerks. Ambient levels of total suspended particulate matter (TSPM) and respirable suspended particulate matter (RSPM) were measured in the repair shop and office. Airborne 1-nitropyrene levels were determined from the collected TSPM. Urine was collected over a 24-hr period on Sunday, Monday, and Tuesday. The authors reported that the cumulative and average excretion of 1-aminopyrene when days are combined (that is, Monday and Tuesday or Sunday, Monday, and Tuesday) were greater in the train mechanics than in the office clerks; however, when the authors compared daily excretion levels on a single-day basis, there were no differences between the two groups of employees. Relating these findings to diesel-engine particulate exposure is problematic. The authors reported that "*a considerable part of the APM* [airborne particulate matter] *is not primarily derived from diesel exhaust."* Furthermore, TSPM and RSPM levels were not consistent with the time and frequency of engine test runs. In addition, in the mechanics, the highest 24-hour average of urinary 1-aminopyrene occurred on Monday, but airborne levels of 1-nitropyrene were not detectable. The authors

provide no information on other sources of nitro-PAHs to which mechanics may have been exposed. The authors did state that this was a preliminary study, and should be treated as such when drawing conclusions about bioavailability.

In section 5.1.2.6 of Chapter 5, Part B, OEHHA mentions the study by Schenker *et al.* (1990), in which urinary mutagenicity was not correlated with exposure to diesel exhaust in 87 railroad workers. The authors obtained measurements of RSP, using personal monitors, and corrected these values for exposure to environmental tobacco smoke. Given the fact that OEHHA appears to be trying to assess the bioavailability of mutagens in section 3.4, Chapter 3, Part B, OEHHA should have also discussed the negative findings of Schenker for railroad workers.

Response: *Although Kanoh et al. (1993) did not provide a quantitative measure of food consumption for the exposure groups, they state that both the diesel exhaust-exposed and control groups ate approximately the same quantity of food. Also, body weight data from carcinogenicity studies indicate that the diesel exhaust exposure either has no significant effect on body weight at exposures of less than 200 days (Heinrich et al., 1986, 1994), or caused less than an approximately 15% decrease in body weight at approximately 50 days of exposure at concentrations of greater than 2.5 mg/m^3 diesel exhaust (Mauderly et al., 1994; Nikula et al., 1995). Exposure duration in this study was only for 56 days. These data indicate that variations in food consumption due to diesel exhaust exposure are unlikely to be responsible for the increased urinary 1-POH levels seen in rats exposed to diesel exhaust.*

Scheeper et al. (1994) did in fact report that "a considerable part of the APM [airborne particulate matter] is not primarily derived from diesel exhaust." However, it would be expected that the ambient non-diesel APM would have the same general composition for both the diesel train-engine mechanics and the office clerks, barring some unobvious additional source for the mechanics. Additionally, corrections were made for the primary possible confounder, environmental tobacco smoke. Finally, although not statistically significant, urinary 1-AP levels obtained from single day urine collections were consistently 1.7 - 2-fold higher for diesel train-engine mechanics compared to office clerks. The single day levels were not significant due to interindividual variability, which is to be expected with n values of 3 and 2 for the diesel train-engine mechanics and office clerks, respectively. Combining the Monday and Tuesday or Sunday and Monday and Tuesday values increased the power of the determination, and was appropriate under the circumstances. These data strongly suggest that 1-AP may be useful as a biomarker of diesel exhaust exposure, and that nitroPAHs contained in diesel exhaust particulate matter may be bioavailable in humans. Section 3.4 (Biomarkers Associated With Diesel Exhaust Exposure) of the document has been changed to reflect this.

Section 3.4 (Biomarkers Associated With Diesel Exhaust Exposure) is intended to describe solely studies describing potential biomarkers of diesel exhaust. Studies describing indications of potential in vivo genotoxicity are beyond the scope of this section. The study by Schenker et al. (1990) which describes attempts to correlate mutagenicity in Salmonella exposed to extracts of urine obtained from diesel exhaust-exposed workers is therefore properly discussed in Section 5.1.2.6. (Bioavailability Under Physiological Conditions).

Comment 25: Presence of DNA Adducts. OEHHA (p. 5, Part B) refers to the presence of lymphocytic DNA adducts in persons occupationally exposed to diesel exhaust as supporting

> *"...the results of epidemiologic studies which describe a positive correlation between human diesel exposure and the induction of lung cancer.*

OEHHA cites the studies by Hemminki *et al.* (1994), Hou *et al.* (1995), and Nielson *et al.* (1996), who investigated DNA adduct levels in peripheral blood cells from healthy, non-smoking males. The subjects were employed as bus garage workers, bus mechanics, or truck terminal workers. It should be noted that the two studies by Hemminki *et al.* (1994) and Hou *et al.* (1995) are on the same workers, who presumably were measured on two occasions. In the first report (Hemminki *et al.*, 1994), the bus garage workers (n = 16), but not the bus mechanics (n = 23), had higher lymphocytic DNA adduct levels than control workers (n = 22); in the second publication, both garage workers and mechanics showed higher DNA adduct levels. *Hprt* mutation frequencies, however, were similar in "exposed" and control persons. Truck terminal workers were evaluated only in the first report and their mean adduct levels were higher than in the control workers. It is very important to stress that exposure to diesel exhaust was presumed and no measurements were taken. In addition, garage workers and mechanics are exposed to diesel fuel during refueling and lubricating oils during engine overhauls; thus, the potential for dermal exposure to PAHs exists and was not taken into account in their analyses. Finally, although data were collected on various social and personal factors, these data were not included in any of the analyses.

Nielsen *et al.* (1996) also examined bus garage workers and bus mechanics. In contrast to the findings of Hemminki, Hou, and coworkers, Nielson observed higher adduct levels in the mechanics than in the garage workers. Again, diesel exhaust concentrations were not measured, so exposure can only be assumed. In fact, the authors state,

> *"Inspection of the working environment gave no indication of significant air pollution from DE, as the garages were very well ventilated and precautions were taken to avoid exposure to engine exhaust "*

Both groups of workers were, however, exposed to lubricating oil. The authors, unlike OEHHA, did not appear to over interpret their findings.

> *"This study demonstrated that bus garage workers and mechanics were exposed to a higher level of genotoxic compounds compared to a nonoccupationally exposed control group ... The source of genotoxins was unclear as well as the route of exposure, but there were indications pointing towards PAH from DE in ambient air and used lubricating oil .. The study indicated that skin absorption of PAH might be an important factor to consider when studying PAH exposure from air pollution sources.*

Response: *The description of the study by Hemminki et al. (1994) in Section 5.4 (Tests Assessing Primary DNA Damage) of the document notes that diesel exhaust exposure levels were not available for the workplaces studied at the time of the study. Lack of availability of direct diesel exhaust particulate matter exposure or an acceptable surrogate measurement would make*

this study unacceptable for a quantitative risk assessment. However, these studies are sufficient for use in contributing to a hazard identification because it is unlikely that the worker groups such as diesel forklift drivers were unexposed to diesel exhaust or that the hospital mechanics and laboratory personnel in the control group were exposed to diesel exhaust. Diesel mechanics and garage workers in the bus maintenance worker group could have potentially been exposed to PAHs through dermal exposure to used lubricating oil and diesel fuel, respectively. The document has been revised to take this into account; these revisions have also been included in the description of the study by Hou et al. (1995). However, it is not apparent that the group having the highest DNA adduct levels (diesel forklift drivers) in the study by Hemminki et al. (1994) had any exposure to either diesel fuel or used lubricating oil. Additionally, both studies indicated increased DNA adduct levels in bus mechanics; the increase was not statistically significant in the study by Hemminki et al. (1994), but was significant in the study by Hou et al. (1995). No difference in hprt mutant frequency was observed between the 47 exposed and 22 control individuals; however, both mutant frequency and adduct level were highest in the 16 most heavily exposed workers, and increased mutant frequency correlated significantly with increased adduct levels.

Nielsen et al. (1996) do state that DNA adduct levels were higher in bus mechanics than in the garage workers; however, the authors did not provide information on whether that difference was statistically significant, or quantitative levels for those two groups as opposed to the combined levels provided in the article. The document has been revised to include these findings. Nielsen et al. (1996) did state that "it appears from the description of working conditions that the garage workers in the Swedish study were more exposed than the individuals in our study", which could well account for mechanics displaying higher levels of DNA adducts than garage workers in their study. The document has been revised to include this information.

To summarize, the data discussed above indicate that workers in occupations which involve maintenance or use of diesel-powered equipment may have increased levels of lymphocyte DNA adducts, that such increases can be demonstrated in workers without likely dermal exposure to either diesel fuel or used lubrication oil, and that lymphocyte hprt mutations may be associated with increased lymphocyte DNA adducts. These data support the results of epidemiologic studies which describe a positive correlation between human diesel exposure and the induction of lung cancer.

Comments on Probability of Threshold of Response

Comment 26: Genotoxic Carcinogens. When justifying the use of linear dose-response models, OEHHA refers to the commonly used practice of assuming the absence of a no-effect level for genotoxic carcinogens. Thus, their assertion that the adsorbed organic compounds are bioavailable and bioactive has important implications for their risk calculations. As noted above, we do not find their cited evidence for *in vivo* genotoxicity persuasive.

***Response**: As noted in Sections 3.4 (Biomarkers Associated With Diesel Exhaust Exposure), 5.4 (Tests Assessing Primary DNA Damage) and 5.5 (Tests Assessing Oxidative DNA Damage), as*

well as the response to Comments 12, 13 and 14 here, data suggest that diesel exhaust is capable of inducing genotoxicity in vivo.

***Comment 27*: Non-genotoxic Carcinogens.** OEHHA cites the analysis by Gaylor and Zheng (1996), which suggests that linear extrapolation is appropriate even for non-genotoxic carcinogens. That is, OEHHA relies on the possibility that non-genotoxic carcinogens do not exhibit a threshold of response (Chapter 6, Part B, OEHHA). OEHHA's dependence on Gaylor's and Zheng's analysis is not justified for several reasons.

Gaylor and Zheng analysis is theoretical and is supported by only one experimental example. The authors use a formula relating tumor incidence to cell kinetic parameters and show that, indeed, "insignificant" (i.e., less than 20%) changes in cell proliferation could result in significant increases in tumor incidence. Although their calculations are internally consistent, they derive from a theoretical model, and should not be used as evidence of linearity for dose-response.

It does not appear that it was the intention of Gaylor and Zheng to validate their model, and they provide only one example to support their theory. The authors cite the studies of Maronpot *et al.,* (1993) and Kociba *et al.* (1978) who evaluated the effects of tetrachlorodibenzo-*p*-dioxin (TCDD) in female Sprague-Dawley rats. According to Gaylor and Zheng, 125 ng TCDD/kg/day did not affect the proliferation of hepatocytes (Maronpot *et al.,* 1993), but 100 ng TCDD/kg/day did increase the incidence of hepatocellular carcinomas (Kociba *et al.,* 1978). However, these results were reported from two different groups of investigators and with such a close dose range (that is, 100 and 125 ng TCDD/kg/day), it would be essential that the study design and experimental methods be identical between the two laboratories before Gaylor and Zheng can conclude that increases in tumor incidence occurred at lower doses than increases in cell proliferation.

Gaylor and Zheng's (1995, p. 221) theory is based on the premise that

> *"a threshold dose is questionable if a nongenotoxic carcinogen acts via a cell receptor. Also, a nongenotoxic carcinogen that increases the cell proliferation rate, via the cell division rate and/or cell removal rate by apoptosis, by augmenting an existing endogenous mechanism is not likely to have a threshold dose.*

However, the authors also state (Gaylor and Zheng, 1995, p. 221),

> *"Nongenotoxic cytotoxic carcinogens that increase cell proliferation rates to replace necrotic cells are likely to have a threshold dose for cytotoxicity below which necrosis and hence, carcinogenesis do not occur. Thus, low dose cancer risk estimates based upon nonthreshold, linear extrapolation are inappropriate for this situation. "*

The current theory for particle-induced tumorigenesis in the rat-inhalation bioassay includes a component of inflammation (Driscoll, 1996). While these inflammatory cells may induce increases in cell proliferation via the production of growth factors and other bioactive components, they also are capable of increasing cell proliferation via cell injury. Inflammatory cells also produce

oxidants, which in turn, form DNA adducts. Furthermore, the presence of adequate quantities of anti-oxidants are protective against oxidant-induced mutations.

Results from Driscoll's laboratory and Mauderly's laboratory demonstrate that a threshold does exist for rat-lung responses linked to particle-induced tumorigenesis. In Mauderly's 1987 diesel exhaust study, which OEHHA used for their unit risk calculations, rats exposed to 3.5 mg/m^3 or 7.0 mg/m^3 diesel exhaust developed lung tumors, but rats exposed to 0.35 mg/m^3 did not. In addition, those animals exposed to the lowest concentration of diesel exhaust did not exhibit any biochemical or cytological changes in their bronchoalveolar lavage fluid or lung tissues (Henderson *et al.,* 1988). The authors concluded (p. 546),

> *"The results suggest that, for the noncarcinogenic health effects reported in this paper, there is a threshold of exposure below which adverse effects were not observed "*

Driscoll and coinvestigators have examined the inflammatory and mutagenic responses of rats exposed to varying concentrations of α-quartz, carbon black, or titanium dioxide (Driscoll *et al.,* 1996, 1997). At particle levels that did not elicit marked inflammation, hprt mutations in epithelial cells did not occur. The investigators (Driscoll *et al.,* 1997, p. 107) concluded,

> *"Specifically, lung doses of non-genotoxic particles that do not produce inflammation, or elicit a degree of inflammation which can be dealt with by lung defenses, may not increase the risk of mutation (and possible lung tumors). That some degree of inflammation may be tolerated without increasing mutation frequency is supported by the results of the present studies ... Overall, these findings indicate that inflammation may play a key role in the in vivo mutagenic effects of particle exposure. Importantly, a role, in whole or in part, for particle-elicited inflammatory cells in the mutagenic effects supports a non-linear relationship between particle exposure and in vivo mutation.*

***Response**: One proposed mechanism for diesel exhaust-induced rat lung tumors is that exposure to diesel exhaust particulate matter at high concentrations exceeds pulmonary clearance capabilities and causes chronic inflammation. This inflammation leads to macrophage and/or neutrophil-induced oxidative DNA damage resulting in mutations which are instrumental in the induction of lung tumors, and also to cell proliferation which may be mechanistically important to the promotion of the rat lung tumors. Several authors (e.g. Driscoll, 1996; Nikula et al., 1997) have hypothesized that this mechanism may have an exposure threshold of action, and tumor induction due to this mechanism would also have a threshold. Gaylor and Zhang (1996) have suggested using the Moolgavkar-Venzon-Knudson clonal expansion carcinogenicity model that small increases in non-necrotic cell proliferation rates **which may be undetectable** may result in significant increases in tumorigenicity, and also state that 1) a nongenotoxic carcinogen that increases the cell proliferation rate via the cell division rate is not likely to have a threshold dose; 2) dose response curves for cell proliferation and tumor incidence do not necessarily mimic each other. These increases in cell proliferation may be effected either by a stimulated increase in cell division or by an inhibition of apoptosis (programmed cell death). The rat diesel exhaust carcinogenicity studies included in this document that have evaluated diesel exhaust-induced lung cell proliferation (Heinrich et al., 1995; Nikula et al., 1995; 1997)*

used an insensitive measure of cell proliferation (histopathological comparison to controls). More appropriate measures for making quantitative comparisons of cell proliferation (e.g. labeling index determinations using bromodeoxyuridine (BrdU) DNA labeling) have not been employed, making it premature to state that a true threshold of diesel exhaust-induced lung cell proliferation has been determined. Also, lung cell necrosis has not been noted in any of the rat diesel exhaust carcinogenicity studies. The studies by Driscoll et al. (1996, 1997) did not study diesel exhaust but rather utilized other insoluble particles (a quartz, carbon black) which, unlike diesel exhaust, have no potentially directly genotoxic component which would have implications for low-dose response and therefore limits their applicability to explaining mechanisms of diesel exhaust-induced rat carcinogenicity. The work of Gaylor and Zheng (1996) is therefore useful in illustrating that cell proliferation, which is one of a number of potential components of the mechanism of rat lung tumor induction by diesel exhaust may not exert a threshold effect on carcinogenicity. This information does not prove that diesel exhaust-induced carcinogenicity exhibits low dose linearity. However, the clonal expansion carcinogenicity modeling data and the diesel exhaust-induced genotoxicity (including oxidative DNA damage) and Ah (dioxin) receptor binding data indicate that diesel exhaust-induced carcinogenicity may exhibit low dose linearity without the existence of a threshold.

Additionally, the rat diesel exhaust lung tumor data discussed in Section 6 of this document is insufficient for the purposes of determining if a threshold for diesel exhaust-induced carcinogenicity exists. As an example, in the study by Mauderly et al. (1987), rats exposed to 350 µg/m^3 diesel exhaust demonstrated an increased but non-statistically significant increase in lung tumor incidence (1.3% compared to 0.8% for controls; relative risk of 1.4). In the comments on this document, Mauderly correctly points out that the problem is a case of sample size. To determine if a difference in lung tumor incidence of that magnitude is significant at a 95% confidence level would require approximately 15,000 animals/group. This suggests that with the data available, a determination that diesel exhaust induces increases in lung tumors at concentrations of less than 2.2 mg/m^3 cannot be made. However, our conclusion is that it also indicates that insufficient data exists for determining that there is a threshold for diesel exhaust-induced rat lung tumors.

