Comments on MSHA's Quantitative Risk Assessment (QRA) for RCMD

FINAL REPORT TO THE NATIONAL MINING ASSOCIATION

Louis Anthony (Tony) Cox, Jr.

Cox Associates

503 Franklin Street
Denver, CO 80218

tony@cox-associates.com

March 11, 2011

Executive Summary

This report comments on the methods and conclusions in MSHA's recent report, entitled *Quantitative Risk Assessment in Support of Proposed Respirable Coal Mine Dust Rule*. We find that MSHA's QRA makes crucial omissions and errors in statistical analysis and interpretation of data; that it does not remove some important sources of bias in its analysis; that it introduces its own upward biases in estimating future exposures and risks; and that it omits essential risk assessment steps needed to reach valid conclusions. Any of these errors and omissions alone would make the QRA's conclusions untrustworthy. Together, they invalidate the QRA's two major claims, to have shown that (a) Currently permitted levels of exposure to respirable coal mine dust (RCMD) cause excess risks of lung disease in miners; and (b) Further reducing permitted levels of RCMD exposure would further reduce risks of lung disease. These claims are based on fatally flawed statistical and risk assessment methods. Neither follows from sound analysis of the data.

Important flaws in the QRA include the following.

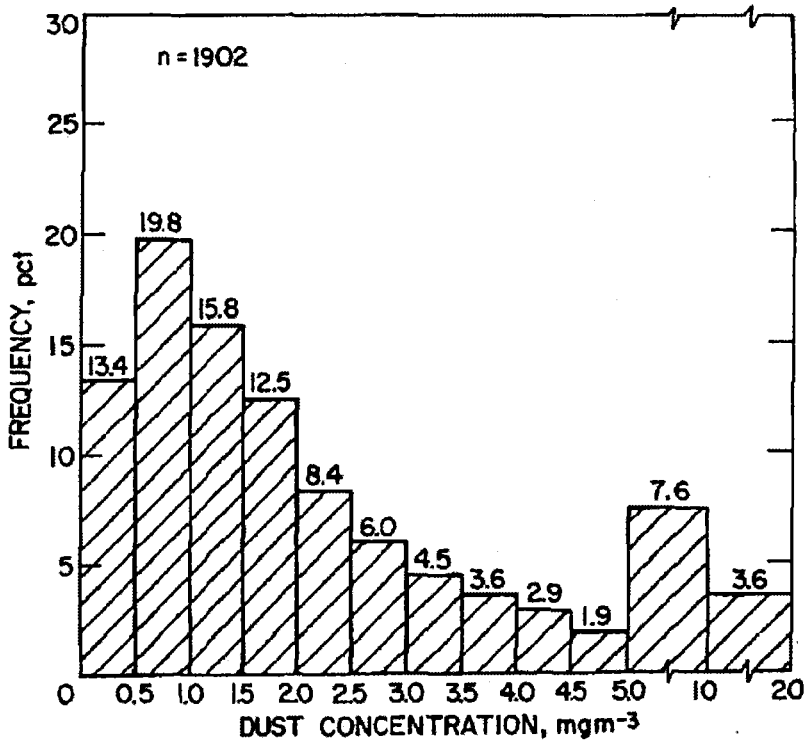
- *Unjustified causal interpretation of correlations and trend data.* Between 1930 and 2000, exposure levels to respirable coal mine dust (RCMD) fell, prevalence of cigarette smoking among adults in the United States declined, prevalence of lung disease among coal mine workers decreased, and per capita consumption of butter fell by about two thirds. *These overlapping trends alone do not justify any conclusions that some of the changes caused others.* Specifically, they do not show that reductions in exposure levels for RCMD *caused* declines in miner lung disease rates, any more than they show that decreases in butter consumption caused declines in lung disease.

  To find out whether changes in one quantity (such as decreases in exposures) might have contributed causally to changes in another (such as decreases in lung diseases), statisticians must apply appropriate statistical tests. For example, they can test whether future changes in the second quantity (the suspected "effect" variable) can be predicted better with knowledge of past values of the suspected cause variable than without that knowledge (e.g., http://en.wikipedia.org/wiki/Granger_causality.) If not, then even a strong correlation between variables may merely reflect their overlapping trends, rather than a true causal relation. That is, two decreasing variables will both tend to have higher values early on and lower values later, but this association provides no evidence that high levels of one cause high levels of the other, or that further decreasing one will cause further decreases in the other.

However, the QRA reports no results of such causal analyses or statistical tests for potential causal relations. Instead, it relies exclusively on correlations, regression coefficients, and other measures of statistical *association,* interpreted (without justification) as if they were valid indicators of *causation*. Presenting such associations as if they were causal is a type of fallacy or "proofiness," i.e., "the art of using bogus mathematical arguments to prove something that you know in your heart is true — even when it's not" (www.nytimes.com/2010/09/19/books/review/Strogatz-t.html). MSHA's QRA makes this fundamental mistake, purporting to show (via regression modeling) that presently permitted mean cumulative levels of RCMD cause excess lung diseases and material impairments in miners, and that decreasing permitted exposure levels would reduce these risks. But the regression models relied on are not valid statistical tools for drawing such causal conclusions. Thus, MSHA's key conclusions do not follow from valid causal analysis of data.

- *Omitted hazard identification.* A key scientific question left unaddressed by the QRA is: Do currently permitted exposure levels create any excess risks of lung disease? The QRA *assumes* that the answer is yes, but it does not justify or validate this assumption. Specifically, it does not address the realistic alternative possibility that excess risks observed in the past were created only by individual exposures that sometimes greatly exceeded currently permitted exposure levels. As shown in Figure 1, estimates of the distribution of exposures in past decades include some levels that are well in excess of currently permitted standards. Whether material impairment results only in workers repeatedly exposed to such higher-than-currently-permitted levels for several decades is not addressed in the QRA; nor does the QRA assess whether current standards, if enforced, would be fully health-protective. In short, MSHA's QRA omits the *hazard identification* section of the standard QRA process (e.g., Goldstein 2005), i.e., the part that should critically assess evidence on whether currently permitted exposure levels pose a hazard, by causing increased risk of disease. Although the QRA does reference other hazard identification efforts, it does not critically assess them, or use them to answer whether currently permitted exposure levels are health-protective. A hazard identification for RCMD risk assessment should summarize and synthesize available evidence on RCMD mode of action and biological causal mechanisms, evidence on exposure-response thresholds, and current knowledge about the relationship between currently permitted exposure levels and empirically observed exposure thresholds for lung diseases caused by mineral dusts (e.g., Porter et al. 2004, 2006). *The QRA does not demonstrate a causal relation between currently permitted exposure concentrations and increased risk of lung disease*, i.e., it fails to show that any hazard exists if current standards are enforced. This omission invalidates the QRA's conclusions about material impairment at current exposure levels (since it has not shown that any impairments are caused by currently permitted exposure levels. It *assumes* that this is the case, and uses statistical regression models to quantify the assumed relationship between mean cumulative exposure and risk of various lung diseases. But this is not valid evidence that a true hazard exists, or that mean exposure levels – rather than much higher past

individual exposures such as those in the right tail of Figure 1 – cause any increase in risk.) Failure to identify a hazard from currently permitted exposure levels likewise invalidates the QRA's projections of human health benefits from further reducing permitted levels. A QRA should include a hazard identification section, with a critical discussion of disease causation by RCMD and other risk factors (such as cigarette smoking or diesel exhaust exposure) associated with RCMD exposures and lung diseases.



Frequency distribution of respirable dust concentrations in operating coal mines.

Figure 1: Past dust concentration distributions include a right tail that greatly exceeds presently permitted levels. *Source:* Jacobson, 1970

- *No uncertainty characterization.* The QRA does not show confidence intervals for its risk estimates, or characterize the effects of data and model uncertainties on its risk estimates. It does not candidly inform readers about how likely its conclusions are to be mistaken, or about the potential for the proposed rule to do more harm than good to human health (e.g., by forcing mine owners to adopt practices with lower means but higher variance in exposures). This one-sided presentation of information (showing substantial health benefits from the proposed rule as the *only* possibility, with no disclosure of relevant uncertainties or trade-offs) might be expected in an advocacy piece attempting to build a case for tighter regulation, but it is inappropriate in a scientific QRA document that should inform regulatory policy, but not seek to manipulate it or to play an advocacy role in building a case for a specific policy position. From this

standpoint, the QRA does not appear to follow recommended good practice for presenting risk analysis information and uncertainties (e.g., Jardine et al. 2003).