Finally, it should be noted that although calculations of human cancer risk using rat lung tumor data have been included in the document on an informational basis, those calculations are no longer being included in the final range of human cancer risk estimates since human cancer risk estimations based on human data are available.

Comments on Extrapolation from Rats-to-Humans

Comment 28: The Rat as an Outlier. The lung tumor response in rats is not particle specific. The development of lung tumors in rats exposed to diesel exhaust is no different than the response observed in rats after lifetime lung overburdening with other particulates such as carbon black, titanium dioxide, talc, iron oxide (rust), and volcanic ash (see Mauderly and McCunney, 1994). That is, laboratory rats respond to the lifetime lung overburden of particulate, and not specifically to diesel exhaust.

Because only rats, and not mice and hamsters, develop lung tumors after chronic inhalation of high levels of a variety of insoluble particles, most scientists are of the opinion that rats have an anomalous response to inhaled, insoluble particles of any kind. Contrary to OEHHA's statement that the results in mice are *"mixed, "* Dr. Heinrich and Dr. Stober, both premier researchers in this field, came to dramatically different conclusions in their comments on OEHHA's 1994 draft. Specifically, Dr. Heinrich states (p. 67, Part C, OEHHA):

> *"Recent thorough studies have shown no increase in lung tumor incidence in two strains of mice. The current conclusion, therefore, is that diesel exhaust does not cause significant elevation of lung tumors in NMRI and C57BL/6N mice.*

Dr. Stober states (p. 145, Part C, OEHHA):

> *["The IARC analysis] removes any significance from the mouse studies.*

Finally, Dr. Mauderly, another eminent researcher in the particle inhalation area, has published, with his colleagues, a thorough study of diesel exhaust exposure in CD-1 mice, which concludes (Mauderly *et al.,* 1996):

> *"The lack of an exposure-related increase of primary lung neoplasm among CD-1 mice exposed chronically to diesel exhaust contrasts with the significant increase observed in F344 rats exposed concomitantly using the same methods and concentrations.*

Thus, OEHHA's characterization of the mouse results as *"mixed"* contrasts sharply with the accepted opinions of the research community.

OEHHA does not adequately address the fact that rats have an anomalous response to the inhalation of particles. The sequelae of particle retention in the rat lung are exaggerated in that species, and, consequently, the lung tumors that develop are not relevant to other species of animals or to humans. Specifically, when rats inhale high levels of particles over extended periods of time, the following mechanisms come into play:

- Lung overload (lung overburdening of particles, resulting in reduced rates of clearance for deposited particles).

- Exaggerated influx of inflammatory cells (both macrophages and neutrophils) into the lungs. Rat neutrophils, in and of themselves, have been shown to be tumorigenic in rat lungs.

- Inadequate levels of lung antioxidants (diminished levels of oxygen free-radical scavengers).

- Alveolar Type II cell epithelial hyperplasia.

Because these observations coincide with lung tumor development in rats, toxicologists propose that they relate mechanistically to rat lung tumorigenesis. Research is currently underway to elucidate the reasons behind the peculiar response in rats. In humans, even at large lung burdens, we do not know, if in fact, alveolar clearance is impaired. In humans occupationally exposed to inert particles, there is no evidence indicating an exaggerated influx of neutrophils or alveolar type II cell hyperplasia (Watson and Valberg, 1996). Most importantly, none of these steps have been shown to be a consequence of diesel-exhaust particle inhalation in humans.

In concordance with these differences in biologic response between rats and humans, the epidemiology of workers exposed to inhaled, insoluble particles has not identified an excess lung cancer risk. Workers occupationally exposed to carbon black, either in its manufacture or use, have been evaluated in a number of epidemiologic studies. Because historical exposures to carbon black in the work environment were known to be elevated, study of these workers provides a good test of possible increases in lung cancer risk. Yet, data from the carbon-black manufacturing industries in the US and the UK do not establish an excess risk of lung cancer (Valberg and Watson, 1996).

Inhalation of insoluble, low-toxicity particulates by other occupational groups has not resulted in excess cancer risk. These groups include TiO_2 workers (Chen and Fayerweather, 1988), workers exposed to nonasbestiform talc (Wergeland *et al.*, 1990), workers exposed to iron oxide (Stokinger, 1984), and coalworkers (National Institute for Occupational Safety and Health, 1986; Mauderly, 1994; IARC, 1997). Coalworkers, in particular, in earlier times are known to have accumulated large burdens of lung-retained coal particles (which are primarily composed of carbon), yet this worker population does not exhibit excess lung cancer risk. IARC recently evaluated inhaled coal dust as ranked it as "Group 3" (unclassifiable as to carcinogenicity in humans due to a lack of evidence of carcinogenicity from either animal or human studies).

Contrary to the position of CalEPA/OEHHA, regulatory bodies in the U.S. have recognized the anomalous response of rats to the inhalation of large quantities of insoluble particles.

The Clean Air Scientific Advisory Committee (CASAC) is a peer-review group for U.S. EPA composed of experts in inhalation toxicology. CASAC (1995) determined that the response of rats to inhaled diesel exhaust particles is **not** useful for U-S-EPA in developing cancer unit risks (that is, for human health risk assessment). CASAC (1995) states:

> *"The cancer-causing mechanism in the rat may be unique to the rat and does not appear to occur in other species including humans. The mechanism in rats is apparently related to particulate overload followed by a sequence of events beginning with inflammation and ending in tumorigenesis. These events are conditional upon particle overload which also occurs in rats exposed to high concentrations of inert dusts as well. Consequently, it appears that these studies are not relevant for human risk assessment."*

In another context, the U.S. EPA (U.S. EPA, 1988) addressed the fact that chronic inhalation of TiO_2 results in lung tumors in rats. U.S. EPA delisted TiO_2, from the toxics release inventory, as a lung carcinogen because,

> *"in the rat bioassay, the dose levels of TiO_2 used overwhelmed the normal clearance mechanisms of the lungs "* and *"the overall weight-of-evidence determination shows there is not sufficient evidence to reasonably anticipate that TiO_2 will cause cancer in humans.*
>
> The Presidential/Congressional Commission on Risk Assessment and Risk Management (CCRARM) expressly identified in a 1996 draft document (and again in the 1997 final report) some of the mechanisms and substances that are **not** predictors of human health effects. CCRARM (1997) singled out the response of rat lungs to inhaled particulates (giving carbon black and TiO_2 as examples) for which inhalation studies, positive for lung tumors in rats, are not likely to be predictive of human cancer risk.

In spite of this wide-spread opinion that the rat inhalation bioassay for insoluble particulate is not appropriate for risk assessment purposes, OEHHA persists with using rat data for their unit risk calculations. Although we do not endorse the use of rat data in such a calculation, we have noted some problems and errors in Chapter 7 of OEHHA, Part B (1997). Specifically,

- 7-1: Particle concentration is selected because it is a *"commonly used measure.* This is inadequate justification. OEHHA needs to carefully discuss the implications of their choice of dose metric. That is, if lung cancer risk in humans cannot be based lung overload mechanisms, then (OEHHA would say) adsorbed organic content is more relevant than the mass of diesel particulate. This choice would likely require that all of the organic material were bioavailable, which is highly unlikely. Furthermore, if the quantity of adsorbed organics is the relevant dose parameter, then the fuel type and operating conditions become very important and cannot be dismissed by the Office. For example, the locomotive fuel and operation of diesel train engines characteristic of worker exposure in the Garshick *et al.* studies are not equivalent to the fuel type and diesel exposure conditions for the California population today.

- For the animal studies, *"the lung burden dose measure was assumed on theoretical grounds to he a better predictor of tumorigenicity." Yet, "lung burden was not used to calculate risk estimates from human study data"* because exposures were not great enough. In OEHHA's 1994 draft, lung burden in humans was described as a more relevant exposure metric, and the reason for the change appears to be ad hoc.

- It is stated that *"lung burden estimates were derived from the model of Yu and associates (1991) "*, yet on Table 7-1, pp. 7-30, Hattis and Silver (1992) are given as the source of the lung burden model.

- The same paragraph that says *"rat data are consistent with ... risk estimates of 16 to 160 cases per million "* also states that *"risk estimates ... differed by less than five-fold. "* Is this discrepancy fall within OEHHA's rubric of *"relatively consistent"?*

- What is the effect of *"censoring of such observations due to any deaths in which lung tumors were not detected"?*

- Are computations that *"gave the value zero for the latency period"* consistent with biology?

- It is stated that the model-derived q_1 refers to 35 hr/wk of rat exposure. Hence, an additional correction is applied to derive a q_1 for continuous human lifetime exposure. However, on the top of p. 7-5 it is made clear that the rat doses that were entered into the model were *"lifetime-mg/m^3*. " Thus, it would seem that the intermittent rat exposure was doubly (and redundantly) corrected.

- 7-8: Perhaps because of OEHHA's (p. 7-7) redundant correction, Table 7.3 and Table 7.7 report different values for what would seem to be the same result. On Table 7.3, the *"95% UCL for human unit cancer risk, based on concentration, "* and developed from the Mauderly study is **9x10^{-5}.** Whereas on Table 7.7i the *"Human 95% UCL for unit risk for diesel exhaust",* based on concentration, and predicted from the Mauderly rat data is **28xl0^{-5}.** This three-fold difference does not give confidence about the precision of OEHHA's modeling procedures.

- 7-12: The calculation comparing lung burdens at the bottom of the p. is incorrect. On the top of p. 7-13, it is stated that the ratio of human lung burden per alveolar surface area, at ambient diesel concentrations, is only 500 times less than the rat lung burden per alveolar surface area for the test chamber atmospheres. Yet, the rat lung burden is 6,220 μg/0.4 m^2 = 15,600 μg/m^2, and the human lung burden is 4.2; μg/135 m^2 = 0.031 μg/m^2. **Thus, the difference is a factor of 500,000, not 500.** Again, this loss of a factor of 1,000 does not give confidence in the modeling procedures.

It is important to stress that correcting these errors will not compensate for the inappropriate use of the rat data to begin with.

***Response**: Sections 6 and 7 of this document state that the mechanism of action by which diesel exhaust induces lung tumors in rats is not established. One proposed mechanism is that exposure to diesel exhaust particulate matter at high concentrations exceeds pulmonary clearance capabilities and causes chronic inflammation. This inflammation leads to macrophage and/or neutrophil-induced oxidative DNA damage resulting in mutations which are instrumental in the induction of lung tumors, and also to cell proliferation which may be mechanistically important to the promotion of the rat lung tumors. This mechanism has also been invoked for carcinogenicity caused by other insoluble particles (e.g. carbon black, titanium dioxide). Rat lung tumor induction due to high dose (2.2 mg/m^3 or higher) exposure to diesel exhaust may share some commonality of mechanism with other carcinogenic insoluble particles; this possibility is discussed in the document. However, genotoxicity due to 1) the PAH and nitroPAH content of diesel exhaust, and 2) possible oxidative DNA damage primarily due to diesel exhaust exposure may play a role in the induction of lung tumors in rats at lower levels of diesel exhaust. A recent report by Borm et al. (1997) indicates that incubating rat lung epithelial-derived cells with human polymorphonuclear lymphocytes (PMN) (either unactivated or activated by preexposure to phorbol myristate acetate) increases DNA adduct formation caused by exposure to benzo[a]pyrene; addition of more activated PMN in relation to the*

number of lung cells further increased adduct formation in a dose-dependent manner. The authors suggest that "an inflammatory response in the lung may increase the biologically effective dose of polycyclic aromatic hydrocarbons (PAHs), and may be relevant to data interpretation and risk assessment of PAH-containing particulates." These data raise the possibility that low dose diesel exhaust exposure may result in levels of neutrophil influx which would not necessarily be detectable via histopathological examination as acute inflammation but which might be effective at amplifying any potential diesel exhaust genotoxic effect. WHO (1996) has noted that modeling of human cancer risk from rat lung tumor data should take into account the effects of both particles (carbon core) and extractable organic matter (PAHs, nitro PAHs).

Additionally, some parameters of the "particle overload" hypothesis are incompletely characterized. No data exists on the claimed inadequacy of rat lung antioxidant levels. Alveolar type II cell epithelial hyperplasia has been noted after diesel exhaust exposure, but the measures of cell proliferation used were relatively crude and unsuitable for use in a quantitative estimate of cell proliferation as would be required for biologically-based modeling. It should also be noted that uncertainties exist regarding the magnitude and biological importance of particle overload for diesel exhaust-induced rat lung carcinogenicity. Hattis and Silver (1994) examined lung burden data from diesel exhaust rat carcinogenicity studies and came to the conclusion that "there is continuing accumulation of diesel-derived dust in the lungs of rats throughout life, even at low doses". They also found that this was not predicted by models developed to represent diesel exhaust particulate matter accumulation under "overload" versus nonoverload conditions. Finally, they have found that at high diesel exhaust exposure levels, the increase in the ration of internal diesel exhaust particulate matter burden to external exposure is not very large, being slightly larger than a factor of 2 at most, and state that "Although dust overloading is a real phenomenon, it is not a very large effect and thus would not be expected to give rise to dramatically lowered estimates of risk at low exposure levels." It is therefore premature to conclude that the carcinogenic response in rats to diesel exhaust is completely nonspecific.

Section 6.1.1.1 of this document lists two studies (Pepelko and Peirano, 1983; Heinrich et al., 1986a) in which the authors describe statistically significant lung tumor induction in mice in response to diesel exhaust exposure. Another study (Takemoto et al., 1986) is also listed in which the authors reported increased tumor incidences which were not statistically significant, but which were later reported by IARC (1990) to be statistically significant after the data were reanalyzed. Section 6.1.3 has been changed to include a study by Ichinose et al. (1997) which describes lung tumor induction in mice after intratracheal exposure to diesel exhaust particulate matter. Of the six mouse diesel exhaust inhalation or intratracheal instillation carcinogenicity studies described in this document, three definitely contain positive results, one probably contains positive results, and two are negative. This indicates that the description of the mouse carcinogenicity results as "mixed" is justified.

The rat lung has been a relatively good predictor of known human lung carcinogens (e.g., beryllium compounds, cadmium compounds, chromium VI compounds, nickel compounds, asbestos, crystalline silica, radon, tobacco smoke, coal tar, bis(chloromethyl)ether). Based on bioassay data, the rat appears to be a better predictor of known human lung carcinogens than

the mouse or hamster, especially for particulate carcinogens. For example, some cadmium and nickel compounds which test positive in the rat show mixed results in mice and are negative in hamsters. Certain forms of asbestos are positive in rat bioassays but negative or mixed in mouse and hamster. Likewise, crystalline silica is positive in rat bioassays but negative in mice. It is not unusual to see different results of carcinogenicity bioassays in different rodents or even in different strains of the same rodent. Thus, the lack of positive results in hamsters and mixed results in mice does not of itself preclude extrapolation of results in rats to humans.

In respect to the specific points critiquing Chapter 7, found in the last portion of the comment:

1) Section 7.1 has been amended concerning the rationale for the choice of particle mass as the measure of diesel exhaust for the risk assessment. The evidence is not clear about the mechanism of causation, and the particle measure may be regarded as the best available surrogate for whatever is associated with the particles that may be causing the lung tumors. That section now concludes: "On the basis of its relation to health studies and its general practicality, that measure is used in this risk assessment."

2) With the exclusion of shopworkers the lung burden in humans appears to be approximately linear in exposure. For this situation it is very convenient to use exposure rather than lung burden, when possible, as in humans with exposures of 50 mg/m^3 or less. The animal exposures that produce tumors are sufficiently high that the use of lung burden is likely to be a better predictor of tumorigenicity. In the 1994 draft the shop workers were included in the analysis; so lung burden in humans was used in that draft.

3) In the 1997 draft at Section 7.2.4.1, the statement about the use of the model of Yu and associates applies only to the comparative study described in that section, the data for which are given in Table 7-2 and the results in Table 7-3. The methods used there followed the US EPA. The statement about using the method of Hattis and Silver in Table 7-1 applies only to the study of Mauderly et al., the data for which is described in that table. The analyses of the comparative study were done more simply than the full time-to-tumor analysis of Mauderly et al.

4) In Section 7.2.4.1 of the 1997 draft, the paragraph intended to describe consistency of the ranges of UCLs apparently did not succeed and accordingly has been rewritten as follows: "The 10-fold range of UCLs obtained for the comparative study, 16 to 160 cases per million persons exposed to 1 µg/m^3 diesel particulate matter over a lifetime, overlaps in its upper end the 5-fold range obtained for the analysis of the Mauderly et al. (1987) study, 54 to 280. When narrowed to consideration of the results for lung burden only, the 5-fold range for the comparative study, 16 to 77 is very close to the 4-fold range for Mauderly et al., 15-54."

5) Censoring prevents the full lifetime observation of some individuals (Lawless, 1982, p. 3). So in this circumstance more individuals would have been counted with tumors if they had not been removed from the study prior to the end of their lifetime. Thus, the censoring would reduce the estimates of risk if not accounted for.

6) The 1997 draft has been revised to make it clear that latency period refers here to detection lag, the time from carcinogenesis to death. We do not have data to confirm the model result that this time is negligible on the scale of the rat's lifetime.

7) The statement of the underlying model in terms of lifetime-mg/m^3 does not imply that the calculations were all carried out with those units. The TSD follows a common practice of using convenient units throughout the calculation and then adjusting the result appropriately.