- *Flawed and biased statistical analyses.* The parts of the QRA process that MSHA has not omitted – exposure assessment and exposure-response modeling – are fatally flawed. The lengthy discussion of exposures focuses on *estimated mean cumulative exposures*. These are inappropriate for predicting lung diseases that occur only when exposure thresholds are exceeded (Porter et al. 2006, Cox 2009), since mean cumulative exposure can decrease even if the disease-relevant exposures (exceeding a threshold) increase. Moreover, the QRA fails to use statistical methods and models appropriate for the *uncertain exposures* in its data sets. This invalidates its statistical estimates of regression coefficients and its risk projections based on them. It introduces unknown biases into the regression models. MSHA also makes upward adjustments in exposures (since some past exposure values may have been under-estimated), but fails to recognize that, if past exposure values were truly under-estimated, then past regression coefficients linking estimated exposures to disease risks will have been correspondingly over-estimated. Adjusting one but not the other biases MSHA's risk estimates upward. Finally, MSHA's QRA mistakenly refers throughout to regression results as being "exposure-response relations," although no valid (causal) exposure-response relationship has been quantified, or shown to exist. MSHA's use of regression models to attribute miner lung diseases to RCMD is unjustified, and this invalidates all of its quantitative conclusions about risks.

To overcome these flaws, we recommend that any future or revised QRA should include a hazard identification section that neutrally summarizes and evaluates evidence on whether currently permitted exposure levels are already fully protective; validated causal models of exposure-response relations; exposure measures that emphasize high (e.g., above-threshold) individual exposures and uncertainty and variability in individual exposures; and an uncertainty characterization section that candidly reveals the potential of the proposed rule to shift exposure distributions without creating any human health benefits from reduced disease risks.

Introduction

This report evaluates MSHA's October, 2010 report, *Quantitative Risk Assessment in Support of Proposed Respirable Coal Mine Dust Rule* (www.msha.gov/regs/QRA/CoalDust2010.pdf), focusing on the following aspects:

(a) Do the conclusions of the analysis follow from its assumptions, data, and models?

(b) Are the QRA's assumptions, models, and methods appropriate for representing and interpreting present knowledge about how exposures to mineral dusts cause lung diseases?

(c) Does the QRA provide valid health risk assessment conclusions and projections of how the proposed rule would change current risk levels?

(d) Are the QRA's conclusions and estimates of current risks, and of potential benefits from the proposed rule, robust (i.e., relatively insensitive) to plausible changes or needed corrections in the QRA's data, assumptions, and methods change its?

Unfortunately, we conclude that the answer to each of these questions is: No. We organize the following comments on the QRA around the traditional steps of quantitative health risk assessment: hazard identification, exposure assessment, exposure-response modeling, and risk and uncertainty characterization. General comments on these areas are followed by some more specific comments on various technical aspects of the QRA..

General Comments on Hazard Identification

MSHA's QRA Omits the Hazard Identification Step and Fails to Show that Currently Permitted Exposure Levels Cause Any Increase in Disease Risk

The hazard identification step in health risk assessment has been described as follows:

> "Hazard Identification is the process of determining whether exposure to a chemical agent can cause an increase in the incidence of a particular adverse health effect (e.g., cancer, birth defects) and whether the adverse health effect is likely to occur in humans. The process examines the available scientific data for a given chemical (or group of chemicals) and develops a weight of evidence to

characterize the link between the negative effects and the chemical agent." (www.epa.gov/oswer/riskassessment/human_health_toxicity.htm)

MSHA's QRA does not present such a hazard identification step for RCMD, discussing the weight of evidence for or against the hypothesis that currently permitted exposure levels cause increased risk of lung diseases among miners. Specifically, it does not discuss the evidence for the alternative hypothesis that only exposures in excess of a certain threshold (for concentration and/or duration) pose a risk of progressive lung diseases, as suggested by some previous studies of epidemiology and experimental data on mineral (e.g., quartz) dusts and lung disease (Porter et al. 2004, 2006). Without this step, the QRA's conclusions about projected risks from current mean exposure levels and projected health benefits from tighter regulation are no more credible or less speculative than any other statistical association-based projections, such as that decreases in per capita consumption of butter cause decreases in teen pregnancy. The problem is that statistical associations between historical trends do not provide a valid basis for the causal interpretations and projections presented in the QRA. A hazard identification step that reviews current knowledge of how mineral dusts affect risks of lung disease, and that relates this knowledge to currently permitted exposure levels and thresholds for disease causation, should be an essential component of a QRA for RCMD. A weight of evidence assessment of whether currently permitted exposure levels cause increased incidence of lung diseases should be a prerequisite for further QRA. If the answer is no, then there is no risk to assess.

Hazard Identification Should Discuss Exposure Thresholds in Light of Mode of Action

We believe that a useful QRA for RCMD should reflect current biological understanding of the inflammatory mode of action (MOA) for lung diseases induced by inhalation of particulates, including coal dust (e.g., Azad et al. 2008, Schins and Borm 1999, Cox 2009, Tuluce et al.2010). Briefly, this understanding of MOA shows that only sufficiently great exposure concentrations and durations induce an influx, and a shift in phenotype, of alveolar macrophages to the lung, triggering a cascade of changes that create an excess of oxidants over anti-oxidants and a high-ROS (reactive oxygen species) and high reactive nitrogen species (RNS) lung environment. The resulting oxidative stress can cause lung diseases in susceptible

individuals (those with inadequate defensive (e.g., antioxidant) and repair capabilities). These adverse effects do *not* occur at lower exposure concentrations, since "Tissues and cells respond to mild oxidative stress by increasing antioxidant defenses. However, high levels of ROS/RNS may overwhelm antioxidant defenses, resulting in oxidant-mediated injury or cell death" (Comhair and Erzerum 2002).

This MOA information should be included in the hazard identification discussion for RCMD. It indicates that *high exposure concentrations (e.g., those that overwhelm antioxidant defenses) are more dangerous than low exposure concentrations, for the same cumulative exposure.* More detailed investigation of specific lung diseases, such as silicosis (Porter et al. 2004, 2006) or COPD (Cox 2009) suggests that there is an *exposure threshold* for disease causation, below which mineral dust exposures are not expected to cause excess risk of inflammatory, progressive lung diseases. This implies that it is crucial for the hazard identification component of the QRA to discuss evidence on whether currently permitted exposure levels are already below the levels that cause increases in disease risk. (Fitting regression models to past exposure estimates and disease rates, as in the current QRA, does not address this crucial question, since past exposures presumably contained a wide variety of individual exposures, including some that were well above currently permitted levels.)

General Comments on Exposure Assessment

MSHA's exposure assessment and modeling, which take up most of the current QRA, makes essential use of the following simplifying assumptions.

- Risk depends only on cumulative exposure
- Estimated risk depends only on estimated *mean* cumulative exposure levels, and not on the rest of the frequency distribution of individual exposures around these estimated means.
- Changes in mean exposure levels can be used to predict changes in risk.
- The risk that is attributed to exposure would be reduced by reducing exposure
- Regression and analysis of covariance (ANCOVA) are appropriate statistical tools for correcting potential biases and estimating effects on risk of exposure and other covariates.

The following comments show that each of these assumptions, which underlie the rest of the analysis, is incorrect. Thus, the QRA's exposure estimates are not appropriate for quantifying risks from currently permitted exposure levels, or for correctly predicting how, if at all, further reductions in currently permitted exposure levels would reduce human health risks. The QRA's attempted use of its exposure estimates for these purposes is not valid.

Use of Cumulative Exposure is Inappropriate

MSHA's analysis explicitly assumes "that health risks associated with RCMD exposures are a function purely of cumulative exposure, regardless of any peaks or valleys in the intensity of dust concentrations that have been experienced over time" (p. 29). Although the QRA recognizes (p. 59) that this assumption is not necessarily valid, it states that "None of the published exposure-response models, however, take any account of exposure patterns. Therefore, this QRA has made no attempt to quantify their effects." However, assuming that risk depends only on cumulative exposure is not justified for inflammatory lung diseases. It is incompatible with current mode of action (MOA) knowledge about inflammatory lung disease causation (e.g., Azad et al. 2008, Schins and Borm 1999, Cox 2009, Tuluce et al.2010) and is inconsistent with both animal and human data on exposure thresholds for progressive lung diseases caused by mineral dusts (e.g., Porter et al. 2004, 2006 for crystalline silica).

Contrary to MSHA's assumption, what matters most for health risk assessment is the prevalence of exceptionally high – not average or cumulative – exposures in the work place (i.e., the upper tail of the frequency distribution of individual exposures). Specifically, any useful exposure assessment should address: *What fraction of workers receive exposures high enough to cause lung disease* (e.g., by creating a high-ROS lung environment and persistent oxidative stress)? The current QRA does not address this key question. Knowledge of estimated mean exposures does not answer it. Thus, although the QRA acknowledges that it does not consider peaks in exposure intensities (i.e., in dust concentrations), and that it "has made no attempt to quantify their effects," it is precisely these unquantified effects that determine risk, assuming that the peaks of interest are those that are high enough to cause lung disease. Thus, the QRA omits those aspects of exposure that are relevant for determining risk.