8) The reason for obtaining two different values for unit risk for the Mauderly et al. (1987) data is that the full calculation used a time-to-tumor approach and the comparative calculation used a "quantal" (end of life) approach (ICF Kaiser, 1993). Even though the time-to-tumor approach is expected to be more accurate, the quantal approach allows comparison among studies when time-to-tumor data are not available. Also, reporting results for different models on the same study gives an idea of the effect of the alternative models.

9) The staff apologizes for three values in Point (6) being incorrect by a factor of 1000. Each was once expressed in milligrams for convenience, but for clarity the three quantities were meant to have been re-expressed in micrograms, as indicated: (1) When converting from milligrams, 15.5 μg/m^2 *should have read 15.5 x 10^3* μg/m^2*, as the comment correctly pointed out. (2) Likewise, 4.2 mg should have read 4.2 x 10^3 mg. (3) Likewise, 3.1 x 10^{-2} mg/m^2 should have read 31 mg/m^2 The comment corrected the first of these errors, but not the second, thus obtaining the extra factor of 1000. Here is the whole calculation with also a minor correction:*
Rat lung burden/area = 6,220 mg/0.4 m^2 = 15,600 mg/m^2, just as in the comment.
But, human lung burden = 3.7 mg/m^3 x 1240 m^3 = 4.6 x 10^3 mg.
Then, human lung burden/area =4.6 x 10^3 mg/m^2 /135 m^2 = 34 mg.
So the ratio is 15,600 mg/m^2/34 mg/m^2 = 460, a little less than given in the 1997 draft.
The text has been corrected.

Comment on Use of Lung Cancer Risk in Diesel-Exposed Rats to Predict Lung Cancer Deaths in Carbon Black Workers

Comment 29: The production of lung tumors in rats after chronic particle inhalation is not predictive of human risk, particularly in the case of workers in carbon-black manufacturing, who do not exhibit an excess lung cancer risk (Valberg and Watson, 1996). OEHHA discounts this observation by referring to *"a new Canadian study showing carcinogenicity in humans exposed to carbon black. "* (p. 7-26, Part B, OEHHA). However, the referred-to publication (Parent *et al.,* 1996) did not report on workers in carbon black manufacturing; the study population had only presumptive exposure to carbon black, and in fact, experienced exposures to many other substances. OEHHA elaborates on this statement in Appendix C, and attempts to question this lack of concordance between rats and humans by citing Parent's case-control study (1996). In addition, the Office presents alternative assumptions for the analysis by Valberg and Watson. OEHHA's discussions in Appendix C are flawed on two counts:

- The cited study does not show what OEHHA attributes to it.

- OEHHA's criticisms of the lack of concordance between rat and human responses to carbon black are inaccurate.

5.2.1 Shortcomings of the Parent *et al* (1996) Study. The Parent *et al., (1*996) data were derived from a Canadian population-based epidemiologic study that evaluated various health indices and occupations, some of which had presumptive exposure to carbon black (Siemiatycki *et al.*, 1991). Parent *et al.* (1996) conducted additional analyses of the lung-cancer cases identified in Siemiatycki's population-based study. Because of the attention drawn to this study by OEHHA, it is important to clarify the degree to which this study can be relied upon for carbon-black, lung-cancer risk assessment.

In Siemiatycki's data base (1991), patients were interviewed to obtain information on work history. A team of hygienists and chemists then assigned possible occupational exposures to the various job categories. Exposure to carbon black was judged to occur only in user industries, such as the painting, printing, and rubber industries (the population did not include any workers in carbon black manufacturing). After adjusting for various factors, the investigators calculated a total of eight odds ratios (ORs) as follows: two target populations *("all workers", "French-Canadians only")* were compared to two control groups *("general population ", "cancer controls")* over two exposure categories *("any", "substantial").* For presumptive carbon-black exposure, two of the eight ORs were reported to be statistically significant at the $p = 0.10$ level. When compared to cancer controls, *"all workers"* having "any" *or "substantial"* exposure history to carbon black experienced an increased lung-cancer risk; all *any* other ORs were not significant.

The major shortcoming of Siemiatycki's data base is exposure assessment. First, actual exposure to airborne carbon black was not documented; exposure was only inferred from patient interview and the assumption of exposure from job descriptions. Second, those workers with an increased risk of lung cancer were primarily in the **printing** and **publishing** industries, which involve exposure to known or suspect organic carcinogens but with no demonstrated exposure to pure airborne carbon black. For example, Table I from Siemiatycki's monograph (1991) notes that workers assigned to exposure to carbon black were also exposed to numerous other compounds, several of which were also associated with an increased risk of lung cancer (see Table 7). However, the authors did not control for confounding due to simultaneous occupational exposures.

In the Parent *et al.* (1996) analyses, lung-cancer patients were matched with population controls and cancer controls. Again, presumed exposure to carbon black was assigned by a team of chemists and hygienists. The authors constructed a cumulative exposure index using variables for concentration, frequency, confidence of exposure, and duration of job. They categorized cumulative exposure as either "*unexposed", "lower", or "higher",* depending on the numerical value of the index. Of the entire study population, only 5.3% were assigned some exposure to carbon black. The majority of such exposure occurred in painters and paperhangers (26%), printing press operators (12%), and motor vehicle mechanics and repairers (8%). Although carbon black was considered present in these occupations, the authors did not distinguish between inhalation, accidental ingestion, or dermal contact. In addition, the chemical and physical form of

the carbon black was not characterized. There were substantial differences, including smoking status, between the lung-cancer cases and the two control groups. The authors adjusted for some of these factors in their analyses. The authors reported that *"some increase in risk for all lung cancers was apparent."* However, Table V from their paper shows that for seven out of eight ORs, the lower 95% CI were less than 1.0, indicating non-significance of any elevated OR's. The authors presented no statistical tests to determine significance levels. Also, because there was a dramatic difference in smoking status between the lung-cancer patients and the two control groups, it is unclear why the crude and adjusted ORs are so similar; they should have been very different.

In summary, Siemiatycki's monograph (1991) reported that he had found a significant increase in lung-cancer ORs for workers with a presumed exposure to carbon black, but no specific carbon-black exposure data were available, no dose-response could be demonstrated, and the workers was also exposed to other potential lung carcinogens, which were not controlled for in the analyses. Furthermore, of eight different ORs, only two were statistically significant at $p = 0.10$. Parent's analyses (1996) also failed to obtain adequate exposure information, and significance was not supported by proper statistical tests. It is not possible to conclude that the results from this study showed any an association between airborne carbon black exposure and lung-cancer risk.

Thus, it is clear that the probative value of the Parent *et al*. (1996) work as a comment on risk of carbon black exposure is poor at best. In fact, the authors acknowledged that exposure to carbon black was probably minor compared to exposures to other substances.

5.2.2. Inaccuracies in <u>OEHHA's</u> criticisms of the Valberg and Watson (1996) analysis. In Appendix C, OEHHA undertakes a complex series of analyses based on alternative assumptions. While some of these different assumptions represent acceptable fine tuning, OEHHA has lost sight of two simple points. One, OEHHA's complex analysis obscures the basic definition of unit. risk. The definition of the unit risk value is that, when it is multiplied by the lifetime average concentration, the result is the lifetime cancer risk. Hence the calculations presented in the Valberg and Watson (1996) paper are correct, notwithstanding OEHHA's efforts to minimize the lack of concordance between rat predictions and human experience in the case of carbon black. It should be emphasized that the same lack of concordance exists in the case of coal dust, where rats have been shown to develop lung tumors after coal dust inhalation, yet the extensive record of coal miners with heavy lung burdens of retained coal dust does not reveal excess lung cancer risk.

Two, OEHHA in its own document, claims to detect an excess lung cancer risk in the Garshick *et al*. (1988) cohort data for railroad workers. The average lifetime concentration of diesel exhaust is given by OEHHA as 64 $\mu g/m^3$ (p. 7-21, Part B, OEHHA). **If** the rat bioassay were valid, and because it predicts equivalent carcinogenicity for diesel exhaust and carbon black, the "cancer signal" in carbon black manufacturing workers should be in proportion to their lifetime exposure. Because the average lifetime concentration of carbon black for the historical cohorts studied was approximately 410 $\mu g/m^3$ (Valberg and Watson, 1996), the hypothetical lung cancer risks for carbon black workers should be **6.4**x **larger** than for Garshick *et al.'s* railroad workers. How could such an excess have been missed? The answer is that the particle-inhalation bioassay in rats is not applicable to predicting human cancer risk.

As pointed out in Section 5.1, worker lifetime exposure to significantly elevated levels of carbonaceous particles (carbon black, coal dust) have not been associated with increased lung cancer risk. Carbon black particles are very similar in size and composition to diesel exhaust, and even though coal dust is larger in size, lifetime inhalation produces significant lung retention of the find particle fraction of coal dust. The fact that human lung reactions to these two particles are so dissimilar to the rat lung response severely undermines the utility of the rat for diesel exhaust risk assessment.

5.3. **Conclusion.** The lung tumor data from rats chronically exposed to high levels of diesel exhaust should not be used for estimating lung cancer unit risk. For insoluble particles, the lung tumor response is rat specific and not particle specific. Other rodents exposed to insoluble particles do not develop lung tumors, and epidemiologic studies of workers exposed to insoluble particles do not report an excess of lung cancer.

***Response**: Appendix C is not meant to be a comprehensive treatise on carbon black and lung cancer. Furthermore, it is not germane to the derivation of the unit risk factor for diesel exhaust. We have removed appendix C from the document and placed it into Part C as it was composed in response to a comment on the 1994 diesel exhaust document. In addition, we are not relying on the animal data to quantitatively evaluate diesel exhaust as a human carcinogen.*

OEHHA has decided following peer review including suggestions by the Scientific Review Panel to provide calculations of human cancer risk using rat lung tumor data on an informational basis, but not include those calculations in the final range of human cancer risk estimates since human cancer risk estimations based on human data are available. We agree that the effects of particle overload impacts on the validity of extrapolation of the rat lung tumor data to humans. However, we do not agree with the argument that the induction of rat lung tumors by diesel exhaust is not particle specific.

Sections 6 and 7 of this document state that the mechanism of action by which diesel exhaust induces lung tumors in rats is not established. One proposed mechanism for diesel exhaust-induced rat lung tumors is that exposure to diesel exhaust particulate matter at high concentrations exceeds pulmonary clearance capabilities and causes chronic inflammation. This inflammation leads to macrophage and/or neutrophil-induced oxidative DNA damage resulting in mutations which are instrumental in the induction of lung tumors, and also to cell proliferation which may be mechanistically important to the promotion of the rat lung tumors. This mechanism has also been invoked for carcinogenicity caused by other insoluble particles (e.g. carbon black, titanium dioxide). Rat lung tumor induction due to high dose (2.2 mg/m^3 or higher) exposure to diesel exhaust may share some commonality of mechanism with other carcinogenic insoluble particles; this possibility is discussed in the document. Several authors (e.g. Driscoll, 1996; Nikula et al., 1997) have hypothesized that this mechanism may have an exposure threshold of action, and tumor induction due to this mechanism would also have a threshold. Gaylor and Zhang (1996) have suggested using the Moolgavkar-Venzon-Knudson clonal expansion carcinogenicity model that small increases in non-necrotic cell proliferation rates which may be undetectable may result in significant increases in tumorigenicity. They also

state that 1) a nongenotoxic carcinogen that increases the cell proliferation rate via the cell division rate is not likely to have a threshold dose; 2) dose response curves for cell proliferation and tumor incidence do not necessarily mimic each other. These increases in cell proliferation may be effected either by a stimulated increase in cell division or by an inhibition of apoptosis (programmed cell death).

The rat diesel exhaust carcinogenicity studies included in this document that have evaluated diesel exhaust-induced lung cell proliferation (Heinrich et al., 1995; Nikula et al., 1995; 1997) used an insensitive measure of cell proliferation (histopathological comparison to controls). More appropriate measures for making quantitative comparisons of cell proliferation (e.g. labeling index determinations using bromodeoxyuridine (BrdU) DNA labeling) have not been employed, making it premature to state that a true threshold of diesel exhaust-induced lung cell proliferation has been determined. Also, lung cell necrosis has not been noted in any of the rat diesel exhaust carcinogenicity studies. The studies by Driscoll et al. (1996, 1997) did not study diesel exhaust but rather utilized other insoluble particles (a quartz, carbon black) which, unlike diesel exhaust, have no directly genotoxic component which would have implications for low-dose response and therefore limits their applicability to explaining mechanisms of diesel exhaust-induced rat carcinogenicity. The work of Gaylor and Zheng (1996) is therefore useful in illustrating that cell proliferation, which is one of a number of potential components of the mechanism of rat lung tumor induction by diesel exhaust, may not exert a threshold effect on carcinogenicity. This information does not prove that diesel exhaust-induced carcinogenicity exhibits low dose linearity. However, the clonal expansion carcinogenicity modeling data and the diesel exhaust-induced genotoxicity (including oxidative DNA damage) and Ah (dioxin) receptor binding data indicate that diesel exhaust-induced carcinogenicity may exhibit low dose linearity without the existence of a threshold. A recent report by Borm et al. (1997) indicates that incubating rat lung epithelial-derived cells with human polymorphonuclear lymphocytes (PMN) (either unactivated or activated by preexposure to phorbol myristate acetate) increases DNA adduct formation caused by exposure to benzo[a]pyrene; addition of more activated PMN in relation to the number of lung cells further increased adduct formation in a dose-dependent manner. The authors suggest that "an inflammatory response in the lung may increase the biologically effective dose of polycyclic aromatic hydrocarbons (PAHs), and may be relevant to data interpretation and risk assessment of PAH-containing particulates." These data raise the possibility that low dose diesel exhaust exposure may result in levels of neutrophil influx which would not necessarily be detectable via histopathological examination as acute inflammation but which might be effective at amplifying any potential diesel exhaust genotoxic effect. WHO (1996) has noted that modeling of human cancer risk from rat lung tumor data should take into account the effects of both particles (carbon core) and extractable organic matter (PAHs, nitro PAHs).

Additionally, some parameters of the "particle overload" hypothesis are incompletely characterized. Alveolar type II cell epithelial hyperplasia has been noted after diesel exhaust exposure, but the measures of cell proliferation used were relatively crude and unsuitable for use in a quantitative estimate of cell proliferation as would be required for biologically-based modeling. It should also be noted that uncertainties exist regarding the magnitude and biological importance of particle overload for diesel exhaust-induced rat lung carcinogenicity.

Mauderly et al. (1994) included data from a rat bioassay on the number of neutrophils/mL present in bronchoalveolar lavage fluid from the exposed and control animals (males and females combined). Active oxygen species generated by activated neutrophils are one component of the inflammatory response to diesel exhaust exposure that might be mechanistically important to the induction of tumorigenesis. The number of neutrophils was increased approximately 50-75% for the high carbon black group compared to the low carbon black group; the increase for the high diesel exhaust group was 20-40% compared to the low diesel exhaust group. However, the tumor incidence (males and females combined) for the high carbon black and diesel exhaust groups were approximately 3-fold greater than that for the low carbon black and diesel exhaust groups, respectively. Similarly, the differences in the severity scores for alveolar macrophage hyperplasia and alveolar epithelial hyperplasia in rats that died or were killed after 18 months of exposure between the low and high diesel exhaust groups (approximately 25 and 20%, respectively) do not correlate well with tumor incidence. It would be expected that a better correlation between tumor incidence and indices of inflammation and cell proliferation would exist if diesel exhaust-induced rat lung tumors were solely due to particle overload.

Hattis and Silver (1994) examined lung burden data from diesel exhaust rat carcinogenicity studies and came to the conclusion that "there is continuing accumulation of diesel-derived dust in the lungs of rats throughout life, even at low doses". They also found that this was not predicted by models developed to represent diesel exhaust particulate matter accumulation under "overload" versus nonoverload conditions. Finally, they have found that at high diesel exhaust exposure levels, the increase in the ratio of internal diesel exhaust particulate matter burden to external exposure is not very large, being slightly larger than a factor of 2 at most, and state that "Although dust overloading is a real phenomenon, it is not a very large effect and thus would not be expected to give rise to dramatically lowered estimates of risk at low exposure levels." It is therefore premature to conclude that the carcinogenic response in rats to diesel exhaust is completely nonspecific.

Furthermore, the rat diesel exhaust lung tumor data discussed in Section 6 of this document is insufficient for the purposes of determining if a threshold for diesel exhaust-induced carcinogenicity exists. As an example, in the study by Mauderly et al. (1987), rats exposed to 350 µg/m^3 diesel exhaust demonstrated a non-statistically significant increase in lung tumor incidence (1.3% compared to 0.8% for controls; relative risk of 1.4). Comment 9 correctly points out that the problem is a case of sample size. To determine if a difference in lung tumor incidence of that magnitude is significant at a 95% confidence level would require approximately 4000 (not 15,000) animals/group. This suggests that with the data available, a determination that diesel exhaust induces increases in lung tumors at concentrations of less than 2.2 mg/m^3 cannot be made. However, it also indicates that insufficient data exists for determining that there is a threshold for diesel exhaust-induced rat lung tumors.