The fact that MSHA's QRA does not quantify relatively high (disease-relevant) exposures, nor model how they would change if the proposed rule were adopted, is sufficient reason by itself to conclude that the QRA's predictions and conclusions lack predictive validity, and that they should not be used as a basis for risk management policy decisions. Although there are many other fatal flaws, this one directly affects the practical value of the substantial effort that the QRA has spent on estimating and modeling mean exposure levels. Mean exposure levels are simply irrelevant for predicting risks, as a well-conducted hazard identification would probably have made clear, and as discussed further below. The QRA's risk predictions based on estimated exposure levels lack any validation as being accurate projections of real-world effects of current or potential future exposure distributions.

Mean Exposures Do Not Predict Risk

The QRA focuses on estimating *mean* cumulative exposure concentrations. But this does not quantify the relatively high individual exposures that are most important for predicting risk correctly. There are thus two related problems with the QRA's exposure metric: both its use of *cumulative* exposures (ignoring peaks, and the fact that a higher concentration for a shorter time may cause diseases even though the same cumulative exposure spread over more years would not); and its focus on *mean* exposures, ignoring the variance of exposure and the occurrence of exceptionally high (far above the mean) cumulative exposures. By failing to quantify exceptionally high exposures and exposure intensities, the QRA loses any ability to explain or predict correctly the real human health risks from any specified level of exposure. The QRA's tables and figures for estimated and predicted risks are irrelevant for quantifying real-world effects of exposures on risk, because they falsely attribute to estimated mean exposure levels the health effects produced by higher exposures (and, indeed, by the full joint frequency distribution of exposures to multiple risk factors).

Changes in Mean Exposure Levels Do Not Predict Changes in Risk

Similar problems, of misattributing past health effects to estimated mean exposure levels instead of to high exposures, arise when the QRA attempts to predict how *changes* in mean

exposure levels will change risks. Risk predictions and policy evaluations based on changes in *mean* exposure levels are seriously misleading when it is changes in *high* exposure concentrations (above the mean) that primarily affect risk.

MSHA's assumption, and repeated assertions, that reducing mean exposures will reduce risk, are not valid. For example, reducing mean exposure levels of RCMD by making already-low (e.g., below the disease causation threshold) exposures even lower might create no additional human health benefit. Conversely, reducing the variance in exposures (by reducing individual exposure fluctuations and outliers that are far above the current mean) could reduce health risks without changing the current mean exposure level. Indeed, a policy can reduce the mean while increasing the variance of exposure concentration, thereby reducing mean exposure while increasing risk. Thus, *there is no clear (or necessary) relation between changes in mean cumulative exposures and changes in human health risk.* Changes in means do not determine corresponding changes in the entire exposure distribution, and thus are inappropriate for predicting changes in risk.

Example: A Rule that Decreases Mean Exposure Can Increase Risk

Suppose that workers can be exposed to any of three concentration levels, 1, 2, or 3 mg/m^3, with corresponding risks of 0.1, 0.4, and 0.9, respectively. (These numbers are for purposes of illustration only.) Initially, among 30 workers, 20 are exposed to the middle level, and 5 to each of the other two levels. The average exposure level for these workers is thus (5*1 + 20*2 + 5*3)/30= 60/30 = 2 mg/m^3, and the excess risk is 13 excess cases among 30 workers, as shown in the following table.

x = exposure concentration (mg/m^3)	p = risk at exposure concentration x	n = number of workers exposed to x	Expected excess cases
1	0.1	5	0.5
2	0.4	20	8
3	0.9	5	4.5
			Total = 13

Now, suppose that a new rule reduces the mean exposure concentration, through rigorous dust control measures that result in lower exposures for most workers, but in higher exposures for workers in locations

where implementation or compliance fail. The following table shows the new exposures and risks. Average exposure concentration has *decreased*, from 2 mg/m^3 to (16*1 + 14*3)/30 = 58/30 = 1.93 mg/m^3, but risk has *increased*, from 13 to 14.2 (= 16*0.1 + 14*0.9) excess cases among 30 workers.

x = exposure concentration (mg/m^3)	p = risk at exposure concentration x	n = number of workers exposed to x	Expected excess cases
1	0.1	16	1.6
2	0.4	0	0
3	0.9	14	12.6
			Total = 14.2

This example illustrates the importance of quantifying not just the *mean* exposure concentration before and after a proposed rule is implemented, but how the *frequency distribution* of exposures will change. It is essential to model the effects of changes on the rest of the exposure distribution (or at least its right tail, if a threshold dose-response relation is used) before valid predictions can be made about how a change in exposures will affect risk. Without such information, risk managers have no valid basis for predicting whether proposed changes will increase or reduce risk. This is one reason that MSHA assumption that a reduction in mean exposure will reduce risk is not valid.

The QRA's Statistical Analyses and Claims are Incorrect and Biased

MSHA's QRA acknowledges that, "Applying an exposure-response model to an occupational average exposure level fails to account for risks in more specific environments when the exposure is above the occupational average" (p. 41). (This is one of many instances in which the QRA emphasizes that true exposures may exceed estimated ones, but without mentioning that the reverse is equally true, and that it is important to account for both to correctly estimate risks at different exposure levels.) However, it then offers the following reassuring, but mistaken, claim (p. 41): "Indeed, when exposure-response relationships are curved upwards as those shown above, *evaluating risk at the average exposure level will always underestimate average risk*" (emphasis added). This claim is false – it is the reverse of the truth, which is that evaluating risk at average exposure levels over-estimates average risks – when, as

in the QRA, "evaluating risk at the average exposure level" is done by estimating excess cases of lung disease at estimated average exposure levels.

Example: Evaluating Risks at Average Exposure Levels Overestimates True Risks

Consider the upward-curving exposure-response relation is: $Risk = (0.1*Exposure)^2$. Suppose that half of a group of workers have zero exposure (and hence $(0.1*0)^2 = 0$ individual risk) and that the other half have an exposure of 10 mg/m^3 (and hence an individual risk of $1^2 = 1$). The average exposure level is $(0 + 10)/2 = 5$ mg/m^3. The average risk, evaluated at the average exposure level of 5 mg/m^3, is: $(0.5*0 + 0.5*1) = 0.5$. This is *twice* as high as the true risk at 5 mg/m^3, namely, $(0.1*5)^2 = (0.5)^2 = 0.25$. Contrary to the QRA's suggestion that true risks are always underestimated when average exposures are used, evaluating risk at the average concentration *overestimates* the true risk at that concentration, since it attributes to the average concentration level risks that are primarily caused by greater-than-average exposures. The QRA's claim to the contrary is misleading, and might make think

More generally, MSHA's QRA applies statistical techniques, including ANCOVA and regression, that are inappropriate when explanatory variables – such as true exposure concentrations, in this case – have not been precisely measured. Although appropriate statistical methods for such "errors in variables" or measurement error situations are well developed and readily available (e.g., Carroll 1989, Murad and Freeman 2007), the QRA does not use them. This omission introduces unquantified, potentially large biases (e.g., Steenland et al. 2000, Hossain and Gustafson 2009) into the QRA's qualitative and quantitative conclusions, rendering them unreliable and useless for valid statistical inference. Like some of the other methodological flaws in the QRA, this one is sufficient, by itself, to make MSHA's risk projections unreliable: they can be incorrect in either direction (too high or too low), by an unknown amount, due to the failure to model measurement errors in RCMD exposures, co-exposures, and covariates.

MSHA's Exposure Estimates and Risk Estimates are Biased Upward

Consistent with the "proofiness" approach of selecting and adjusting facts and data to fit a preferred policy or advocacy position (e.g., tighter regulation), MSHA's QRA unjustifiably "adjusts" exposure estimates to correct for possible occasional underestimation of true exposures

(p. 25 and Appendix F), but without performing any symmetrical adjustments to correct for equally possible occasional overestimation of true exposures. *This adjustment process is statistically invalid*, for at least the following reasons.

- *MSHA's "adjustment" process systematically overestimates exposures, even when the original exposure estimates are unbiased.* For example, consider the case where operator and MSHA samples are both unbiased for at least some WLs and job categories, with independent, identical, normally distributed sampling errors. Even in this ideal case of unbiased estimates, MSHA's "adjustment" procedure still adjusts half of the values upward (since there is a 50% probability that operator sample means will exceed MSHA sample means when neither one is biased). As the QRA only makes upward adjustments, and never symmetrical downward ones (in effect, ignoring or denying the obvious possibility that the operator samples may sometimes be too high), the net effect of these one-sided adjustments is to systematically over-estimate exposures "Adjusting" exposure estimates upward without correcting for resulting biases amounts to deliberate manipulation of data to produce inflated exposure and risk estimates.
- *When exposure estimates are "adjusted" upward, then potency estimates should be symmetrically counter-adjusted downward, to avoid biasing risk estimates upward.* MSHA has failed to make these needed counter-adjustments in its exposure-response models. The QRA's combined use of "adjusted" exposure estimates together with unadjusted exposure-response models is not valid, and biases all of its projected risks upward. To see why, consider the simplified model $Risk = b*Exposure$, where b is a potency parameter estimated from *Exposure* and *Risk* data. If *Exposure* has truly been systematically *under-estimated* in MSHA data, as the QRA argues (e.g., because mines improve when they are being inspected), then b must have been correspondingly *over-estimated* (since observed excess diseases will have been caused by higher-than-estimated exposure levels). Subsequently adjusting *Exposure* upward without symmetrically adjusting the estimated potency parameter b downward is inconsistent, and inflates estimates of risk.

Perhaps recognizing that its adjustment procedures lack statistical justification, the QRA notes (footnote, p. 26) that "MSHA believes that its use of operator data in the AS estimation procedure as applied to specific WLs serves, on balance, to reduce rather than increase the potential for overall bias." While MSHA may entertain such beliefs, it is clear

mathematically that the adopted procedure incorporates at least the two sources of upward bias just described. Neither source of bias has been corrected for in MSHA's analysis, suggesting that its stated beliefs on this point follow only from wishful thinking

- *Deliberate exclusion of abatement-related exposure estimates biases exposure estimates and risk estimates upward.* Abatement is real. Suppose that a mine moves through cycles of gradually increasing dust levels followed by abatement measures that restore dust levels to well below the permitted exposure limit. For the QRA to include periodic and support measurements (p. 24), while excluding post-abatement measurements, deliberately biases average exposure estimates upward (by excluding the low values, which are just as real as the other values). If this biased procedure is used, then the exposure-response relations used to project risk should be adjusted to reflect the fact that the exposure input data fed into them have been censored to exclude low (abatement) values.

General Comments on Exposure-Response Assessment

Exposure-Response Relations Used in the QRA are Unjustified

The QRA relies on previously published, peer-reviewed statistical regression relationships to estimate health risks caused by currently permitted mean RCMD exposure levels, and to project reductions in adverse health effects that it claims would be caused by reducing mean exposure levels. Unfortunately, these are not valid applications of the previously published regression relationships.

Although one can always regress adverse health effects against other quantities, including RCMD exposures (or, for that matter, annual sales of golf balls), and thereby produce significant-looking regression coefficients, if the variables involved have trends, such regression does not establish a valid exposure-response relation between them. A valid exposure-response relation is supposed to quantify the probability of an adverse health response *caused* by each level of exposure, i.e., the *toxicity* of exposure, not just the regression coefficient (which merely reflects the slope of the scatter plot plotting adverse health effects against one or more other

variables); see e.g., www.epa.gov/oswer/riskassessment/human_health_toxicity.htm. Causal relations and regression relations between exposure and response variables may be entirely different. It is only the former, and not the latter, that is properly referred to as an exposure-response relation. Throughout its discussion, MSHA has incorrectly called regression relations and coefficients "exposure-response relations." In reality, however, MSHA's QRA has not identified or quantified any true (causal) exposure-response relation, but has incorrectly used statistical regression relationships to fill this role.

Example: Regression Does Not Yield a Valid Exposure-Response Relation

To illustrate why it is inappropriate for MSHA's QRA to treat regression models as if they were true (causal) exposure-response relations, or as if they could be used to predict how changing exposures would affect risk, consider the following simple, two-equation, example of a causal model:

$$Risk = Age - Exposure$$
$$Exposure = 0.5*Age \qquad \text{(Simple causal model)}$$

In this pair of causal equations (called "structural equations" in causal modeling), changing variables on the right side of an equation causes corresponding changes in the variable on its left side. Thus, the first equation shows that *Risk* increases with *Age*, but decreases with *Exposure* (for any given value of Age). The second equation shows that *Exposure* increases with *Age*. But, substituting the second equation into the first yields the following non-causal ("reduced form") regression relation between *Exposure* and *Risk*:

$$Risk = Age - Exposure = Age - 0.5*Age = 0.5*Age = Exposure, \text{ or simply}$$

$$Risk = Exposure \qquad \text{(Regression model).}$$

The regression model relating observed values of *Risk* and *Exposure* shows a direct positive relation (a positive statistical association) between them. Yet, an increase in *Exposure* would not increase *Risk*: as shown by the first causal equation (*Risk = Age – Exposure*), an increase in *Exposure* would decrease *Risk*. The causal and regression relations are entirely different, and even have opposite signs.

MSHA's QRA for RCMD makes the fundamental error of treating regression models *as if* they were causal models, and then using them to project how changes in *Exposure* would change *Risk*, even

though the regression models used in the QRA are not causal models, and are not valid for this purpose. As a result, the QRA's projected changes in risk from the proposed rule are as meaningless as the (mistaken) prediction that increasing *Exposure* will increase *Risk* in the above simple example.

MSHA's Exposure-Response Modeling and Risk Estimates are Biased

The fact that MSHA has not quantified or validated a genuine exposure-response relation makes the rest of the analysis and the discussion of results and conclusions somewhat moot: they are based on such a fundamental error (mistaking regression results for causal exposure-response relations) that they lack any validity for use in risk assessment or risk management. However, even if this fatal flaw is ignored, there are numerous other important errors and biases in the exposure-response modeling, including the following.

- *Unmodeled variance in exposure durations creates unquantified biases into MSHA's estimated exposure-response relations.* Steenland et al. 2000 caution that "Assignment of job-specific mean exposure levels from a sample of workers causes an upward bias in the estimated exposure-response trend when there is little variance in the duration of exposure but causes a downward bias when duration has a large variance. This bias can be substantial (e.g., 30-50%)." MSHA's QRA neglects to correct for the effects of variance in exposure durations within and between occupational categories, leaving, the direction and magnitude of biases in its risk projections are unknown. This makes them unsuitable for use in supporting regulatory risk management decisions. Appropriate statistical methods for quantifying and reducing biases from using job-specific (or WL-specific) mean exposure levels for all workers in the group (e.g., Kim et al. 2010) have not been used.
- *Failure to model errors in exposure estimates bias MSHA's risk estimates.* The regression models used by MSHA to model exposure-response relations have not been corrected for errors in exposure estimates. This biases the estimated exposure-response relation. The general statistical issue has recently been described as follows:

  "In most epidemiological investigations, the study units are people, the outcome variable (or the response) is a health-related event, and the explanatory variables are usually environmental

> and/or socio-demographic factors. The fundamental task in such investigations is to quantify the association between the explanatory variables (covariates/exposures) and the outcome variable through a suitable regression model. The accuracy of such quantification depends on how precisely the relevant covariates are measured. In many instances, we cannot measure some of the covariates accurately. Rather, we can measure noisy (mismeasured) versions of them. In statistical terminology, mismeasurement in continuous covariates is known as measurement errors or errors-in-variables. *Regression analyses based on mismeasured covariates lead to biased inference about the true underlying response-covariate associations*." (Hossain and Gustafson, 2009, emphasis added).

Appropriate statistical techniques to correct for such biases are widely available, as discussed in the preceding references, but have not been used in the QRA or the models on which it depends (e.g., the Attfield and Kuempel model in Appendix K).

- *Unquantified bias in risk projections due to omitted confounders.* Pages 134-135 of the QRA acknowledge that coal rank is confounded with other geographic differences. In fact, there are numerous potential confounders that have been omitted from the model. For example, is poverty (or lower income or education or socioeconomic status, etc.) associated both with work in locations with higher dust concentrations, and also with less access to health care, or with increased risk of respiratory diseases, even in the absence of coal-related occupations? If so, how much of the effect attributed to coal dust in the QRA is actually caused by such confounding factors? The QRA and its supporting epidemiological models do not answer these questions. Hence, they do not provide a valid basis for attributing differences in disease rates to differences in exposures, nor for interpreting statistical projections of lung disease mortality and morbidity as indicating causal effects of exposures.
- *Unquantified bias in risk projections due to model specification errors.* The exposure-response relations used in the QRA do not pass basic consistency checks for yielding valid risk predictions. For example, as stated on p. 135, "Even with cumulative coal mine dust exposure set to zero, the Attfield-Kuempel exposure-response model produces relative risk estimates of 4.4 and 1.2 for miners regionally associated with anthracite and high rank bituminous coal... This suggests... that the relative and excess risk shown for NMRD mortality at WLs with anthracite and high rank bituminous coal should be interpreted with

extreme caution." This is an understatement. A model that attributes a large relative risk (4.4) to coal in the absence of any exposure is not suitable for risk assessment or for supporting regulatory risk management decision-making. Appendix K's subsequent argument that perhaps regional effects can be cancelled out, and that "The portion that is not cancelled is attributable to occupational exposure" is completely specious. (The part that does not cancel could be due to model specification errors, unmodeled interactions among variables, omitted covariates and confounders, or other factors that are in no way "attributable to occupational exposure".)

- *Unquantified upward bias in risk estimates due to ignored model uncertainties.* The results of different models used by the QRA conflict with each other, showing that they cannot all be valid. For example, "The COPD/17 model predicts a 15% increase (RR=1.15) in the risk of death from COPD for 80-year old former miners who have been exposed for 45 years at an average 1.0 mg/m3. For the same exposure, the Attfield/Kuempel model predicts a 34% increase in risk." (p. 40) Such disagreement among predictive models indicates that there is substantial *model uncertainty* (i.e., uncertainty about which model, if any, might be correct). The QRA does not apply appropriate statistical techniques for estimating and correcting for such model uncertainty (such as Bayesian Model Averaging). This typically increases the rates of false-positive associations (mistakenly concluding that risks are significantly elevated when they are not); see e.g. the discussion of false positive associations between air pollution and mortality at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.164.6048.