Finally, Section 6.1.1.1 of this document lists two studies (Pepelko and Peirano, 1983; Heinrich et al., 1986a) in which the authors describe statistically significant lung tumor induction in mice in response to diesel exhaust exposure. Another study (Takemoto et al., 1986) is also listed in which the authors reported increased tumor incidences in mice which were not statistically

significant, but which were later reported by IARC (1990) to be statistically significant after the data were reanalyzed. Section 6.1.3 has been changed to include a study by Ichinose et al. (1997) which describes lung tumor induction in mice after intratracheal exposure to diesel exhaust particulate matter. Of the six mouse diesel exhaust inhalation or intratracheal instillation carcinogenicity studies described in this document, three definitely contain positive results, one probably contains positive results, and two are negative. This indicates that the description of the mouse carcinogenicity results as "mixed" is justified, and it cannot be accurately stated that the only rodent species susceptible to diesel exhaust-induced lung tumors is the rat.

Comments on behalf of the Engine Manufacturer's Association
prepared by Tony Cox of Cox Associates, August 21, 1997

Comments on Executive Summary

OEHHA's 1997 draft risk assessment for diesel exhaust (DE) retains essential features of its 1994 draft risk assessment. OEHHA received many technical criticisms of its 1994 risk assessment in public comments, e.g., as reflected in the current draft's discussion of "uncertainties" about its methods and conclusions (especially Sections 7.2.8 and 7.3.3). A key question is whether the technical objections that have been raised threaten the validity of OEHHA's main conclusions, or whether they only point out ways to further improve an analysis that is basically sound. For example, OEHHA now acknowledges the following sources of "uncertainty" in their analysis:

- Use of approximate instead of exact model formulas;
- Ignored exposure uncertainties and measurement errors (p. 7-22);
- Treatment of assumed models as if they were known to be true;
- Restriction of the set of models considered to those that are low-dose-linear (p. 7-20);
- Ignored heterogeneity in individual exposures and response parameters (p. 7-22).

But are they more than just uncertainties - are they outright mistakes that invalidate OEHHA's main conclusions? How might one tell? This document examines each of OEHHA's main conclusions, and the arguments supporting them, from the perspective of how OEHHA has addressed or dismissed the above-noted uncertainties in their analysis of animal data, epidemiological data, meta-analysis, and causal interpretation. We find that most of their key conclusion about risk are not implied by (or even always consistent with) available data, but that instead that they arise primarily from modeling assumptions and practices corresponding to what OEHHA acknowledges as areas of "uncertainty". It appears that the key conclusions that OEHHA has drawn are assumption-driven rather than data-driven. The facts and data are, on their own merits, more consistent with the conclusion that DE creates no significant excess risk in humans at low exposure levels than with OEHHA's assumption-based conclusion that DE poses a significant risk (approximately proportional to cumulative exposure) even at low doses.

A point that critics of OEHHA's DE risk assessment have so far failed to persuade OEHHA Staff to accept is that many of the identified areas of uncertainty could be resolved relatively easily using more appropriate statistical methods. Such methods appear throughout the modern statistical and biostatistical literatures and can be implemented easily using widely available statistical software. Technically, models that represent uncertainty in exposure estimates, allow for interindividual heterogeneity, and are flexible enough to admit the possibility of low-dose nonlinearity would appear to be unambiguously more appropriate for modeling DE risk data than models that don't. However, OEHHA has chosen not to use such methods, instead opining that

the techniques they have used, despite their recognized errors and limitations, produce answers that might not be very inaccurate and that OEHHA considers "adequate" (C-OEHHA, 167, 168, 170). OEHHA's criteria for model adequacy are not stated. The basis for preferring simpler, less correct models to more complex, more accurate models is unclear, given the capabilities of modern statistics software. Tables 3, 4, and 6 offer suggestions for applying more accurate statistical methods to correct some of the main errors/limitations in OEHHA's modeling approach

The main purpose of this document is to see what new can be added to the discussion of human risks from DE exposure to simplify and clarify the main policy-relevant issues. The following new points go well beyond technical niceties. They address the central logic of OEHHA's analysis and conclusions.

Comment 1: *OEHHA's analysis of animal data is seriously flawed by unjustified aggregation of rat tumor data across sexes.* This is statistically invalid and creates a serious error of aggregation. When the Mauderly *et al.* tumor data for male and female rats are analyzed separately, both contain apparent response thresholds, contradicting OEHHA's findings based on the pooled data. The threshold hypothesis is better supported by the data than OEHHA's assumption of low-dose linearity. Correctly analyzed (i.e., without pooling tumors across sexes), the Mauderly *et al.* rat data do not support OEHHA's conclusions about low-dose risk. This finding is not new. Mauderly *et al.* originally stated that their data tends to support the hypothesis of a threshold for response. What is new is OEHHA's use of aggregation to obscure the threshold patterns in the data and contradict the findings of the original researchers.

***Response**: OEHHA has decided, following peer review including suggestions made by the Scientific Review Panel, to provide calculations of human cancer risk using rat lung tumor data on an informational basis, but not include those calculations in the final range of human cancer risk estimates since human cancer risk estimations based on human data are available.*

In Table 2 of this comment the commentator presents data separated into male and female rats, stated to be from Mauderly et al. (1987). It should be noted that Mauderly et al. (1987) themselves present an aggregation of rat lung tumor data across sexes in Table 3 of their study. The data as presented in the comment do not appear in the cited Mauderly paper. Staff note that all rats were included for this table, even the ones which were not examined for tumors. Thus the tumor incidence will be underestimated since the denominator includes many rats for which any contribution to the numerator is unknown.

As noted in Section 6.1.5, the rat bioassay data are insufficient to determine if sex differences exist in the development of lung tumors in rats after exposure to diesel exhaust. Four of the seven positive diesel exhaust inhalation rat bioassays used both male and female F344 rats. Brightwell et al. (1986; 1989) reported total lung tumor incidences of 44% and 96% in males and females, respectively, in animals sacrificed after the end of the 24 month exposure period. However, mortality data was not provided, so it cannot be determined if differences in survival between the two sexes affected tumor incidence rates. Ishinishi et al. (1988) found total lung tumor incidence to be greater in males than females at the highest two exposure levels for each of the two diesel engine types tested. Mortality rates were similar for both sexes. Mauderly et

al. (1986) reported similar mortality rates and total lung tumor prevalence rates for male and female rats. Nikula et al. (1995) noted that female rats were more susceptible than male rats to developing lung tumors after exposure to diesel exhaust; total lung tumor incidence rates for male rats were 4.8% and 8.5% for the 2.5 and 6.5 mg/m^3 diesel exhaust groups, respectively. Corresponding total lung tumor incidence rates for female rats were 7.6% and 27.4% for the 2.5 and 6.5 mg/m^3 diesel exhaust groups, respectively. However, mortality rates for male rats were also greater than female rats. At 23 months, percent cumulative mortality in males was 86% and 94% for the 2.5 and 6.5 mg/m^3 diesel exhaust groups, respectively; percent cumulative mortality in females was 69% and 73% for the 2.5 and 6.5 mg/m^3 diesel exhaust groups, respectively. As described in Section 6.1.1.2 of this document, the authors noted that logistic regression modeling did not show significant differences between the tumor responses to diesel exhaust and carbon black for either sex. However, they also noted that slope estimate errors in the logistic regression models were large; the errors were particularly large for the males, because of their much shorter lifespan compared to the females. Lung tumor incidence in the females increased rapidly towards the end of their lifespan, when most of the males had died of other causes. Also, the lack of lung tumors in the control females may have increased the estimated slopes for the treatment group females. The authors stated that these factors make it difficult to determine if there were true gender differences in neoplastic incidence or differences in the responses of males to diesel exhaust and carbon black. After considering this information, we agree that the decision by Mauderly et al. (1987) to present an aggregation of rat tumor data across sexes in Table 3 of their study was appropriate, and should be used to provide calculations of human cancer risk using that rat lung tumor data on an informational basis.

In general, the rat diesel exhaust lung tumor data discussed in Section 6 of this document (including the Mauderly et al. (1986) study) are insufficient for the purposes of determining if an exposure threshold for diesel exhaust-induced carcinogenicity exists. As an example, in the study by Mauderly et al. (1987), rats exposed to 350 µg/m^3 diesel exhaust demonstrated a non-statistically significant increase in lung tumor incidence (1.3% compared to 0.8% for controls; relative risk of 1.4). The problem in this case is sample size. To determine if a difference in lung tumor incidence of that magnitude is significant at a 95% confidence level would require approximately 4000 (Mauderly cites 15,000 in a comment on this document) animals/group. Another study (White et al., 1983) lists tumor incidences of 0/30, 1/30, and 3/30 at diesel exhaust concentrations of 0, 0.25 and 0.75 mg/m^3, respectively. The p value for the 0.75 mg/m^3 group is 0.12 (Fisher exact test); this value is less than the normal 0.05 cutoff, but comes close enough to significance to be suggestive. These studies suggest that with the data available, a determination that diesel exhaust induces increases in lung tumors at concentrations of less than 2.2 mg/m^3 cannot be made. However, it also indicates that insufficient data exists for determining that there is a threshold for diesel exhaust-induced rat lung tumors.

Comment 2: *OEHHA's reanalysis of the Garshick et al. data is flawed by failure to correct for the confounding effects of factors such as year of birth and age at death* (which are positively associated with both lung tumor rate and average DE exposure.) When the effects of such confounding are removed, DE concentration is negatively associated with lung cancer rate. Thus, the Garshick *et al.* data do not support OEHHA's conclusions of a positive statistical association (nor of a causal relation) between DE exposure and human lung cancer. This is consistent with

Garshick's own finding and reanalysis of his own data. As described by Cohen and Higgins (1995), "Recently, Garshick reanalyzed these data and found that when the effect of age was allowed to vary within birth cohorts, the apparent upward trend in the relative risk for cumulative exposure disappeared". (It is also consistent with independent work by Dr. Kenny Crump, who has repeatedly pointed out to OEHHA, using different analyses, that the *Garshick et al.* data do not support OEHHA's interpretation of a significant positive association between DE exposure and lung cancer risk.) The reanalysis of the Garshick *et al.* cohort data with correction for confounding has been peer-reviewed and will be published in *Risk Analysis* later this year (Cox, 1997). Its main methods and findings were shared with OEHHA in 1995 and 1996.

***Response**: The analyses of Appendix E in the 1997 draft TSD, Appendix D of the current draft, included the removal of effects of confounding due to age and year of birth, as well as calendar year. Comments on the new analyses of the commenter are responded to below. The findings of Drs. Garshick are responded to in connection with their respective comments and are also discussed in Appendix F of the current TSD.*

Comment 3: *OEHHA's new meta-analysis is flawed by failure to correctly calculate p-values* to correct for false positives due to multiple comparisons and multiple hypothesis testing. This problem also occurs in many of the individual studies cited by OEHHA, including the studies of Garshick *et al.* The result is that a pattern of consistently elevated relative risks is expected, whether or not DE exposure has a positive effect on lung cancer risk. Since this is the pattern that has been observed, OEHHA's meta-analysis offers no evidence either for or against the hypothesis that DE exposure has a genuine causal association with human lung cancer risk (as opposed to merely a statistical association due to improperly controlled false positives).

***Response:** The commenter suggests that multiple comparison bias is responsible for the findings of multiple point estimates of risk greater than unity. The theoretical underpinning of this statement is that, if multiple comparisons between exposures and outcomes are undertaken in a given epidemiological study, this increases the likelihood that there will be positive results based on chance alone. For example, if in a given study, 10 comparisons are made (e.g., between diesel exhaust exposure and cancers of the lung, stomach, bladder, brain, kidney and other organs), then the probability of at least one statistically significant association occurring will be $1 - (1-\alpha)^{10}$, where α = the given statistical significance level. If $\alpha = 0.05$, the conventional (though arbitrary) cutoff level for statistical significance, then the probability of a positive result = 0.40, assuming that the underlying null hypothesis is true (i.e., that there is in reality no association between the exposure(s) and the outcome(s) under study). Therefore, to avoid such theoretical false-positive results, some statisticians have recommended statistical adjustments for multiple comparisons such as those suggested by the commenter.*

There are several problems with the commenter's suggestion. The most important is that it invokes the universal null hypothesis - i.e., that all associations observed in a given data set are random and can be attributed to chance. As Rothman (1990) has observed, "To entertain the universal null hypothesis is, in effect, to suspend belief in the real world and thereby to question the premises of empiricism…In a body of data replete with associations, it may be that some are explained by what we call "chance," but there is no empirical justification for a hypothesis that

all the associations are unpredictable manifestations of random processes." In other words, the mechanical application of "correction" for multiple comparisons advocated by some statisticians is premised on an assumption that runs contrary to the foundations of empirical science. While such corrections will guard against inappropriate conclusions based on false positive results, they do so at the expense of rejecting real associations (that is, by increasing the rate of false negatives). In the case of diesel exhaust exposure, there are several sound biological reasons to suspect that occupational exposure to diesel exhaust would be related to lung cancer: to reject associations between these variables because the authors failed to make adjustments for multiple comparisons would be foolish. Again, to cite Rothman (1990), "[I]t is always reasonable to consider each association on its own for the information it conveys. This is not to say that the setting in which the observations are made should be ignored, but only to emphasize that there is no formula that can substitute for critical evaluation of each association or observation that comes to attention." Therefore, OEHHA respectfully disagrees with the commenter's suggestion.

Comment: 4. *OEHHA's causal interpretation of the relation between DE exposure and human lung cancer (Section 6.2.4) is unsupported by any formal statistical tests for causation.* The reported associations are expected based solely on the statistical methods used, even if DE exposure has no effect on lung cancer. Thus, OEHHA's meta-analysis does not support the conclusion that DE exposure contributes to human lung cancer risk.

Response*: The commenter is correct in stating that the meta-analysis deals only with statistical associations and that no statistical tests for causation were performed. Causal inference in chronic disease epidemiology involves an assessment of statistical associations, but requires an evaluation of a variety of other factors as well, including (among others) the consistency of the findings among multiple studies, whether the findings are likely to be due to bias or chance, biological plausibility, and the existence of exposure-response relationships. These and other considerations are discussed at length in section 6.2.4, "Causal inference for diesel exhaust exposure and lung cancer."*

In summary, none of the three data sources that OEHHA uses - rat, Garshick et al., and meta-analysis - is sufficiently robust to allow a conclusion that DE creates low-dose cancer risks in humans. Nor has OEHHA performed statistical tests of this hypothesis. Of course, they may simply assume that DE exposure causes human cancer risks. But then it should be made explicit to decision-makers that this conclusion rests solely on OEHHA's opinions and modeling assumptions and is not dictated either by correct analysis or by facts and data. The following sections develop these points more fully.

Comments on analysis of rat data

Section 7.2 of OEHHA's draft risk assessment applies a traditional multistage model and a "simplified Moolgavkar model" to the 1987 data of Mauderly *et al.* and concludes (Table 7.7, p. 7-38) that the 95% upper confidence limits for extrapolated unit risks in humans should fall in the range from 0.5×10^{-4} to 2.8×10^{-4} per microgram-per-cubic meter of DE concentration in inhaled air, depending on what assumptions are made about the appropriate dose metric. OEHHA

interprets this outcome as reinforcing their findings based on epidemiological data, and uses it to bolster their conclusion that relatively low levels of DE may create a substantial risk of lung cancer in exposed human populations.

Since 1994, many commentators have questioned OEHHA's use and interpretation of these data. Table 1 lists representative technical comments and summarizes OEHHA's responses. The following additional points are intended to simplify the discussion by noting that the experimental rat data do not address the low-dose issues of practical interest, that they do not support (and are not required for) OEHHA's conclusions about low-dose risks, and that the modeling issues about them are therefore irrelevant and can be disregarded. OEHHA's conclusions must stand or fall based on their epidemiological data analysis.

Comment 1. Available rat data do not address low-dose risks and provide no evidence of increased risk at low doses. We believe that, as a matter of logic, the Mauderly et al. data cannot be used to draw sound inferences about low-dose risks for DE. Our reasoning is as follows.

Premise: The Mauderly *et al.* data only show significantly elevated risks at the two highest dose levels. Table 2 recapitulates the original Mauderly *et al.* data, including squamous cysts. (As noted by OEHHA, inclusion or exclusion of the cysts makes little difference to the conclusions.)

Note that OEHHA's Table 7.1 aggregates these data across the two sexes. This masks the fact that the lowest non-zero dose level is associated with a *decrease* in observed tumors among male rats, rather than with an increase as predicted by both of OEHHA's models (Weibull multistage and simplified Moolgavkar).

Among female rats, a ten-fold increase in concentration from 0.35 to 3.5 is matched by only a two-fold increase in risk, but a further doubling of concentration is then matched by a quadrupling of observed tumor risk. If the dose-response relation has a conventional sigmoid shape, then this data pattern suggests that at low doses, there is substantial background risk of lung cancer among female rats, with only sampling variability observed at the two or three lowest concentrations. A significant positive (upward-curving) effect of concentration on lung cancer takes place only above 0.35, making it plausible that, at the two lower concentrations, there is no effect of dose on tumor rate. Thus, these data do not support OEHHA's claim that tumor risks are elevated at the lowest dose level (implied by both of the two models, Weibull multistage and simplified Moolgavkar that OEHHA has examined), and that no evidence of a threshold can be found.