In summary, the QRA throughout mistakenly treats biased and mis-modeled statistical associations and regression relations as if they were known to be valid causal relationships. For example, Part 1 of the QRA concludes that "This part of the QRA has shown that coal mine dust exposure at currently experienced occupational averages poses significant risks of material impairment in nearly all occupational categories." This misinterprets what has actually been shown, which is only that lung diseases occur in coal miners in nearly all occupational categories. How this would be affected by reductions in exposure levels has *not* been quantified, although that is what the QRA claims to have done. The associations and regression models used are not causal models. As previously discussed, reductions in mean exposures can result

from changes in exposure distributions that increase risk; hence, regression relations that show an always-increasing relation between them may be simply irrelevant for accurate risk prediction.

General Comments on Risk Characterization

The "Risks" Quantified by MSHA Do Not Exist

In light of its missing hazard identification, inappropriate use of cumulative exposure, failure to model or correct for errors in estimated mean exposures, failure to model or correct for variance in exposure concentrations and durations, use of one-sided "adjustments" and exclusions of data that bias estimated exposures and risks upward, improper use of regression relations in place of true (causal) exposure-response relations, and failure to identify or quantify any causal relation between current (or proposed future) permitted levels of RCMD and any adverse human health effects, it may seem superfluous to add that MSHA's QRA has not quantified any real risks to human health. Nonetheless, figures and tables can look impressive when their basis is not clear. It is this impressive and convincing look conveyed by quantitative displays that sometimes allows "proofiness" to substitute for sound risk analysis in shaping perceptions and steering risk management policies. Thus, it is perhaps worth emphasizing that the figures and tables in the QRA, purporting to show that risks of adverse health effects and material impairments increase with mean RCMD exposures, and would be substantially reduced by adopting the proposed reductions in mean RCMD exposures, are entirely fictitious. They are based on a series of unvalidated modeling assumptions, incorrect analyses, imaginary relationships, and logically inconsistent and *ad hoc* risk attribution procedures that have no known relation to reality. Despite their impressive appearance, the QRA results have no legitimate claim to be considered even approximately (or directionally) correct, or to have any specific relation to real-world risks and health effects of exposures. They are useless for informing risk managers about the probable size of real risks, or for predicting how those risks would change (if at all) if the proposed rule were adopted.

Use of Single-Shift Sampling Would Further Exacerbate Poor Risk Analysis

One of the flaws in MSHA's QRA, as already discussed, is that it fails to model the *variance* in RCMD exposures and exposure estimates. The true risk from a frequency distribution of exposure concentrations will typically depend on the variance of the distribution, as well as its mean, contrary to the QRA's assumptions. Because the QRA ignore variance, it is unable to quantify how means and variances trade off against each other, or interact with each other, in determining risks. Nonetheless, one of the proposals that the QRA has been used to support is the suggestion that *single-shift sampling* should be used to assess compliance with the proposed new standard. In the absence of explicit modeling of the relation between exposure means, variances, and risks, this is an altogether irresponsible suggestion. If the goal is to assure that the exposure cumulative distribution function (CDF) lies to the left of some desired target CDF, for example (guaranteeing that high exposures are rare), then this can be accomplished by applying appropriately designed statistical tests (e.g., extended Kolmogorov-Smirnov tests, http://sticerd.lse.ac.uk/dps/em/em433.pdf). But such tests are best performed by collecting more data, not less, and by adjusting for serial correlations that may be apparent in multiple-shift samples but not in single-shift samples. Using data from only a single shift appears designed to increase the probability of type 1 error (falsely concluding that a mine is out of compliance), since sample variance around the mean is maximized by using only one shift, yet no compensating increase in sample means has been proposed to correct for this increase in sample variance. In short, the characterization of mines as being "out of compliance" or posing unacceptably large risks is likely to be made more error-prone, not less, by the use of single-shift sampling. This use should therefore be discouraged. A sound analysis, applying appropriate tools of statistical decision theory, would require a much more explicit discussion of objectives, statistical tests, and relations between sample variances and risks, than the QRA provides.

General Comments on Uncertainty Characterization

MSHA's QRA Omits Uncertainty Characterization, Needed for Responsible Decision-Making

The exposure-response regression model results presented in Figures 10 and 11 of the QRA, and in its appendices, completely fail to characterize uncertainties about estimated and predicted risks. This violates recommended principles of practice for good risk analysis (e.g., Jardine et al. 2003). A risk assessment without uncertainty characterization is importantly incomplete. It is not appropriate to use such an incomplete assessment to guide regulatory decision-making, as it does not candidly disclose the uncertainty information needed to support a well-informed decision.

The curves in Figures 10-13 of the QRA do not show confidence bands, so there is no way to use them to determine what level of confidence, if any, should be placed in different purported risks estimated from them. No sensitivity analyses or other techniques (e.g., Bayesian model averaging) have been used to show how *model uncertainties* affect the conclusions. As a result, risk managers are presented only with one-sided information, consistently suggesting a strong positive relationship between RCMD and risk of lung disease, without any of the important qualifying information about confidence and uncertainty (e.g., do better models show no such relationship?) needed to support well-informed decision-making. The QRA does not disclose the value-of-information from collecting further data, but presents its preferred answers as if there were no doubt about their correctness and no possibility that better information would change the recommended course of action.

By presenting only upward-sloping curves, the exposure-response regression models seem to make a promise to risk managers – that reducing mean RCMD concentrations will necessarily decrease health risks (perhaps dramatically) – that goes well beyond what a responsible interpretation of the modeling and data could justify. In reality, mean RCMD concentrations have fallen in the past, yet health risks have *increased* in some age categories (e.g., Bang et al. 1999, MMWR 2009). This is contrary to the predictions of the QRA's risk modeling, and shows that the model predictions are far from certain, and may be completely incorrect. Responsible uncertainty characterization requires informing readers of the QRA about the probability or potential that its conclusions and recommendations are mistaken, and that

passing the proposed rule could produce no additional health benefits, or could even do health harm (e.g., by encouraging mining of coals with lower mean RCMD concentrations but more hazardous dusts). By failing to show any possibilities except that lower mean RCMD concentrations would produce lower risks (contrary to historical experience), the QRA fails to illuminate the risks and trade-offs involved in different risk management policy options. It thus fails to provide the crucial, policy-relevant information that uncertainty characterization should provide. The QRA should be revised to include the missing step of uncertainty characterization.

Conclusions

Our main conclusions, based primarily on the preceding comments, as well as on the more detailed comments in the following sections, are as follows.

- *The premises of MSHA's QRA analysis (its assumptions, data, statistical methods, and models) do not imply its conclusions.* For example, the QRA uses non-causal models (e.g., risk attribution models, regression models) to draw unwarranted causal conclusions – specifically, that currently permitted exposure levels place workers at risk, and that reducing exposures would reduce risk.
- *The QRA's premises are not sound*, and they do not accurately reflect present knowledge, data, and risk assessment methods and models for lung diseases caused by respirable coal mine dust (RCMD). For example, the assumption that risk depends only on mean cumulative estimated exposure is inconsistent with available mode of action evidence.
- *The QRA's premises do not provide a sufficient basis for valid health risk assessment*, nor for projecting how the proposed rule would change current risk levels. The risk estimated and predicted by the QRA are fictitious, based on flawed assumptions and models and incorrect statistical methods; they have no known relevance to real-world health effects.
- *Improving the risk assessment methods, as we have recommended, could completely change (or even eliminate) estimated current risks, and estimated benefits from the proposed rule.*

More Detailed Technical Comments

p. 19 *"This QRA addresses imbalances in the number of available samples by developing separate exposure estimates for each WL. Results are then aggregated by occupational category, assigning equal weight to the mean dust concentration observed at each WL."*

Comment: It is not clear how equal weighting "addresses the imbalances," rather than exacerbating them. No statistical justification is given for the use of equal weights at different WLs. (For example, it is not shown that this procedure reduces, rather than increasing, errors in cumulative exposure estimates.) Assigning equal weights to sample mean dust concentrations observed at different WLs does not correct for differences in sample variances across WLs. In general, techniques that minimize mean squared errors in estimates, such as Generalized Least Squares (GLS) or weighted least squares (WLS), require that sample means be combined using weights that reflect sample variances at different locations. By ignoring sample variances in its weighting scheme, the QRA departs from standard statistical procedures, and possibly introduces unknown errors and biases into its exposure estimates. In particular, MSHA has supplied no proof or reason to believe that its equal-weighting procedure produces unbiased estimates of cumulative exposures for workers.