In summary, OEHHA has aggregated two dose-response patterns, one for each sex, each of which is more consistent with the hypothesis of a response threshold than with the hypothesis of low-dose linearity, to obtain a composite data set in which there does not appear to be a response threshold. Such statistical sleight-of-hand is now well understood. Aggregation can often be used to create statistical patterns that contradict the underlying truth that holds in each of the aggregated groups (see e.g., J. Gurland and J. Sethuraman, "How pooling failure data may reverse increasing failure rates' *Journal of the American Statistical Association*, 90, 432, 1995, 1416-1423, and references therein.) Proper procedure is to examine the dose-response pattern for each sex separately. If both sex-specific dose-response curves are consistent with the

hypothesis of a concentration threshold for carcinogenic responses, as in Table 1, then the correct conclusion is that such a threshold is possible.

***Response**: As noted in the response to the comments on the Executive Summary, OEHHA has decided, following peer review including suggestions made by the Scientific Review Panel, to provide calculations of human cancer risk using rat lung tumor data on an informational basis, but not include those calculations in the final range of human cancer risk estimates since human cancer risk estimations based on human data are available.*

Additionally, as noted above in the response to the comments on the Executive Summary, we agree with the decision by Mauderly et al. (1987) to present an aggregation of rat lung tumor data across sexes in Table 3 of their study, and believe that use of their data to provide calculations of human cancer risk on an informational basis is appropriate. We also believe that the rat diesel exhaust lung tumor data discussed in Section 6 of this document (including the Mauderly et al. (1986) study) are insufficient for the purposes of determining if an exposure threshold for diesel exhaust-induced carcinogenicity exists.

Comment 2: All available scientific evidence is consistent with the hypothesis that the elevated risks observed at the highest dose levels in the Mauderly *et al.* experiment are explained by a non-chemical carcinogenic process, relevant only at high doses in which lung tissue is repeatedly damaged by mechanical abrasion from soot deposits that have not been cleared from the lung. Meanwhile, protective enzymes (such as GSH) that normally protect cells against the damage inflicted by such repetitive mechanical trauma are depleted by the very high, sustained exposures for which increases in lung tumors are observed. This mechanistic description fully explains the available data, but is presumably irrelevant at lower doses (Driscoll *et al.*, 1996; Nikula *et al.*, 1996).

Conclusion: The elevated risks observed at the highest dose levels in the Mauderly *et al.* experiment are irrelevant to the question of whether tumors might occur at the much lower doses of practical interest, presumably by a different (e.g., genotoxic) mechanism. Although no such low-dose mechanism has been discovered for DE, despite vigorous and sophisticated searches (Driscoll *et al.*, 1996; Nikula *et al.*, 1996), its existence cannot be logically disproved by the failure to find it. But the Mauderly *et al.* data neither support nor refute conjectures about possible low-dose effects. Thus, we recommend that the rat data not be used for purposes of drawing inferences about low-dose risks. OEHHA's claim that the rat data support their low-dose risk estimates is based on a statistically invalid aggregation of dose-response patterns across sexes.

***Response:** As noted in the response to the comments on the Executive Summary, OEHHA has decided, following peer review including suggestions made by the Scientific Review Panel, to provide calculations of human cancer risk using rat lung tumor data on an informational basis, but not include those calculations in the final range of human cancer risk estimates since human cancer risk estimations based on human data are available.*

Sections 6 and 7 of this document state that the mechanism of action by which diesel exhaust induces lung tumors in rats is not established. One proposed mechanism for diesel exhaust-induced rat lung tumors is that exposure to diesel exhaust particulate matter at high concentrations exceeds pulmonary clearance capabilities and causes chronic inflammation. This inflammation leads to macrophage and/or neutrophil-induced oxidative DNA damage resulting in mutations which are instrumental in the induction of lung tumors, and also to cell proliferation which may be mechanistically important to the promotion of the rat lung tumors. This mechanism has also been invoked for carcinogenicity caused by other insoluble particles (e.g. carbon black, titanium dioxide). Rat lung tumor induction due to high dose (2.2 mg/m^3 or higher) exposure to diesel exhaust may share some commonality of mechanism with other carcinogenic insoluble particles; this possibility is discussed in the document. Several authors (e.g. Driscoll, 1996; Nikula et al., 1997) have hypothesized that this mechanism may have an exposure threshold of action, and tumor induction due to this mechanism would also have a threshold. Gaylor and Zhang (1996) have suggested using the Moolgavkar-Venzon-Knudson clonal expansion carcinogenicity model that small increases in non-necrotic cell proliferation rates which may be undetectable may result in significant increases in tumorigenicity. They also state that 1) a nongenotoxic carcinogen that increases the cell proliferation rate via the cell division rate is not likely to have a threshold dose; 2) dose response curves for cell proliferation and tumor incidence do not necessarily mimic each other. These increases in cell proliferation may be effected either by a stimulated increase in cell division or by an inhibition of apoptosis (programmed cell death).

The rat diesel exhaust carcinogenicity studies included in this document that have evaluated diesel exhaust-induced lung cell proliferation (Heinrich et al., 1995; Nikula et al., 1995; 1997) used an insensitive measure of cell proliferation (histopathological comparison to controls). More appropriate measures for making quantitative comparisons of cell proliferation (e.g. labeling index determinations using bromodeoxyuridine (BrdU) DNA labeling) have not been employed, making it premature to state that a true threshold of diesel exhaust-induced lung cell proliferation has been determined. Also, lung cell necrosis has not been noted in any of the rat diesel exhaust carcinogenicity studies. The studies by Driscoll et al. (1996, 1997) did not study diesel exhaust but rather utilized other insoluble particles (a quartz, carbon black) which, unlike diesel exhaust, have no directly genotoxic component which would have implications for low-dose response and therefore limits their applicability to explaining mechanisms of diesel exhaust-induced rat carcinogenicity. The work of Gaylor and Zheng (1996) is therefore useful in illustrating that cell proliferation, which is one of a number of potential components of the mechanism of rat lung tumor induction by diesel exhaust, may not exert a threshold effect on carcinogenicity. This information does not prove that diesel exhaust-induced carcinogenicity exhibits low dose linearity. However, the clonal expansion carcinogenicity modeling data and the diesel exhaust-induced genotoxicity (including oxidative DNA damage) and Ah (dioxin) receptor binding data indicate that diesel exhaust-induced carcinogenicity may exhibit low dose linearity without the existence of a threshold. A recent report by Borm et al. (1997) indicates that incubating rat lung epithelial-derived cells with human polymorphonuclear lymphocytes (PMN) (either unactivated or activated by preexposure to phorbol myristate acetate) increases DNA adduct formation caused by exposure to benzo[a]pyrene; addition of more activated PMN in relation to the number of lung cells further increased adduct formation in a dose-dependent

manner. The authors suggest that "an inflammatory response in the lung may increase the biologically effective dose of polycyclic aromatic hydrocarbons (PAHs), and may be relevant to data interpretation and risk assessment of PAH-containing particulates." These data raise the possibility that low dose diesel exhaust exposure may result in levels of neutrophil influx which would not necessarily be detectable via histopathological examination as acute inflammation but which might be effective at amplifying any potential diesel exhaust genotoxic effect. WHO (1996) has noted that modeling of human cancer risk from rat lung tumor data should take into account the effects of both particles (carbon core) and extractable organic matter (PAHs, nitro PAHs).

Additionally, some parameters of the "particle overload" hypothesis are incompletely characterized. Alveolar type II cell epithelial hyperplasia has been noted after diesel exhaust exposure, but the measures of cell proliferation used were relatively crude and unsuitable for use in a quantitative estimate of cell proliferation as would be required for biologically-based modeling. It should also be noted that uncertainties exist regarding the magnitude and biological importance of particle overload for diesel exhaust-induced rat lung carcinogenicity. Mauderly et al. (1994) included data from a rat bioassay on the number of neutrophils/mL present in bronchoalveolar lavage fluid from the exposed and control animals (males and females combined). Active oxygen species generated by activated neutrophils are one component of the inflammatory response to diesel exhaust exposure that might be mechanistically important to the induction of tumorigenesis. The number of neutrophils was increased approximately 50-75% for the high carbon black group compared to the low carbon black group; the increase for the high diesel exhaust group was 20-40% compared to the low diesel exhaust group. However, the tumor incidence (males and females combined) for the high carbon black and diesel exhaust groups were approximately 3-fold greater than that for the low carbon black and diesel exhaust groups, respectively. Similarly, the differences in the severity scores for alveolar macrophage hyperplasia and alveolar epithelial hyperplasia in rats that died or were killed after 18 months of exposure between the low and high diesel exhaust groups (approximately 25 and 20%, respectively) do not correlate well with tumor incidence. It would be expected that a better correlation between tumor incidence and indices of inflammation and cell proliferation would exist if diesel exhaust-induced rat lung tumors were solely due to particle overload.

Hattis and Silver (1994) examined lung burden data from diesel exhaust rat carcinogenicity studies and came to the conclusion that "there is continuing accumulation of diesel-derived dust in the lungs of rats throughout life, even at low doses". They also found that this was not predicted by models developed to represent diesel exhaust particulate matter accumulation under "overload" versus nonoverload conditions. Finally, they have found that at high diesel exhaust exposure levels, the increase in the ratio of internal diesel exhaust particulate matter burden to external exposure is not very large, being slightly larger than a factor of 2 at most, and state that "Although dust overloading is a real phenomenon, it is not a very large effect and thus would not be expected to give rise to dramatically lowered estimates of risk at low exposure levels." It is therefore premature to conclude that the carcinogenic response in rats to diesel exhaust is completely nonspecific. It should also be noted that no data exists indicating that exposure to diesel exhaust causes lung tissue to be repeatedly damaged by mechanical abrasion

from soot deposits that have not been cleared from the lung, or that protective enzymes (such as GSH) that normally protect cells against the damage inflicted by such repetitive mechanical trauma are depleted by the very high, sustained exposures for which increases in lung tumors are observed.

Table 1. Past Criticisms of OEHHA's Animal DE Risk Assessment.

CRITICISM	OEHHA's RESPONSE	REJOINDER/ RECOMMENDATION
1. The draft risk assessment ignores scientifically relevant information about the mechanism of DE cancer induction in rats, which does not apply to humans at realistic exposure levels (C-OEHHA-146, Comment 9)	DE particles and associated organic are genotoxic and potentially might contribute to a low-dose cancer risk. The epidemiological studies discussed in the TSD provide strong evidence that DE-associated cancer occurs in humans.	Do not use the rat data, since they are irrelevant to OEHHA's conjectured low-dose mechanisms. Do not claim that DE causes cancer in humans based on statistical associations that are not causal (see Table 4)
2. OEHHA's selection of cumulative exposure as a dose metric is not justified by experimental data in rats, which suggests that there is strong, nonlinear time-dependence and concentration-dependence in the observed cancer response.	The TSD's assumption of cumulative exposure is plausible and quite customary (C-OEHHA, P.159)	Treat concentration and duration of exposure as two separate risk factors, rather than multiplying them together. The mechanism of high-dose rat lung cancer is not customary, and concentration-duration pairs with the same product may create very different risks.
3. OEHHA has selected inappropriate mathematical risk models that have not been validated and that ignore relevant mechanistic information (C-OEHHA-159-161)	The selected risk models are standard models from the TOX-RISK program. More realistic/accurate models are not. (C-OEHHA-161) Using model-free methods would depart from established practices in risk assessment.	Use model-free estimation methods (Table 4) that do not require preconceived theories of carcinogenesis (since a correct theory for low doses is unknown.) The unusual mechanism of observed DE cancer induction justifies departing from established default practices.

4. OEHHA uses a retracted set of models that ignore the possibility of zero or negative responses at low doses.	Using a wider set of models is outside the realm of practicality (C-OEHHA, 164)	It is practical and easy with many nonlinear regression packages to consider a fuller range of possible models. Let the data pick the best model (which should include OEHHA's as one possibility).
5. The TSD ignores model uncertainty (C-OEHHA-164)	The models used in the TSD are the most plausible available. They are both generally accepted and widely used.	The TSD models are not widely used for DE cancer risk modeling, nor were they designed for DE. Use model-free estimation methods instead, since what is known about DE cancer mechanisms is not described by available models.
5. The TSD ignores model uncertainty (C-OEHHA-164)	The models used in the TSD are the most plausible available. They are bother generally accepted and widely used.	The TSD models are not widely used for DE cancer risk modeling, nor were they designed for DE. Use model-free estimation methods instead, since what is known about DE cancer mechanisms is not described by available models.

TABLE 2. LUNG TUMOR RISKS (PREVALENCE AT DEATH) IN RATS CHRONICALLY EXPOSED TO DIESEL EXHAUST (Source: Mauderly *et al.*, 1987)

CONCENTRATION (mg/m^3)	ALL LUNG TUMORS*	
	Males	Females
0	0.01 = 2/182	0 = 0/182
0.35	0.005 = 1/184	0.01 = 2/183
3.5	0.02 = 4/182	0.02 = 4/182*
7	0.07 = 13/183*	0.09 = 16/181*

* = significantly elevated compared to control group ($p < 0.05$ based on chi-square test with Yates correction)

Comment 3: *OEHHA's selection of theoretical mathematical risk models for dose-response extrapolation is unjustified for the DE rat tumor data. Model-free methods such as nonparametric regression should be used; instead* OEHHA has selected two mathematical risk models, the Weibull multistage and simplified Moolgavkar models, both of which are supported by the TOXRISK@ software package. However, neither model was designed to describe the events (e.g., lung over-burdening, repetitive lung tissue wounding, GSH depletion, proliferation of injured cells) that have been shown experimentally to be associated with tumorigenesis at the high DE concentrations where lung tumors occur. Moreover, both models lead to low-dose-linear dose-response relations - an assumption that cannot be justified by the data and that tends to be undermined by the observed nonlinearities in the experimental rat data. Both models are generic - they ignore the specific knowledge about high-dose DE carcinogenesis that are relevant for the data to which they are applied.

The key issue in model selection for DE is that standard models (such as the Weibull multistage and Moolgavkar models) were developed to describe different biological phenomena from those involved in experimental DE-induced rat lung carcinogenesis. OEHHA admits that low-dose responses would presumably be based on different, as-yet only conjectured, biological mechanisms. Therefore, there is no biological justification for pre-selecting the Weibull-multistage and Moolgavkar forms for purposes of extrapolating from the high-dose responses to hypothesized risks at lower doses. Modern statistical methods allow a range of practical, desirable alternatives, including not specifying any particular theorized parametric model in advance. This seems to be desirable, given that OEHHA frankly admits that low-dose mechanisms of DE carcinogenesis are unknown and speculative.

OEHHA could reduce the expected error introduced by its preselection of only two possible mathematical model forms by considering a wider range of risk models that allow for the possibility that the dose-response function is zero or sub-linear at sufficiently low doses. Practical ways to do this include the following:

(a) *Model-averaging and model-weighting techniques (Buckland et al.,* 1997; Berger and Pericci, 1996). This approach deals with model uncertainty by allowing for a wide set of possible theoretical models and using the experimental data to judge their relative plausibility. Buckland *et al.* (1997) describe simple versions for use in applied work, directly addressing OEHHA's expressed concerns about the complexity involved in doing a better job.

(b) *Model-free estimation methods* e.g., nonparametric regression models (Hall and Turlach, 1997), model-free curve fitting, and computationally intensive smoothing methods that only require weak assumptions, e.g., that the dose-response curve be smooth, or that it be monotonic, or s-shaped, etc. These methods deal with uncertainty about the correct model by making very few assumptions and solving for the dose-response curve that best describes the empirical data points, without imposing any very strong theoretical preconceptions.

(c) *Computationally intensive model selection methods* (e.g., Shao, 1996). This strategy searches for the dose-response model that minimizes estimated prediction errors, based on the available data.

However, OEHHA has chosen to consider only the Weibull multistage and simplified Moolgavkar models. In defending this choice, OEHHA states (p. 7-10) that 'The analysis works with models that are considered to be the most plausible, and is not concerned with a mathematically complete set of alternatives that have no previous justification. However, the mathematical alternatives are difficult to rule out and may be considered to be a source of uncertainty." This reflects a misunderstanding of the nature of modern techniques such as model-free curve-fitting and nonparametric regression. The goal of these techniques is not to introduce unjustified alternatives to be ruled out, but rather to avoid introducing unnecessary theoretical assumptions in fitting dose-response curves to experimental data. As much as possible, the data should be allowed to determine the model that is used to describe the dose-response relation. It should select from a large set of *a priori* possibilities, with enough flexibility to adequately reflect the data (something that the Weibull multistage model has been criticized for not doing). Rather than either selecting or rejecting mathematical models that imply low-dose linearity *a priori*, for example, modem techniques attempt to let the experimental data determine the weight to be given to linear vs. nonlinear possibilities.

OEHHA does not know how (or whether) DE could cause cancer at low doses, so claiming that it has selected models that "are considered to be most plausible" (p. 7-10) is unwarranted. Many scientists, including Mauderly, have suggested that threshold or low-dose nonlinear models are more plausible than the ones that OEHHA has selected. When the most plausible models are not known, model-free techniques seem appropriate and should be used in addition to, or in preference to, pre-defined parametric models.