Comment: The QRA's exposure estimation emphasizes quantifying a weighted sum of mean concentration levels. This has no known or validated relevance for predicting how the proposed rule will affect human health risks, which presumably depend on changes in the *distribution* of individual exposures, rather than only on *mean* exposure levels. The QRA's assumption to the contrary lacks justification. We believe that the weighted mean exposures quantified in the QRA have little or no relevance for predicting effects of the proposed rule on human health, as they do not indicate how the right tail of the exposure distribution – the part that is most relevant for risk – will be affected, or how the resulting changes in exposure will affect risk. (For example, although Appendix H distinguishes between shifts on which the new standard would or would not be exceeded, it does not inform risk managers whether eliminating shifts that exceed the *existing* standard would suffice to eliminate excess risk.) The QRA should be revised to discuss

what is known about the mode of action of dust-induced respiratory diseases, and the extreme importance of higher-than-average exposure concentrations in determining risk.

pp. 19-20: [A]n analysis of covariance (ANCOVA) was performed to obtain unbiased estimates of the mean dust concentration associated with each occupation on shifts sampled by MSHA."

Comment: MSHA suggests that ANCOVA yields unbiased estimates of mean dust concentrations in this data set. No justification is provided for this claim. It is contradicted by statistical theory showing that *ANCOVA yields biased results* in the presence of measurement errors, or "errors in variables," such as those arising from job classification errors and ambiguities, imperfect compliance with applicable standards, or unmodeled heterogeneity within occupational categories (see e.g., Carroll 1989; Raaijmakers and Pieters 1987 for the case where the dependent variable is linearly related to underlying variables that are measured with errors). Thus, the ANCOVA procedure adopted by MSHA "to obtain unbiased estimates" may instead create biases (*ibid*).

MSHA not only attributes to ANCOVA properties that it does not have (i.e., producing unbiased estimates despite realistic measurement errors), but also it fails to follow good statistical practice by performing and reporting the results of diagnostic tests to show whether the assumptions underlying ANCOVA are at least approximately satisfied in this application (e.g., by showing the results of tests for normality and homoscedasticity of errors in the dependent variable, or by correcting for errors in the independent variables and covariates). Although Table 52 presents results, it does not discuss the methods used to produce them. Standard packages typically use least-squares algorithms, which can give biased results. Other algorithms (e.g., method of moments, regression-calibration) have been developed to overcome the biases due to standard ANCOVA and regression in the presence of measurement errors (e.g., Murad and Freedman 2007), but there is no evidence that the QRA used these techniques. The QRA should be revised to specify how the ANCOVA analysis was performed, and should either justify or withdraw its claim that the analysis yields unbiased estimates.

p. 22, Table 8: "Examples of work locations showing significantly higher than normal average dust concentrations."

Comment: It is not clear whether or how the QRA corrected for multiple testing and multiple comparisons biases (which lead to false positives when researchers "cherry pick" the high or low results from a large number of statistical tests or comparisons, such as the large number of work locations considered in the QRA) before asserting that these work locations have "significantly" higher than normal average dust concentrations. (For example, did MSHA use step-down or other procedures to adjust the p-values used to define "significance" to correct for these biases? The QRA does not say.) Nor is it clear why the effects of work locations with significantly *lower* than average concentrations were not treated symmetrically, to arrive at an unbiased assessment of cumulative exposures.

p. 24: "Therefore, undue influence of measurements in earlier years constitute one source of potential bias with respect to estimates of current exposure conditions. ... One way to address these potential biases is to restrict the analysis to 2008 data..."

Comment: The QRA presents no justification, and no empirical validation, for selecting 2008 data as more relevant that earlier measurements for predicting future risks. The QRA uses exposure-response relations estimated from data collected in earlier years, which have no known predictive power for the specific exposure conditions in 2008. The claim that restricting the analysis to 2008 data is a way to "address" potential biases is asserted without explanation or justification, and without noting that such restriction can potentially create new, larger biases by deliberately giving undue – in fact, exclusive – influence to measurements in 2008. Recent methodological study of proportional hazards modeling with time-dependent exposures in case-control studies suggests that using all years of data, with subject-specific weights reflecting age-specific conditional probabilities of disease, significantly reduces bias compared to alternative approaches (Ledffondre et al. 2010). The QRA does not offer any justification for its claim that excluding data from years other than 2008 will reduce biases. The QRA should be revised to either justify or withdraw its suggestion that selecting 2008 estimates reduces prediction biases.

p. 29: "The objective in this portion of the QRA is to assess risks associated with actual current exposure levels."

Comment: Actual current exposure levels are not known and have not been quantified in the QRA. The QRA focuses on *estimated mean* exposure levels, not *actual* exposure levels. It does not address how frequently levels of exposure actually needed to cause disease occur (with or without the proposed rule). It uses exposure-response relationships developed for conditions that held decades ago, with unknown validity or relevance for actual current (or future) exposure conditions. For example, part of the risk attributed to RCMD in Appendix J would not exist in the absence of cigarette smoking. As noted by the QRA on p. 132, "[C]urvature in the joint exposure-response relationship expressed by Equations 12 and 13 amplifies the predicted response to RCMD exposures for smokers. (This is an inherent characteristic of the logistic model employed.)" However, the exposure-response relations used reflect smoking patterns that held decades ago; they have not been updated to apply to *current* (or relevant future) smoking patterns, following smoking cessation and smoke-free workplace initiatives introduced in the 1980s and 1990s. Hence, these old models attribute to RCMD risks that would hold only for smoking patterns that no longer exist. Moreover, according to the QRA's own assumptions, the risk associated with actual current exposure levels depends on actual past and future exposure levels of RCMD experienced by workers. But, actual future exposures (and past exposures other than in 2008) are not included in the model. Hence, it is misleading to suggest that the QRA assesses "risks associated with actual current exposure levels." What the QRA does instead it to assess *hypothetical* risks for *hypothetical* exposure scenarios, using decades-old exposure-response models that have not been validated for current actual exposures (and that also contain numerous errors and upward biases, as previously noted, such as those from unmodeled measurement errors or "errors-in-variables" in exposure and covariate estimates). It is inaccurate to state otherwise. The QRA should be revised to withdraw this suggestion that it has quantified "risks associated with actual current exposure levels.".

p. 30: "Since the models show risk increasing with age, that portion of the risk attributable to the accumulated exposure is obtained by calculating the difference in risk calculated with and without exposure."

Comment: No justification is provided for the claim that "the difference in risk calculated with and without exposure" is "that portion of the risk attributable to the accumulated exposure". The difference in risk calculated with and without exposure may be caused by many factors other than exposure. Among them are the following.

- *Other unmodeled covariates and confounders*, such as low income, low education, poor access to medical care, hazardous geographic location, etc., that might be associated with exposure and that contribute to increase risk. Attributing the difference in calculated (hypothetical) risks to differences in exposures will overstate the effect of exposure, if such unmodeled confounders are present.
- *Failure to use statistical models that correct for errors and variance in estimated exposures.* This might produce a difference in risk calculated with and without exposure, even in the absence of any real (causal) effect of exposure on risk. As cautioned by Cain et al. (1992) for models of the effects of changes in risk factors (such as exposure) on changes in health risks, "The problem occurs when the true value of the risk factor relates to the outcome, and the measured value differs from the true value due to measurement error. *We may find the observed change in the risk factor significantly related to the outcome when there is in fact no relationship between the true change and the outcome*" (emphasis added). We are not aware of any investigation of similar biases specifically for the Attfield-Seixas and Kempel models in Appendices I-K of the QRA. However, the fact that these models do not correct for effects of measurement (or estimation) errors in exposure estimates and covariates raises the possibility that an unknown fraction (up to 100%) of the risk attributed to differences in exposures may in reality be due to unmodeled errors in variables.
- *Ignored interactions among variables.* For lung diseases with multi-factorial etiologies (e.g., having coal dust exposure, quartz dust exposure, cigarette smoking, etc. as contributing causes), it is appropriate to calculate a *partial* attributable risk (Rämsch et al. 2009) that partitions the total difference in risk between two groups among the various differences in the factors that affect risk, taking into account any interactions among factors. Socioeconomic and demographic factors that affect exposure or risk (such as age, seniority, income, etc.), as well as interactions among them, should be included in the model and used in the calculation of partial attributable risks. Otherwise, attributing the difference in risks solely to differences in exposures, as in the QRA, can lead to spuriously high estimates of the risks attributed to

exposure, and these attributable risks can be very different from the risks that are caused by exposure, or that are preventable by reducing exposure.

Example: Ignoring Interactions among Risk Factors Invalidates Attributable Risks

As a simplified example, suppose that the risks for individuals depend on joint exposures to coal dust (exposed *vs.* unexposed) and smoking (non-smoker *vs.* smokers), as shown in the following table.

Hypothetical example showing disease probabilities (and number of workers) for each combination of two factors

	Unexposed	**Exposed**
Non-smoker	0 (100 workers)	0 (0 workers)
Smoker	0.4 (100 workers)	0.8 (200 workers)
Total	40 cases among 200 workers	160 cases among 200 workers

Thus, in this example, an unexposed smoker has a probability of 0.4 of disease, and there are 100 such workers (lower left cell). An exposed smoker has a 0.8 probability of disease (perhaps due to poorer access to medical care at locations for exposed workers compared to unexposed workers). Non-smokers have zero risk. The QRA's definition of attributable risk would attribute the difference between risks of exposed and unexposed workers to exposure, thereby attributing (160 – 40)/(160 + 40), or 60%, of total cases to exposure. This ignores the fact that exposure without smoking creates zero excess risk. The same attribution procedure would also attribute 100% of all cases to smoking. Thus, it allocates 160% of the total number of cases to the two risk factors, thereby inflating the total number of cases being attributed by 160%. Partial attributable risk calculations are designed to prevent such anomalies and exaggerations in risk, but have not been used in MSHA's QRA.

The models relied on by the QRA do not model interactions among factors. (For example, the logistic regression model in Table 54 (p. 131) of the QRA does not even include terms for interactions among smoking, age, and cumulative exposure.) *The QRA does not calculate the partial attributable risk (if any) that is specifically due to exposure,* but instead mistakenly attributes *all* of the difference in projected risks calculated with and without exposures to the effects of exposure, ignoring all other variables and interactions. Yet, as noted on page 132, such interactions are present: "[T]he portion of emphysema risk attributed to dust exposure is

greater for smokers than for non-smokers, by an amount that increases with the intensity and duration of smoking." We suggest that the QRA's modeling inappropriately attributes to RCMD some of the share in risk that should be attributed instead to cigarette smoking. Partial attributable risk calculations are needed to avoid over-attribution of total risks to exposure.

More generally, there is no necessary relation between the *attributable risks* calculated for exposures, and the *risks that would be prevented* by removing exposure. The QRA mistakenly conflates these two distinct concepts. It does not calculate what it claims to: the change in risk that would occur (i.e., be caused) if the proposed rule were implemented. Rather, it attributes to RCMD risks that are caused in part by other things (such as smoking, or unmodeled measurement errors), and then mistakenly presents these as risks as if they were solely caused by, and preventable by reducing, RCMD exposures. The QRA should be revised to withdraw its claim that "the difference in risk calculated with and without exposure" is "that portion of the risk attributable to the accumulated exposure," and should discuss partial attributable risks. It should also clarify that the attributable risks it has calculated have no known or demonstrated relation to risks *caused* by RCMD exposure.

Comment: The QRA does not provide statistical confidence intervals for its attributable risk. Thus, it leaves unanswered the key question of whether its estimated attributable risks are significantly different in different exposure groups. The QRA should be revised to show confidence intervals specifically for the risks caused by RCMD exposure.

p. 31: *"As suggested by the difference in vertical scales for Figures 10 and 11, Attfield and Seixas found the effect of exposure to be significantly elevated at high rank coal mines."*

Comment: It is misleading to state that Attfield and Seixas "found the effect of exposure to be significantly elevated," since they did not identify any causal effects of exposure. As just explained, the Attfield-Seixas model does not actually quantify a (causal) effect for exposure. It is a regression model, not a causal model. It uses a specific epidemiological model to *attribute* shares in hypothetical risk projections to various sources, including high-rank coal, but this is more an accounting exercise (deciding how much blame to allocate to RCMD, based on implicit

modeling and policy judgments) than an empirical finding demonstrating an actual causal effect of exposure on risk. The risk attribution exercise can be done in different ways, producing different quantitative results, as noted in our discussion of partial attributable risks. Thus, the QRA should not refer to such attributed risks as if they were findings of actual effects of exposure, when they are only results of more or less *ad hoc* attribution procedures.

p. 32 *Estimated relationships between average RCMD concentration experienced over a 45-year working lifetime and excess risk of CWP and PMF.*

Comment: The QRA should clearly state that the estimated relationships presented in Figures 10 and 11 are *attributed* to RCMD exposure via a regression model, but not necessarily *caused* by RCMD exposure. These regression relationships have not been validated as predictive or causal models, and it is incorrect to interpret them as showing how changes in average RCMD concentrations would change risks (causal interpretation). The extent to which these curves reflect the effects of unmodeled confounders, measurement errors, interactions among factors, model specification errors, use of obsolete smoking assumptions, etc. has not been quantified or corrected for, so they should not be interpreted as if they represented true causal relationships between average RCMD concentrations and lung diseases. That some of the models attribute a large relative risk (4.4) to coal even when exposure is zero (p. 135) further suggests that causal interpretation of model results is not warranted. The QRA should be revised to clarify that the regression-based risk estimates it has calculated have no known or demonstrated relation to risks *caused* by RCMD exposure.

Additional Brief Comments

The following comments apply the major points already discussed to other parts of the QRA text that we believe may be misleading to many readers. They do not re-state our reasoning in detail, but briefly comment on recommended revisions.

pp 34-35: *"These relative risks represent an expected 20 to 70-percent increase in pneumoconiosis mortality, attributable to the exposure."*

Comment: For reasons discussed in earlier comments, this claim should be revised to clarify that "attributable to" does not mean "caused by," and that the increases in mortality said to be "attributable to the exposure" are based on unvalidated models and calculations that can attribute substantial risk even to zero exposure. It is important that the text not mislead readers into believing that it has shown that exposure *causes* a 20- to 70-percent increase in pneumoconiosis mortality.

p. 37 Figure 13: *"Estimated relationship between average coal mine dust concentration exposure... and predicted average reduction in FEV_1..."*

Comment: This figure has no clear meaning. It lacks confidence bands (or data points) to show whether its suggested straight-line relation is empirically supported, and it does not say what kind of "relationship" it is estimating. Simply plotting one quantity against another does not establish a causal or statistical "relationship" between them (e.g., showing that the one on the x axis is a useful predictor of the one on the y axis). It seems that even a beneficial effect of exposure (such as reducing early-onset diseases and prolonging life, so that age-dependent loss of FEV_1 has a greater opportunity to operate) might lead to such a diagram. Figure 13 should be either much better explained, or else removed.

p. 38: *"According to this model, the effects of dust exposure increase with age. After a 45-year exposure averaging 1.0 mg/m^3, the excess risk of severe emphysema is predicted to be 50 excess cases per thousand (ECPT) for 65-year old miners."*

Recommendation: Revise to remove the phrase "the effects of dust exposure" (which are not known and have not been quantified), and to explain that the 50 excess cases per thousand are not actually caused solely by dust exposure, but include cases that would not occur in the absence of cigarette smoking.

p. 39 "Both studies report a statistically significant exposure-response relationship..."

Recommendation and Comment: Replace "exposure-response relationship" with "association, under the simplifying assumptions of the models used". The word "relationship" is highly ambiguous, and could be interpreted mistakenly to mean "predictive relationship" or "causal relationship." The text should clarify that the specific "relationship" found is one of *statistical association*, based on a regression coefficient in an unvalidated statistical model that did not correct for measurement errors, interaction terms, effects of omitted confounders and covariates, or model specification errors and uncertainties. The claim of "statistical significance" is supportable only if it is *assumed* that the models used are correct (since model uncertainty has not been quantified). It is unlikely that the models are correct as specified, in light of anomalies such as their conflicts with historical data, with each other, and with common-sense constraints (such as that zero risk should be attributed to zero exposures). Thus, their claims of statistical significance, which are contingent on model validity, are not trustworthy. Even if a statistically significant association between exposure and response exists, however, it has not been shown to be a genuine (causal) exposure-response relationship, but is simply an association between two variables, possibly explained by confounding, biases, model specification errors, or other non-causal sources of association. Calling the association an "exposure-response relationship" is misleading, as no genuine exposure-response relation has been shown to exist.

p. 41 ""Indeed, when exposure-response relationships are curved upwards as those shown above, *evaluating risk at the average exposure level will always underestimate average risk*".

Comment: As previously discussed, *estimating* risk from data on numbers of cases at each estimated average exposure level actually *over-estimates* the true risk at each exposure level.

The QRA's reversed conclusion would hold only if the exposure-response relation were already known with certainty, rather than being estimated from data with significant variance around each estimated average exposure level.

p. 54, Table 16: *"Estimated excess risk of developing severe emphysema by age 73... (cases attributable to occupational exposure per 1000 exposed workers).*

Recommendation and Comment: Replace "attributable to occupational exposure" with "attributed to (but not necessarily caused by) occupational exposure". As previously discussed, these "attributable" risk numbers have no necessary relation to causal effects of exposure, and would not necessarily be reduced if exposure were reduced.

p. 55 "The increased effect of exposure to high rank coal, and especially anthracite, on excess NMRD mortality risk is readily apparent by age 73 (Table 17)..."