***Response**: Other models give results similar to what OEHHA obtained in the document (Table 7-7). Staff analyzed the separate male and female rat data provided in the comment. Using the linearized multistage model with the concentrations given in the Mauderly et al. (1987) report extrapolated to a continuous concentration (e.g., 0.35 mg/m^3 x 7 h/24 h x 5 d/7d x 1000 = 72.8*

mg/m^3) and scaling to the 2/3 power (the default risk assessment methodology used by USEPA and CalEPA) resulted in an inhalation unit risk of 4x10^{-4} (µg/m^3)$^{-1}$ using the female rat data and a risk of 3x10^{-4} (µg/m^3)$^{-1}$ using the male rat data. Using the LED$_{10}$ approach with scaling to the 3/4 power as suggested in the 1996 USEPA draft cancer risk assessment guidelines resulted in an inhalation unit risk of 3x10^{-4} (µg/m^3)$^{-1}$ using the female rat data and a risk of 2x10^{-4} (µg/m^3)$^{-1}$ using the male rat data. All 4 values are compatible with the values of 0.5 x 10^{-4} to 2.8 x 10^{-4} per microgram-per-cubic meter in Table 7-7. All values indicate theoretical upper bounds on the risk. USEPA (1994) and WHO (1996) used models similar to those of OEHHA.

With an even simpler procedure offered by Gaylor and Kodell, the 95% upper confidence limit of the tumor incidence at the lowest exposure concentration giving a statistically significant increase in tumors is linearly extrapolated to zero to give an upper bound on risk. For Mauderly et al. 3.5 mg/m^3 (728 mg/m^3 extrapolated to continuous exposure) gave 4/182 (0.02198) lung tumors. The 95% UCL on 0.02198 is approximately 0.05. The UCL on individual risk is 0.05/728 = 7x10^{-5} (µg/m^3)$^{-1}$. Scaling from rat to man with scaling to the 3/4 power would result in an inhalation risk of approximately 3x10^{-4}, similar to other estimates.

These methods all assume a no threshold mechanism for carcinogenesis and linearity of response at low doses. Other models or model-free curve estimation might give different results. However, following public and peer review, including review by the Scientific Review Panel, OEHHA is focusing on the epidemiological data for quantitative risk assessment. Unit risk factors based on rat lung tumor data will not be included in the final range of risks for humans. Thus, OEHHA does not want to spend more resources looking at unit risk estimates based on model-free curve estimation methods with the rat data as suggested by the commentator, especially when the various methods evaluated this far by OEHHA and other agencies produced similar results.

Comment 4: *The data make it more likely than not that there is a DE concentration threshold below which lung tumor risks are not elevated in rats. Both parametric and model-free methods give dose-response curves with this property when applied to either the male or the female rat data.*

OEHHA claims that the rat data provide no evidence to support the hypothesis of a threshold for carcinogenic responses. This is an artifact of the way in which they have chosen to aggregate and model the data, and it contrasts with the interpretation of the original authors, who stated that (Mauderly *et al.,* 1987):

> At the higher exposure levels, rats accumulated lung burdens of soot greater than those which would be predicted from results at the low exposure level. Vestal (1986) suggested that there is a threshold in the relationship between cumulative exposure (concentration x time) and this particle clearance “overload” phenomenon and that there should also be a threshold in the relationship between lung tumor incidence and dose (exposure concentration, cumulative exposure, or lung burden of soot). The results of the present study appear to support this hypothesis.

In defense of their model selection, OEHHA notes (p. 7-10) that the high degree of non-linearity exhibited by the bioassays suggests that the use of [other, Armitage-Doll type] models would be impractical because of the complex calculations which would require estimation of many parameters. Other possible models might also give more accurate low-dose extrapolation. Such questions of model specification are a further source of uncertainty." In effect, OEHHA acknowledges that their risk models may be incorrect, but suggests that obtaining a more correct answer is too difficult to be practical.

In reality, however, it is easy to use widely available software packages to perform nonlinear regression modeling for a variety of nonlinear models that involve no more parameters than the models that OEHHA has selected in their Table 7-4. Doing so shows that, contrary to OEHHA's findings, the Mauderly *et al.* data set leads to positive threshold concentrations below which no excess risk is predicted, for both male and female rats, in multistage risk models no more complicated than those selected by OEHHA.

Figure 1 presents an example in which nonparametric regression (a simple distance-weighted least squares or loess algorithm) is used to fit a smooth curve to the male rat data in Table 2. In the absence of OEHHA's preconceived theoretical restrictions, this data-smoothing technique indicates no evidence whatsoever of increased tumor risk for male rats ("RISKMALE") at concentrations below about 2 mg/m^3.

Figures 2 and 3 show the results of applying parametric models to the same data set, while Figure 4 shows an analogous figure for the female rat data. Figures 3 and 4 allow for the possibility of a threshold, using the technique suggested by OEHHA in their Table 7.4 (p. 7-33). This technique allows for the possibility of a threshold (indicated by an initial flat horizontal segment of the dose-response curve) and then estimates its value. If there is no threshold, the estimated value can be zero. Although OEHHA does not show the estimated value of the response thresholds in their tabulation of Weibull multistage model parameter estimates, our calculations (Figures 3 and 4) show a threshold between 1 and 2 for both male and female rats. (For male rats, attempts to include a positive linear term led to non-convergence of the estimation algorithm. Figures 1 and 2 suggest that this may be because no such term exists in reality.) Thus, in contrast to OEHHA's conclusions, the Mauderly *et al.* data suggest that a response threshold is not only possible, but is more plausible than OEHHA's assumption of low-dose linearity. Their claim to have used the "most plausible" models is therefore not supported by these data.

***Response**: As noted in the response to comment 1, OEHHA has decided, following peer review including suggestions made by the Scientific Review Panel, to provide calculations of human cancer risk using rat lung tumor data on an informational basis, but not include those calculations in the final range of human cancer risk estimates since human cancer risk estimations based on human data are available.*

In general, the rat diesel exhaust lung tumor data discussed in Section 6 of this document (including the Mauderly et al. (1987) study) are insufficient for the purposes of determining if an exposure threshold for diesel exhaust-induced carcinogenicity exists. As an example, in the study by Mauderly et al. (1987), rats exposed to 350 $\mu g/m^3$ diesel exhaust demonstrated a non-

statistically significant increase in lung tumor incidence (1.3% compared to 0.8% for controls; relative risk of 1.4). The problem in this case is sample size. To determine if a difference in lung tumor incidence of that magnitude is significant at a 95% confidence level would require approximately 4000 (Mauderly cites 15,000 in a comment on this document) animals/group. Another study (White et al., 1983) lists tumor incidences of 0/30, 1/30, and 3/30 at diesel exhaust concentrations of 0, 0.25 and 0.75 mg/m^3, respectively. The p value for the 0.75 mg/m^3 group is 0.12 (Fisher exact test); this value is less than the normal 0.05 cutoff, but comes close enough to significance to be suggestive. These studies suggest that with the data available, a determination that diesel exhaust induces increases in lung tumors at concentrations of less than 2.2 mg/m^3 cannot be made. They also indicate that insufficient data exist for determining that there is a threshold for diesel exhaust-induced rat lung tumors.

The comment addresses the question of what can be said about threshold using the overall dose response relationship. Figures 1 and 2 suggest that such an investigation might be warranted. Figures 3 and 4 purport to fit a Weibull model to the data. However, there is no indication of how the time dependence in the formula referred to in Table 7-4 was implemented. With no time dependence, this is not a Weibull model. The Weibull time-to-tumor model, used in the draft TSD is inherently more accurate than the end of life analysis with no time dependence. In any case a 95% UCL on the slope at the origin is customarily obtained in quantitative risk assessment, but this was not mentioned.

Comment 5: *OEHHA's use of cumulative exposure as a dose metric to extrapolate high-dose tumor risks to low-dose risk is unreasonable.* OEHHA extrapolates the low-dose linear models that it has fit to the aggregated Mauderly *et al*. rat data to project human risks by assuming that equivalent cumulative lifetime exposures create equivalent risks. But most tumors occur very late in the lives of affected rats. This poses a problem for the usual logic of cumulative-exposure extrapolation. If 7 mg/kg for one lifetime cause increased tumors in the last few weeks of life, then is it really plausible that 0.7 mg/kg for ten lifetimes would be expected to cause an equivalent increase in tumor risks in the last few weeks of life? Clearly not. The cumulative exposure hypothesis is not realistic because it would require extending exposure for several lifetimes to obtain equivalent risks. Less extremely, OEHHA's use of a cumulative dose metric ignores the age-dependence of tumor rates, and this appears to be an essential aspect of experimental tumors. Thus, OEHHA's assertion (C-OEHHA-159) that "The TSD's assumption of cumulative exposure in ppm-weeks as a dose metric in the TSD is plausible and quite customary in risk models for animals and humans" is not justified for DE, as opposed to chemical carcinogens in general. Yet without it, the high-dose rat data cannot be extrapolated to much lower human exposures.

***Response**: Alternate dose metrics give similar results. As mentioned above, other models give results similar to what OEHHA obtained in the document (Table 7-7). Staff analyzed the separate male and female rat data provided in the comment. Using the linearized multistage model with the concentrations given in the Mauderly et al. (1987) report extrapolated to a continuous concentration (e.g., 0.35 mg/m^3 x 7 h/24 h x 5 d/7d x 1000 = 72.8 mg/m^3) and scaling to the 2/3 power (the default risk assessment methodology used by USEPA and CalEPA) resulted in an inhalation unit risk of 4 x 10^{-4} (mg/m^3)$^{-1}$ using the female rat data and a risk of 3x10^{-4} (mg/m^3)$^{-1}$ using the male rat data. Using the LED$_{10}$ approach with scaling to the 3/4 power as*

suggested in the 1996 USEPA draft cancer risk assessment guidelines resulted in an inhalation unit risk of $3x10^{-4}$ $(mg/m^3)^{-1}$ using the female rat data and a risk of $2x10^{-4}$ $(mg/m^3)^{-1}$ using the male rat data. All 4 values are compatible with the values of 0.5 x 10^{-4} to 2.8 x 10^{-4} per microgram-per-cubic meter in Table 7-7. All values indicate theoretical upper bounds on the risk.

Comments on Risk Assessment Based on Reanalysis of the Garshick *et al.* Studies

A primary basis for OEHHA's quantitative risk assessment for DE is reanalysis of epidemiological data from the cohort and case-control studies of Garshick *et al.* As with the animal data, OEHHA has received and responded to many public comments on their analysis of these data, but its 1997 draft risk assessment remains very close in approach and results to the 1994 draft risk assessment.

Since publication of the studies, Garshick has updated his data and analysis, in part to remedy errors and omissions in the original data set, and has reported that the trend reported in his publications of increasing lung cancer risk with increasing cumulative exposure no longer holds (Garshick, 1991). Various investigators, most prominently Dr. Kenny Crump (1995), have argued that OEHHA's interpretation of the Garshick *et al.* data is not supported by the data. Other technical criticisms, summarized in Table 3, have been made, and either dismissed or discussed by OEHHA, but without changing the principle features of their approach or their most important conclusion - that the data somehow justify an inference that DE exposure increases lung cancer risk in humans, and with a potency higher than in any other species.

In continuing hope of persuading OEHHA to use technically correct statistical methods to reanalyze the Garshick *et al.* data - methods that do not ignore exposure uncertainties, that do allow for model uncertainty (including the possibility of nonlinearities at low doses, as well as the possibility of linearity), and that allow different individuals to have different dose-time-response relations, for example - Table 4 recommends practical methods for overcoming the various technical statistical problems encountered in analyzing the Garshick *et al.* data set. The cited references provide algorithms and discuss available software packages that can carry out the required calculations without placing an excessive burden on OEHHA's staff. Table 5 outlines in more detail various factors that can create a statistical association between DE exposure and lung cancer in epidemiological studies such as those of Garshick's. All except the first (a true causal relationship) have been found and documented to hold in the Garshick *et al.* studies (e.g., Cox, 1997). Table 6 recommends appropriate formal statistical tests for whether observed statistical associations are causal We urge OEHHA to apply these formal methods before drawing and promulgating policy-relevant conclusions about causation from epidemiological data, bearing in mind that there are many possible explanations for systematically elevated risk ratios (see Table 5) and that the Hill criteria (consistency, strength of association, biological plausibility, etc.) relied on in OEHHA's current draft "have not been as successful in sorting out the signal from the noise as might have been hoped some 30 years ago" in resolving the association-vs.-causation dilemma (N.E. Breslow, Statistics in Epidemiology: The Case Control Study", 1996 R.A. Fisher Lecture, *Journal of the American Statistical Association* 91, 433, 14-28).

Table 3. Past Criticisms of OEHHA's Risk Assessment Based on the Garshick *et al.* Studies

CRITICISM	OEHHA's RESPONSE	REJOINDERS AND RECOMMENDATIONS
1. OEHHA's risk model begs the key question of low-dose nonlinearity and thresholds.	OEHHA maintains that the linear relative risk model is valid for purposes of quantitative risk assessment (C-OEHHA-167).	Let the data influence the weight given to different (e.g., linear vs. Nonlinear) modeling possibilities (Lee 97, Gonzelez-Monateiga 96)
2. OEHHA's risk model ignores exposure measurement errors and uncertainties that can lead to inflated risk estimates.	In simple linear regression, the bias from neglected measurement error is downward. OEHHA expects biases to be small (<10%) (C-OEHHA-168)	In threshold models, the bias is upward and can be large (Carroll, 1997). OEHHA should use a model with exposure uncertainties, since exposures are unknown.
3. OEHHA model assumes that all individuals are equally susceptible to lung cancer. This is wrong and can bias risk estimates upward.	The reviewer has not presented a corrected analysis. It would complicate OEHHA's analysis to do so. OEHHA believes that an uncorrected model is adequate for their purposes (C-OEHHA-168)	Use appropriate statistical models (e.g., Ahn & Chen, 1997, Becker, 1997) that allow for interindividual heterogeneity.
4. OEHHA uses cumulative exposure as a dose metric This inconsistent with data on concentration vs. Lung tumor.	The models used in the TSD are the most plausible available (C-OEHHA-164)	Treat exposure concentration and exposure duration as two separate factors in risk modeling. Let the data determine whether only their product affects risk; don't assume it.
5. OEHHA's own calculations indicate a threshold or strong nonlinearity in exposure-response. A nonlinear model (e.g., multistage model with no linear term fits the data better with fewer parameters than OEHHA's straight-line mode.	Four dose groups is too few to make it prudent to fit a nonlinear multistage model (C-OEHHA-167). Multistage theory predicts low-dose linearity. The Garshick exposure data are highly uncertain, justifying a forced linear model (C-OEHHA-170)	OEHHA routinely fits multistage models to data from 4 dose groups (e.g., the Mauderly *et al.*, data.) Multistage theory may not apply to DC carcinogenesis. It does not justify fitting a straight line to nonlinear data. *Recommendation*: Pick the most appropriate model for the data via goodness-of-fit or other formal criteria (see Table 4), and /or use nonparametric regression.
OEHHA has only tested for statistical associations between DE exposure and lung cancer. They have not tested whether the associations are causal	The TSD's new meta-analysis supports a dose-response relation bolsters the argument against alternative causes. Epidemiological studies indicate that DE-associated cancer is observed in humans (C-OEHHA, P.147,152,172)	Apply relevant tests for causality (Table 6). The individual studies cited in the meta-analysis do not establish a causal relation between DE and lung cancer and do not test the hypothesis of alternative causes such as multiple comparisons bias.

7. OEHHA has not calculated or combined p-values correctly. Their analysis is flawed by multiple comparisons and multiple hypothesis testing.	Not addressed by OEHHA. (New to the meta-analysis.)	Correct for multiple hypothesis testing bias by using appropriate p-value adjustments (Efron, 1996; Toman, 1996; Westfall, 1997)
8. OEHHA has not resolved contradictory p-values in the literature.	Not discussed by OEHHA. (New to the meta-analysis.)	Re-estimate p-values using Bonferroni or other corrections for upward biases.

Table 4. Some Common Statistical Problems and Suggested Modeling and Data Analysis Methods for Dealing with Them

STATISTICAL PROGRAM	APPROPRIATE METHOD	REFERENCES
Exposure estimation error and/or exposure classification errors	Errors-in models Measurement error models	Judge *et al.*, 1985: Stefaski and Cook, 1995; Nakamura, 1992 (for Cox model); Carroll, 1997
Interindividual heterogeneity in exposures or response parameters	Mixture distribution models Classification tree analysis Em algorithm	Lancaster, 1990 Ahn & Chen, 1997 Becker, 1997
Model from unknown. Linear model inappropriate. Multi-way, nonlinear interactions among risk factors	Multivariate model-free methods (e.g., CART) Nonparametric regression Model selection techniques Non-parametric survival data analysis	Bachetti & Segal, 1995 Ahn and Chen, 1997 Gasser & Kneip, 1995; Lee, 1996 Buckland *et al.*, 1997 Lin, 1997
False positives due to multiple comparisons/simultaneous hypothesis test	Bonferroni-type adjustments of reported p-values. Monte-Carlo estimation and significance testing	Chesseman & Oldford, 1994; Hjorth, 1994 Buckland *et al.*, 1997
Causal analysis of associations in multiple time series.	Granger-Sims causality test	Granger, 1980; see also Table 1
Causal analysis of multivariate associations among multiple risk factor and end-points	Linear causal analysis Nonlinear multivariate causal analysis via causal graphs and conditional independence relations Non-experimental data analysis	Kenny, 1979, Heise, 1975 Pearl, 1996; Yao and Tritchler, 2996; Shafer, 1996. Swanson and Granger, 1997 Blalock, 1961; Campbell & Stanley, 1963
Attribution to DE of effects due to mixtures or interactions	Multivariate classification tree analysis	Biggs *et al.*, 1991; Michie *et al.*, 1994; Ahn and Chen, 1997

Table 5. Possible Sources of Significant Positive Associations Between DE and Lung Cancer in the Garshick *et al.* Data Set.

1. True causal relation:	DE causes lung cancer: DE → lung cancer risk.
2. Confounding.	DE ← other factors →lung cancer. (DE = diesel exhaust exposure. 'Other factors" may include year of birth, age at retirement, age at death, and so forth.).
3. Model selection bias:	Investigators try different statistical models (e.g., various exposure groups, exposure assumptions, model formulas, and effect definitions) until one is found that yields a "significant" positive relation. Using the data to select models may create false positives (Buckland et al., 1996; Hjorth, 1994).
4. Multiple comparisons bias:	Investigators examine many subsets of variables (pollutants, seasons, weather conditions) and subsets of people (by ages, medical status, etc.) until "significant" positive relations are found.
5. Extrapolation and attribution biases:	A statistical model (linear, logistic, or Poisson regression; proportional hazards, etc.) is used that falsely attributes positive effects at high concentrations and/or due to synergy among multiple factors to lower concentrations and/or to DE.
6. Sampling, selection, recall, and reporting biases:	Investigators interview subjects (e.g., families of deceased workers) who may not represent the population for which inferences are drawn.

Table 6. Some Formal Tests for Causality.

DATA	HYPOTHESIS TESTED	TEST	PRINCIPLE
Two time series	TEMPORAL CAUSATION: The association between two time series is causal	Granger-type tests (Granger, 1980; Sims, 1990; Boudjellaba, 1993; Hosoya, 1997)	The cause occurs before the effect and contains unique information about it.
Multiple variables in multiple periods	EXOGENEITY: A variable is determined from outside a system of equations (i.e., from outside a model)	Tests for exogeneity (Geweke, 1984; Ericsson and Irons, 1994)	Future values of erogenous variables do not help too predict past values of endogenous ones.
Multiple variables, enough observation to calculate joint and conditional frequency distributions.	CONDITIONAL INDEPENDENCE: One set of variables (e.g., health effects) is conditionally independent of a set of proposed causes, given the values of intervening variables.	Directed graph tests, tests for d-separation (Jensen, 1996; Pearl, 1996; Pearl, 1996; Yao and Tritchler, 1996).	If X causes Y and Y causes Z, then the positive association between X and Z should disappear when conditioned to the level of Y.
Correlations among multiple variables	PATH COEFFICIENTS: A model (system of linear equations relating variables) is consistent with a postulated causal structure, represented by a path diagram.	Path analysis (for linear models) (Heise, 1975; Kenny, 1979; Yao and Tritcler, 1996).	Linear effects are
Multi-equation model relating values of variables	CAUSAL ORDERING: One variable precedes another in the causal graph showing what is determined from what.	Simon-type algorithms for partial causal ordering of model variables (Simon, 1977; Yao and Tritcler, 1996)	Some subsets of variables suffice to determine their own values and the values of other variables. Thus, a system of equations creates a causal partial ordering among variables.

Reanalysis of the Garshick *et al.* study data using conditional independence tests (see Table 6) shows that any statistical association between DE concentration and lung cancer risk is not causal, insofar as lung cancer risk is conditionally independent of DE exposure concentration, given the values of other (specifically, age-related) variables.

OEHHA has already expressed little enthusiasm for using technically correct methods (described by OEHHA as "mathematically more complete") to obtain more informative and probably more accurate risk estimates, on the grounds that doing so "would unnecessarily complicate the TSD's presentation" (C-OEHHA-163). Nor have they been anxious to rigorously test their key hypotheses, e.g., by letting the data determine how much weight should be given to different possible models. Instead, they have observed that assumptions "would no longer be assumptions" if they were tested, and that they believe in their current conclusions without seeing any need for formal testing. Therefore, it seems likely that they will not heed the recommendations or apply the methods identified in Tables 3-6.

The following points are intended to establish that, in the -absence of further analysis, the Garshick *et al.* data do not support OEHHA's conclusions of a positive relation - either statistical or causal - between DE exposure and human cancer risk.

Comment 1: *OEHHA claims (p. 7-15) that "The quantitative risk assessments below derive slopes that estimate the increase in cancer risk for increase in diesel exhaust exposure." This causal interpretation of statistical associations is unwarranted* OEHHA has not shown that an increase in diesel exhaust exposure would increase human cancer risk. Instead, they only describe statistical associations. Such associations are not evidence of causation: they might be expected to occur whether or not there is a causal Association, for the reasons listed in Table 5. Interpreting statistical Associations as evidence of causation without testing this assumption rigorously (see Table 6) does a disservice to decision-makers, as the purported link between changes in DE exposure and resulting changes in public health impacts has not been established.

Response*: The commenter considers the sentence in the draft TSD to be stating an unwarranted causal interpretation of statistical associations. Issues related to causal inference are presented in Chapter 6 of the TSD. A primary goal of Chapter 7 is to provide estimates of the slope of the relationship of risk to exposure.*

Comment 2: *The hypothesis of a causal relation is not supported by the Garshick et al. cohort study*. The Garshick *et al.* (1988) study involves many sets of variables that are mutually correlated. For example, worker age at death is positively associated with both average DE concentration and with lung cancer incidence rate. When the confounding effects of such associations are removed, there is no remaining (potentially causal) association between DE concentration and lung cancer (Cox, 1997). Details were sent to OEHHA's Dr. Stan Dawson (personal communications from Dr. Tony Cox) in l995. Dr. Kenny Crump has independently arrived at a similar conclusion using different methods. OEHHA's insistence that the Garshick *et al.* study supports their conclusions and causal interpretations is not objectively warranted by the data.

Response*: Specific references to the 1995 details are needed in order to assess the applicability of that work. There is no specific demonstration as to how the associations may invalidate the analyses in Section 7.3 and Appendix E (now Appendix D).*

Causal inference in epidemiology is rarely, if ever, based on the results of a single study, and OEHHA staff members have not made any such inferences based only on the Garshick (1988) study. Issues related to causal inference for diesel exhaust and lung cancer are presented in Chapter 6 of the TSD.

Comment 3: *Concentration of DE is not positively associated with lung cancer risk in the Garshick ' et al. (1988) study.* Indeed, it has a non-significant negative association. This undermines any plausible causal interpretation of DE as a human lung carcinogen. OEHHA obscures this fact by only discussing *cumulative exposure* as an indicator of DE exposure history. Since duration of employment is positively associated with lung cancer, OEHHA is able to hide the non-positive association between DE concentration and cancer risk behind the overall positive association between cumulative exposure (to DE and all other concurrent occupational factors, based on duration) and lung cancer.

It is easy to regress individual lung cancers against multiple factors, including estimated duration and average concentration of DE exposure, to estimate their separate contributions. (As discussed in Cox, 1997, the actual relations among variables are nonlinear in several cases, so that multiple linear regression is only a useful starting point.) Table 7 summarizes the results of multiple linear regressions in which each variable in the first column is regressed against the other column variables. The numbers are standardized beta coefficients, indicating the estimated contribution of each independent variable to each dependent variable while linearly adjusting for the contributions of the other variables. Coefficients not in parentheses are highly statistically significantly different from zero. (When an F-test is used to select variables for inclusion in the model via standard forward subset selection, the variables without significant coefficients in Table 7 drop out, but the remaining coefficients are almost unchanged) Inspecting the row for CONC (= estimated average DE exposure concentration) shows that it is not significantly positively associated with LUNG1 (human lung cancer). Year of retirement (RET) is positively associated with both DE exposure (DURATION and CONC) and with lung cancer. More general nonlinear analysis (Cox, 1997) shows that age at death is also positively associated with lung cancer risk, as well as with DE exposure concentration; thus, death age is a confounding factor that could provide a non-causal explanation of any positive statistical association between DE exposure and lung cancer. OEHHA's data analysis has not accounted for such confounding effects, undermining their causal interpretation of the epidemiological data.

Table 7. Multiple Linear Regression Models for the Garshick et al. Cohort Study.

Dependent	Age59	Ret	Duration	Conc	Deathage	Lung 1	R^2
Age59	--	-0.25	-0.10	0.84	0.04		0.98
Ret	-0.28	--	0.88	0.18	0.06	0.003	0.98
Duration	-0.12	0.95	--	0.14	-0.03	(-0.0004)	0.98
Conc	1.06	0.20	0.14	--	0.05	(-0.001)	0.97
Deathage	0.47	0.59	-0.25	0.47	--	(-0.04)	0.76
Lung1	(0.08)	0.112	(-0.02)	(-0.04)	(-0.02)	--	0.004

Source: Cox, 1997

***Response**: The comment overstates the relevance of the new analyses in Cox (1997) with regard to a plausible causal interpretation of DE as a human lung carcinogen. These new analyses do not take into account important aspects of the cohort data. The first is the need to subtract background from the given concentrations in order to have an unexposed group with zero concentration. The second is the need to reconstruct a time-varying exposure pattern. The third is the need to exclude shopworkers because the proportion of those with the measured exposure is unknown. The fouth is the need to exclude the last four years of follow up because of apparent missing reports in those years.*

Comment 4: *OEHHA's own analysis of the Garshick et al. data indicates that a threshold model is much more plausible than a linear low-dose model for human data.* Figure 7-3, page 7-46 of the draft risk assessment shows that relative risks do not increase for the three lowest cumulative exposures, but increase dramatically for the fourth. This data pattern is consistent with a threshold model: the observed pattern fits the definition of a response threshold perfectly, but provides a relatively poor fit to the linear model (indicated by the straight line in Figure 7-3) assumed by OEHHA.

OEHHA suggests that the Garshick *et al.* data support a linear model over a threshold model and that "Although tests of other models might show somewhat better fits, a simple linear relationship appears to be the most reasonable choice at present for humans with no evidence of real sublinearity" (C-OEHHA-170). These suggestions are flatly contradicted by the data. Formal statistical tests confirm what is visually apparent in Figure 7-3: that a threshold model fits the data significantly better than a linear model. OEHHA responds that "consistent with the theoretical constraint', a linearized multistage model would (by definition) include a positive low-dose linear component (ibid.). But there is no theoretical constraint in the multistage model that requires a positive linear term. The linearity that OEHHA refers to as a "theoretical constraint" is imposed as a regulators convention (unjustified by statistical theory) in constructing confidence bands *et al.* In truth, if a nonlinear (e.g., purely quadratic or cubic) model were known to be correct, then correctly computed upper confidence limits would not be linear, but would approach zero at the origin.

OEHHA's risk model and calculations (Table 7.10) are highly idiosyncratic. The use of logtransformed relative risks and simple linear regression are not standard in risk analysis and have no obvious biological rationale. Rejecting better-fitting threshold and sub-linear models *a priori* in favor of an *ad hoc* log-linear model appears to violate OEHHA's own espoused principle of using goodness-of-fit in model selection and evaluation.

In summary, we recommend that OEHHA not enforce a straight-line fit to the nonlinear data. This methodological choice drives the rest of their risk analysis. It is based purely on an *ad hoc* assumption rather than on data or sound, clearly applicable theory. A better, equally practical alternative would *be model-averaging* (Buckland *et al.,* 1997), in which the true form of the relationship between exposure and response is treated (realistically) as unknown, and the data are used to weight different possible options, including linear and nonlinear possibilities.

Finally, how much numerical difference would a more flexible modeling approach be expected to make in OEHHA's quantitative risk estimates based on the Garshick data? As a very rough approximate bound, suppose that there are k alternative models that are considered at least as plausible as OEHHA's linear model. If these alternative models specify zero increased risk at low doses, then OEHHA's risk estimate should be reduced by at least 1/k (and further if the alternatives are more plausible than the linear model). In our opinion, a value of at least k = 4 is realistic, since there are at least three alternative models (quadratic, cubic, and threshold) that are at least as plausible as OEHHA's linear model Thus, we would expect that accounting for model uncertainty in the Garshick data reanalysis via model-averaging would reduce OEHHA's risk estimates (MLE and UCL) by at least a factor of 4.

Response*: The revised draft TSD no longer uses the trend shown in Figure 7-3 of the 1997 draft to calculate risks in Table 7-10. The results of a new analysis based on the individual data are now shown in Figure 7-3 of the revised draft. A summary of results of all the accepted analyses has become the new Table 7-10 of the revised draft.*

The idea of averaging over models to get a slope at the origin is an interesting one, apparently not yet tried, per se, in risk assessment. A common approach seems somewhat similar, taking the geometric mean of all candidate values for unit risk. The current work simply gives the range of all the candidate values for upper confidence limit on unit risk, based on analyses that assume an essentially linear relationship of hazard through a variety of categorical trends with exposure.

Comments on Risk Estimates Based on Meta-Analysis

OEHHA claims that its analysis and interpretation of the Garshick *et al.* data are bolstered by a meta-analysis of many other epidemiological studies. However, their meta-analysis is flawed in its treatment of significance levels and in its approach to causal evidence and interpretation. The purpose of this section is to explain why the meta-analysis provides no support for OEHHA's conclusions and fails to bolster the rest of the risk assessment.

Comment 1: OEHHA's claim that "Support for the Finding of a carcinogenic effect of *diesel exhaust also comes from the meta-analysis* in ***Appendix D (C-OEHHA-152) is not justified***. The meta-analysis deals only with statistical associations, rather than with cause and effect. No statistical tests for causation have been performed (see Table 6). The individual studies cited by OEHHA in their meta-analysis suffer from the artifacts listed in Table 5, so that they are expected to produce false positives (and hence the appearance of a small but consistent pattern of elevated risks in exposed populations) even in the absence of a causal relation between them. Thus, observing such a pattern provides no evidence of a causal association between DE exposure and lung cancer.

Response*: The commenter is correct in stating that the meta-analysis deals only with statistical associations and that no statistical tests for causation were performed. Causal inference in chronic disease epidemiology involves an assessment of statistical associations, but requires an evaluation of a variety of other factors as well, including (among others) the consistency of the findings among multiple studies, whether the findings are likely to be due to bias or chance, biological plausibility, and the existence of exposure-response relationships. These and other considerations are discussed at length in section 6.2.4, "Causal inference for diesel exhaust exposure and lung cancer."*

Comment 2: ***OEHHA (p. 6-47) notes that point estimates of relative risk tend to exceed 1 in many studies of DE exposure and cancer risk and states that "if these Findings were due to chance, one would expect a more nearly equal distribution of point estimates of risk above and below unity. 'This is an error. It confuses findings being 'due to chance' with findings being "unbiased (equally likely to fall above or below*** 1). Findings due entirely to chance may nonetheless contain biases that tend to make them systematically fall above 1 rather than below 1. For example, most investigators, as well as OEHHA in its meta-analysis, have engaged in "subset analysis" in which multiple subsets of workers are examined (e.g., based on age, job category, duration of exposure, etc.) and those subsets that produce statistically significant positive associations are reported. However, such analyses tend to systematically produce false positives (point estimates above 1) unless statistical significance levels are reduced to control for multiple comparisons / multiple hypothesis testing bias. Statistical techniques for appropriately reducing significance levels are available (e.g., simple, approximate Bonferroni inequality adjustments or more sophisticated and accurate Monte-Carlo methods) but do not appear to have been used by OEHHA or in the individual studies included in OEHHA's meta-analysis. Therefore, false positives due to chance alone (in conjunction with improper setting of p-values and confidence limits) are expected to produce a consistent tendency for relative risks to be greater than 1 in the studies examined by OEHHA. OEHHA is mistaken in claiming that this observed pattern is evidence against a chance explanation.

As a second example of how findings due to chance alone can systematically tend to produce relative risks greater than 1, suppose that exposure has no effect on cancer risk but that there is some heterogeneity in individual cancer risks. For example, suppose that the probability of death with lung tumor is 0.2 among sensitive people and 0.1 otherwise, and that half the population is sensitive (independent of DE exposure). Randomly matching exposed individuals with similar unexposed controls and computing relative risk would give four possible relative risk ratios:

0.2/0.2 = 1, 0.1/0.2 = 0.5, 0.2/0.1 = 2, and 0. 1/0. 1 = 1. These four outcomes are equally likely, since the distribution of risks is identical in the exposed and unexposed populations. Hence, the average relative risk obtained from a large number of such matchings will be (1 + 0.5 + 2 + 1)(1 /4) = 4.5/4 = 1. 1 25. In other words, the point estimate of the relative risk exceeds 1 even though exposure has no effect on risk. This simple example illustrates a principle that holds more generally: relative risk calculations that ignore heterogeneity in individual response probabilities within groups may be biased upward. Both OEHHA's proposed models and the risk models used in key studies relied on by OEHHA (such as those of Garshick *et al.)* make this mistake.

Response: *The commenter suggests that multiple comparison bias is responsible for the findings of multiple point estimates of risk greater than unity. The theoretical underpinning of this statement is that, if multiple comparisons between exposures and outcomes are undertaken in a given epidemiological study, this increases the likelihood that there will be positive results based on chance alone. For example, if in a given study, 10 comparisons are made (e.g., between diesel exhaust exposure and cancers of the lung, stomach, bladder, brain, kidney and other organs), then the probability of at least one statistically significant association occurring will be* $1 - (1-a)^{10}$*, where* a *= the given statistical significance level. If* $a = 0.05$*, the conventional (though arbitrary) cutoff level for statistical significance, then the probability of a positive result = 0.40, assuming that the underlying null hypothesis is true (i.e., that there is in reality no association between the exposure(s) and the outcome(s) under study). Therefore, to avoid such theoretical false-positive results, some statisticians have recommended statistical adjustments for multiple comparisons such as those suggested by the commenter.*

There are several problems with the commentator's suggestion. The most important is that it invokes the universal null hypothesis - i.e., that all associations observed in a given data set are random and can be attributed to chance. As Rothman (1990) has observed, "To entertain the universal null hypothesis is, in effect, to suspend belief in the real world and thereby to question the premises of empiricism…In a body of data replete with associations, it may be that some are explained by what we call "chance," but there is no empirical justification for a hypothesis that all the associations are unpredictable manifestations of random processes." In other words, the mechanical application of "correction" for multiple comparisons advocated by some statisticians is premised on an assumption that runs contrary to the foundations of empirical science. While such corrections will guard against inappropriate conclusions based on false positive results, they do so at the expense of rejecting real associations (that is, by increasing the rate of false negatives). In the case of diesel exhaust exposure, there are several sound biological reasons to suspect that occupational exposure to diesel exhaust would be related to lung cancer: to reject associations between these variables because the authors failed to make adjustments for multiple comparisons would be foolish. Again, to cite Rothman (1990), "[I]t is always reasonable to consider each association on its own for the information it conveys. This is not to say that the setting in which the observations are made should be ignored, but only to emphasize that there is no formula that can substitute for critical evaluation of each association or observation that comes to attention." Therefore, OEHHA respectfully disagrees with the commentator's suggestion.

Comment 3: As a second example of how findings due to chance alone can systematically tend to produce relative risks greater than 1, suppose that exposure has no effect on cancer risk but that there is some heterogeneity in individual cancer risks. For example, suppose that the probability of death with lung tumor is 0.2 among sensitive people and 0.1 otherwise, and that half the population is sensitive (independent of DE exposure). Randomly matching exposed individuals with similar unexposed controls and computing relative risk would give four possible relative risk ratios: 0.2/0.2 = 1, 0.1/0.2 = 0.5, 0.2/0.1 = 2, and 0. 1/0. 1 = 1. These four outcomes are equally likely, since the distribution of risks is identical in the exposed and unexposed populations. Hence, the average relative risk obtained from a large number of such matchings will be (1 + 0.5 + 2 + 1)(1 /4) = 4.5/4 = 1. 1 25. In other words, the point estimate of the relative risk exceeds 1 even though exposure has no effect on risk. This simple example illustrates a principle that holds more generally: relative risk calculations that ignore heterogeneity in individual response probabilities within groups may be biased upward. Both OEHHA's proposed models and the risk models used in key studies relied on by OEHHA (such as those of Garshick *et al.)* make this mistake.

Response*: The commenter proposes a hypothetical situation in which genetic susceptibilities for lung cancer are equally distributed between exposed and unexposed groups. He then presents a situation of random individual matchings of individuals in the exposed and unexposed groups, which result in a theoretical average risk to be derived from this scenario of 1.125. However, the commenter puts forth no empirical evidence supporting his assumptions. Moreover, despite his superficially appealing example, it is well accepted in epidemiological theory that a potential confounder (e.g., genetic susceptibility) that is independent of exposure does not meet the definition of a confounder and will not influence the estimate of the relationship between the outcome variable and the exposure (Rothman 1986). Ordinarily, an equal distribution of this genetic risk factor in the exposed and unexposed populations would control for potential confounding, and would not affect the relative risk estimate associated with an exposure of interest, such as diesel exhaust. In a given study, however, it is possible that genetic susceptibilities might be unevenly distributed between exposed and unexposed groups, which could lead to confounding. This could cut both ways, so that susceptibles could in one study be more common among the exposed, and in another, more common among the unexposed, leading to positive and negative confounding, respectively. It is unlikely that only positive confounding would occur, that is, that those genetically susceptible to developing lung cancer would be over-represented only in the exposed populations in all the diesel exhaust studies that reported lung cancer risk estimates greater than one.*

Comment 4: ***OEHHA states (p. 6-47) that "In the studies with the more complete diesel-related exposure and duration of employment information, several identified exposure-response relationships, including the two studies by Garshick et al." But no such exposure-response relationships have been unambiguously identified***. For example, in their cohort study, Garshick *et al.* (1988, p. 823) conclude that "In this study we demonstrate an association between diesel exhaust exposure and lung cancer." However, as described by the authors, "With recent exposure included, no evidence of a consistent exposure duration-response relationship was obtained... When exposure in the year of death and the 4 years before were disregarded... the group with at least 15 years of exposure (with current exposure not included) had a relative risk

of lung cancer of 1.72 (95% CI = 1.27, 2.33)", emphasis added.) For the authors to exclude the most recent four years worth of data is an *ad hoc* truncation of the data that generates a positive result in this study but not in the case-control study, where "the relative odds ratio of lung cancer decreased slightly with recent exposure disregarded" (ibid., p. 823). A positive result created only by selectively discarding data (or, equivalently, selecting a subset of the data to analyze), with the selection being made differently in different studies to maximize positive results, clearly runs the risk of being a false positive. For OEHHA to assert that such ambiguous evidence "identified an exposure-response relationship" is misleading.

It is also misleading to characterize the two studies of Garshick *et al.* as having more complete diesel-related exposure information", since no exposure information whatsoever was available for the individuals in these studies. The apparent exposure-response relationship may be due partly to ignored exposure measurement error (Carroll, 1997). OEHHA has deliberately refused to use appropriate measurement-error models. [See page C-OEHHA-168. Here, a discussion of measurement errors in simple linear regression models of doubtful relevance to binary outcomes (lung cancer or no lung cancer) is followed by the statement that "OEHHA staff, then, do not agree that the realism of the present approach needs to be improved" by allowing for exposure measurement errors.]

***Response**: As noted in responses to several other commentators as well, there was no concurrent industrial hygiene measurement undertaken on any of the populations under study, so that indirect measures of exposure, such as duration of employment were used by the various study authors. In characterizing the Garshick studies as having "more complete diesel-related exposure and duration of employment information," OEHHA was referring to the extensive characterization of exposure for numerous railroad job classifications undertaken by the same group of Harvard investigators during the early 1980s (Woskie et al. 1988a, b). This industrial hygiene information was then used to classify qualitatively the likely exposures of the study subjects in the Garshick cohort and case-control studies. However, OEHHA staff agree with the commenter that the characterization of this information in the prior draft could be somewhat misleading and have modified the text in the revised version. Additional explanatory text has also been inserted to describe the quality of the exposure-response relationships in the Garshick studies.*

Comment 5: ***OEHHA states (p. 6-48) that "'The meta-analysis identified evidence of exposure-response relationships in the subgroup analyses based on duration of exposure. "The claimed relationships are likely outcomes of improper statistical methodology in the individual studies*** - something that OEHHA should have identified and discussed in deciding which studies to include in their meta-analysis. For example, Garshick *et al.* (1986, p. 1242) report that, in their case-control study, 'Workers 64 years of age or younger at the time of death with work in a diesel exhaust exposed job for 20 years had a significantly increased relative odds (odds ratio = 1.41, 95% CI = 1.06, 1.88) of lung cancer.' This presumably contributes to OEHHA's claimed "evidence of exposure-response relationships." But is it based on unsound analysis. The statement is an instance of a whole family of statements of the form "Workers who were A years or younger at the time of death and who were exposed to diesel exhaust for Y years had a significantly increased relative odds ratios for lung cancer. The probability of at least one

false positive occurring among the multiple hypotheses in this family corresponding to different combinations of A (e.g., no more than 54, 59, 64, 69, 74, 79, etc. years old at death) and durations of exposure (e.g., Y = 5, 10, 15, 20, 25, etc. years) is not limited to 5% when each combination of A and Y values is tested at a p = 5% significance level. For example, if 30 different (A, Y) combinations are considered, each independently having a 5% probability of a false positive (i.e., a reported 5% significance level), then the probability of at least one false positive occurring in the study as a whole is p = 1 - (1 - 0.05)30 = 78%. This p-value for the whole study is more than 15 times greater than the reported significance level of 5%. OEHHA cites such results as evidence for a statistically significant (or causally significant) exposure-response relationship without noting that p-values in the individual studies have not been correctly calculated. They have been inadequately critical in selecting results for inclusion in their meta-analysis.

Response*: The commenter suggests that in the Garshick (1987) study, the authors' failure to "correct" statistically for multiple comparisons invalidates not only their results, but also any work that OEHHA has done incorporating the Garshick study. As noted in the response to Dr. Cox's comment # 2, above, this mechanical statistical approach to epidemiology is based on a flawed underlying assumption of the universal null hypothesis. OEHHA staff respectfully disagree with the commenter.*

OEHHA References:

1. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. Epidemiology 1994;5:189-196.
2. Balarajan R and ME McDowall: Professional drivers in London: a mortality study. Brit J Ind Med, 45, 483-486, 1988.
3. Blair *et al*. Guidelines for application of meta-analysis in environmental epidemiology. Reg Toxicol Pharmacol 1995;22:189-197.
4. Boffetta P, SD Stellman and L Garfinkel: Diesel exhaust exposure and mortality among males in the American Cancer Society prospective study. Am J Ind Med, 14, 403-415, 1988.
5. Coggon D, B Pannett and ED Acheson: Use of job-exposure matrix in an occupational analyses of lung and bladder cancers on the basis of death certificates. J Natl Cancer Inst, 72, 61-65, 1984.
6. Garshick E, MB Schenker, A Munoz, M Segal, TJ Smith, SR Woskie, SK Hammond and FE Speizer: A case-control study of lung cancer and diesel exhaust exposure in railroad workers. Am Rev Respir Dis, 135, 1242-1248, 1987a.
7. Garshick E, MB Schenker, A Munoz, M Segal, TJ Smith, SR Woskie, SK Hammond and FE Speizer: A retrospective cohort study of lung cancer and diesel exhaust exposure in railroad workers. Am Rev Respir Dis, 137, 820-825, 1988.
8. Greenland S. Quantitative methods in the review of epidemiologic literature: Epidemiol Rev 1987;9: 1-30.
9. Greenland S: Invited commentary: a critical look at some popular meta-analytic methods. Am J Epidemiol 1994;140: 290-296.
10. Gustafsson L, S Wall, LG Larsso, and B Skog: Mortality and cancer incidence among Swedish dock workers--a retrospective cohort study. Scand J Work Environ Health, 12, 22-

26, 1986.
11. Howe GR, D Fraser, J Lindsay, B Presnal and SZ Yu: Cancer mortality (1965-77) in relation to diesel fume and coal exposure in a cohort of retired railway workers. J Natl Cancer Inst, 70(6), 1015-1019, 1983.
12. Kauppinen TP, TJ Partanen, SG Hernberg, JI Nickels, RA Luukkonen, TR Hakulinen and EI Pukkala: Chemical exposures and respiratory cancer among Finnish woodworkers. Brit J Ind Med, 50, 143-148, 1993.
13. Last JM. Dictionary of Epidemiology, 2nd ed., Oxford University Press, New York, 1988
14. Menck HR and BE Henderson: Occupational differences in rates of lung cancer. J Occup Med, 18(12), 797-801, 1976.
15. Monson RR. Occupational epidemiology, 2nd ed., CRC Press: Boca Raton, FL, 1990, p. 88.
16. National Institutes of Health. Respiratory health effects of passive smoking: lung cancer and other disorders. The report of the U.S. Environmental Protection Agency. Smoking and tobacco control monograph 4. NIH Publication No. 93-3605, 1993
17. National Research Council, Environmental tobacco smoke. Measuring exposures and assessing health effects. Washington, D.C.: National Academy Press, 1986
18. Petitti DB. Meta-analysis, decision analysis, and cost-effectiveness analysis. Methods for quantitative synthesis in medicine. Monographs in Epidemiology and Biostatistics. 1994. Vol 24. Oxford University Press.
19. Rothman KJ. Modern epidemiology. Boston: Little, Brown & Company, 1986.
20. Rothman KJ. No adjustments are need for multiple comparisons. Epidemiology 1990;1:43-46.
21. Swanson GM, CS Lin and PB Burns: Diversity in the association between occupation and lung cancer among black and white men. Cancer Epidem Bio Prev 2, 313-320, 1993.
22. U.S. Department of Health and Human Services. Reducing the health consequences of smoking. 25 years of progress. A report of the Surgeon General. U.S. Department of Health and Human Services: Rockville, MD, 1989, p. 39.
23. U.S. Department of Health and Human Services. The health consequences of involuntary smoking. A report of the Surgeon General. Washington, D.C.: U.S. Government Printing Office, 1986
24. Wegman DH and JM Peters: Oat cell lung cancer in selected occupations. A case-control study. J Occup Med, 20, 793-796, 1978.
25. Woskie SR, TJ Smith, SK Hammond, MB Schenker, E Garshick and FE Speizer. Estimation of the diesel exhaust exposures of railroad workers: I. Current exposures. Amer J Ind Med, 13, 381-394, 1988a.
26. Woskie SR, TJ Smith, SK Hammond, MB Schenker, E Garshick and FE Speizer: Estimation of the diesel exhaust exposures of railroad workers: II. National and historical exposures. Amer J Ind Med, 13, 395-404, 1988b.

Comments on Causation Not Demonstrated

Comment 1: Consideration of the above-noted topics leads to a different set of conclusions than those obtained by OEHHA. The final conclusion articulated by OEHHA, that of a probable causal link between DE exposure and human lung cancer, does not follow from application of data-driven analyses, as described above and as summarized here. OEHHA's draft risk assessment for DE asserts that the Mauderly *et al.* rat data, the Garshick *et al.* human data, and

multiple studies considered in their meta-analysis provide mutually consistent evidence of a no-threshold dose-response relation, implying that low levels of DE increase human lung cancer risks in proportion to cumulative exposure. Our reexamination of these three data sources reveals opposite conclusions. The rat data provide stronger evidence for a threshold relation than for OEHHA's low-dose linear models, precisely as stated by the original investigators. The Garshick *et al.* data as analyzed by OEHHA also suggest a threshold model much more strongly than a linear model. A reanalysis of the Garshick *et al.* cohort data using concepts of causal analysis shows that there is no causal association between DE and lung cancers. The studies in the meta-analysis are ambiguous and provide no clear evidence for or against the hypothesis of a causal Link between DE exposure and human lung cancer risk. Thus, OEHHA's claims that their risk estimates are backed by multiple sources of data and evidence is unjustified. Their risk assessment is entirely dominated by one extreme assumption - that relative risks are related to cumulative exposure by a straight line. This key assumption lacks theoretical or biological justification and is contradicted by both the animal and the human data. If the data were used to help select appropriate risk models, then the best-supported models would predict a threshold or low-dose sub-linear dose-response relationship, implying that DE does not create a human health risk at the exposure levels of interest. Any other conclusion reflects prior convictions or untested assumptions rather than available facts and data.

Response: *As noted in the response to comments 1 and 4, OEHHA has decided, following peer review including suggestions made by the Scientific Review Panel, to provide calculations of human cancer risk using rat lung tumor data on an informational basis, but not include those calculations in the final range of human cancer risk estimates since human cancer risk estimations based on human data are available.*

In general, the rat diesel exhaust lung tumor data discussed in Section 6 of this document (including the Mauderly et al. (1987) study) are insufficient for the purposes of determining if an exposure threshold for diesel exhaust-induced carcinogenicity exists. As an example, in the study by Mauderly et al. (1987), rats exposed to 350 µg/m^3 diesel exhaust demonstrated a non-statistically significant increase in lung tumor incidence (1.3% compared to 0.8% for controls; relative risk of 1.4). The problem in this case is sample size. To determine if a difference in lung tumor incidence of that magnitude is significant at a 95% confidence level would require approximately 4000 (Mauderly cites 15,000 in a comment on this document) animals/group. Another study (White et al., 1983) lists tumor incidences of 0/30, 1/30, and 3/30 at diesel exhaust concentrations of 0, 0.25 and 0.75 mg/m^3, respectively. The p value for the 0.75 mg/m^3 group is 0.12 (Fisher exact test); this value is less than the normal 0.05 cutoff, but comes close enough to significance to be suggestive. These studies suggest that with the data available, a determination that diesel exhaust induces increases in lung tumors at concentrations of less than 2.2 mg/m^3 cannot be made. They also indicate that insufficient data exist for determining that there is a threshold for diesel exhaust-induced rat lung tumors.

The comment addresses the question of what can be said about threshold using the overall dose response relationship. Figures 1 and 2 of Comment 4 suggest that such an investigation might be warranted. Figures 3 and 4 of Comment 4 purport to fit a Weibull model to the data. However, there is no indication of how the time dependence in the formula referred to in Table

7-4 was implemented. With no time dependence, this is not a Weibull model. The Weibull time-to-tumor model, used in the draft TSD is inherently more accurate than the end of life analysis with no time dependence. In any case a 95% UCL on the slope at the origin is customarily obtained in quantitative risk assessment, but this was not mentioned.

The commenter states that the studies in the meta-analysis are ambiguous and provide no clear evidence for or against a causal relationship between diesel exhaust exposure and human lung cancer risk. The commenter's concerns underlying this statement are addressed in responses to his comments # 1 - 3 on the meta-analysis, above. The meta-analysis provides evidence consistent with the hypothesis that exposure to diesel exhaust is associated with an increased risk of lung cancer. The pooled estimates derived from a variety of subset analyses clearly reflect the existence of a positive relationship between diesel exhaust and lung cancer in a variety of diesel-exposed occupations, which is supported when the most important potential confounder, cigarette smoking, is measured and controlled.