Recommendation and Comment: Replace "effect of exposure" with "attribution of risk to exposure." As previously discussed, the regression models used to attribute risks to coal are not causal models. It is misleading to describe the excess case numbers that they attribute to coal as "effects" of exposure to coal, since it has not been shown that exposure to coal causes or affects these cases.

p. 58 "*...[C]urrent exposure conditions place miners at a significant risk of incurring each of the material impairments considered.*"

Recommendation and Comment: Replace "Current exposure conditions place" with "Current exposure conditions are assumed in our risk attributions to place," or else withdraw this claim, which is not supportable based on the data and analyses presented. This statement makes an explicit causal claim (that current exposure conditions "place miners at a significant risk," which has not been demonstrated. The conclusion does not follow from the mix of attributable risk calculations, unvalidated regression models, selections of data, one-sided "adjustments" of

exposure estimates, and assumptions used in the QRA. It is a misinterpretation of what has been shown.

p. 58: "Sufficient homogeneity within WL clusters... [V]ariability of exposure levels at different WLs within each cluster is [assumed to be] small enough that the exposure-response models are approximately linear within the range of exposures represented by the cluster."

This assumption is implausible, based on the high variability observed in measurements of exposures, even for the same occupations (see Table 11), and on the presence of important unmodeled sources of variability across work locations (e.g., in RCMD particle size distributions and coal ranks, as acknowledged on p. 59 of the QRA). It is also certainly incorrect for threshold (or approximately threshold-like) exposure-response relations, such as those for respiratory diseases (see point 5 above). Moreover, it can be shown that the effect of failing to model exposure uncertainties (and variability within exposure classes) when the true exposure-response relation is threshold-like is to *over-estimate risks at low exposures* (and to under-estimate them at high exposures). This invalidates the QRAs projection of health benefits (Section 3, starting on p. 59) from further reductions in exposure standards.

Comment: Heterogeneity in exposure-response relations has not been modeled, yet is well documented (www.cdc.gov/niosh/docket/review/docket210/pdfs/0005-Coal-CIB-08-09-2010-Public-Comment%20mda.pdf, Section 2.2, p. 9). This further undermines the key assumption of "sufficient homogeneity within WL clusters."

p. 60 *"There are two parts to the simulation procedure, corresponding to the proposed reduction in exposure limit and the proposed change in enforcement policy."*

Comment: The simulation procedure mistakenly treats ANCOVA "effects" and regression model and risk attribution "effects" as if they were *causal* effects in valid causal models. It therefore has no demonstrated or likely predictive validity.

Table 1. Summary of Some Important Limitations of MSHA's QRA

Topic	Main Comments	More Detailed Comments
Hazard identification	No hazard identification has been presented	Mode of action, biological evidence, and thresholds are never discussed.
	No mode of action (MOA) information has been presented	MOA information suggests an exposure-response threshold
Exposure assessment	MSHA's use of cumulative exposures is unjustified and exaggerates risks	Cumulative exposure is inconsistent with MOA information
		Cumulative exposure is also inconsistent with human and animal data for dusts and lung disease
	Mean cumulative exposures do not provide a sound basis for predicting risks.	Exceptionally high exposures, not average (mean) exposures, drive risk
	MSHA's exposure estimates are biased upward by making "adjustments" in exposures without making needed corresponding adjustments in estimated exposure-response relationships.	MSHA deliberately omits post-abatement measurements and adjusts exposure estimates upward (but not downward) when sources disagree. This biases exposure estimates.
	MSHA's QRA does not model errors in exposure and other variables. This crucial omission creates biases in results.	Unmodeled errors in exposure estimates can create spurious positive associations between estimated exposure and disease.
Exposure-Response model	No true (causal) exposure-response model is presented	
	The statistical regression relations used are not valid exposure-response models	The statistical regression models are biased by unmodeled effects of variance in exposure durations, estimation errors, omitted confounders, model uncertainty, and model specification errors (which overstate risks)
Risk Characterization	The QRA does not *characterize* risks, but uses incorrect formulas and unvalidated regression models to *attribute* risks to RCMD.	The regression models used by MSHA to attribute risks to RCMD do not adjust for declining smoking, and hence over-estimate risks from RCMD.
		The risk attribution formulas used are incorrect, as they do not account for contributions from multiple factors.
Uncertainty Characterization	The QRA does not characterize uncertainties. This makes its results inappropriate for use in regulatory risk management decision-making.	The QRA results are stated without confidence intervals, sensitivity analyses, or validation tests (comparing model predictions to reality).

REFERENCES

Azad N, Rojanasakul Y, Vallyathan V. Inflammation and lung cancer: roles of reactive oxygen/nitrogen species. J Toxicol Environ Health B Crit Rev. 2008 Jan;11(1):1-15. http://pdfserve.informaworld.com/643704__789269849.pdf

Bang KM, Althouse RB, Kim JH, Game SR. Recent trends of age-specific pneumoconiosis mortality rates in the United States, 1985-1996: coal workers' pneumoconiosis, asbestosis, and silicosis. Int J Occup Environ Health. 1999 Oct-Dec;5(4):251-5.

Cain KC, Kronmal RA, Kosinski AS. Analysing the relationship between change in a risk factor and risk of disease. Stat Med. 1992 Apr;11(6):783-97.

Carroll RJ. Covariance analysis in generalized linear measurement error models. Stat Med. 1989 Sep;8(9):1075-93

Cheng D, Branscum AJ, Stamey JD. Accounting for response misclassification and covariate measurement error improves power and reduces bias in epidemiologic studies. Ann Epidemiol. 2010 Jul;20(7):562-7.

Comhair SA, Erzurum SC. Antioxidant responses to oxidant-mediated lung diseases. Am J Physiol Lung Cell Mol Physiol. 2002 Aug;283(2):L246-55.

Cox LA Jr. A mathematical model of protease-antiprotease homeostasis failure in chronic obstructive pulmonary disease (COPD). *Risk Analysis* 2009 Apr;29(4):576-86.

Goldstein BD. Advances in risk assessment and communication. Annu Rev Public Health. 2005;26:141-63.

Hossain S, Gustafson P. Bayesian adjustment for covariate measurement errors: a flexible parametric approach. Stat Med. 2009 May 15;28(11):1580-600.

Jacobson, M. Respirable dust in bituminous coal mines in the U.S. Inhaled Part. 1970;2:745-56.

Jardine C, Hrudey S, Shortreed J, Craig L, Krewski D, Furgal C, McColl S. Risk management frameworks for human health and environmental risks. J Toxicol Environ Health B Crit Rev. 2003 Nov-Dec;6(6):569-720.

Kim HG, Richardson D, Loomis D, Van Tongeren M, Burstyn I. Bias in the estimation of exposure effects with individual- or group-based exposure assessment. J Expo Sci Environ Epidemiol. 2010 Feb 24.

Leffondre K, Wynant W, Cao Z, Abrahamowicz M, Heinze G, Siemiatycki J. A weighted Cox model for modelling time-dependent exposures in the analysis of case-control studies. Stat Med. 2010 Mar 30;29(7-8):839-50.

Lu C, Lyles RH. Misclassification adjustment in threshold models for the effects of subject-specific exposure means and variances. Proceedings of the Joint Statistical Meeting, 2008. http://www.amstat.org/sections/srms/proceedings/y2008/Files/302008.pdf

MacNee W. Pulmonary and systemic oxidant/antioxidant imbalance in chronic obstructive pulmonary disease. Proc Am Thorac Soc. 2005;2(1):50-60.

Miller BG, MacCalman L. Cause-specific mortality in British coal workers and exposure to respirable dust and quartz. Occup Environ Med. 2010 Apr;67(4):27

MMWR Morb Mortal Wkly Rep. 2009 Dec 25;58(50):1412-6. Coal workers' pneumoconiosis-related years of potential life lost before age 65 years - United States, 1968-2006. Centers for Disease Control and Prevention (CDC).

Mossman BT. Mechanisms of action of poorly soluble particulates in overload-related lung pathology. Inhal Toxicol. 2000 Jan-Feb;12(1-2):141-8.

Murad H, Freedman LS. Estimating and testing interactions in linear regression models when explanatory variables are subject to classical measurement error.Stat Med. 2007 Oct 15;26(23):4293-310.

Oberdörster G. Toxicokinetics and effects of fibrous and nonfibrous particles. Inhal Toxicol. 2002 Jan;14(1):29-56.

Pease JE, Sabroe I. The role of interleukin-8 and its receptors in inflammatory lung disease: implications for therapy. Am J Respir Med. 2002;1(1):19-25.

Porter DW, Hubbs AF, Mercer R, Robinson VA, Ramsey D, McLaurin J, Khan A, Battelli L, Brumbaugh K, Teass A, Castranova V. Progression of lung inflammation and damage in rats after cessation of silica inhalation. Toxicol Sci. 2004 Jun;79(2):370-80.

Porter DW, Millecchia LL, Willard P, Robinson VA, Ramsey D, McLaurin J, Khan A, Brumbaugh K, Beighley CM, Teass A, Castranova V. Nitric oxide and reactive oxygen species production causes progressive damage in rats after cessation of silica inhalation. Toxicol Sci. 2006 Mar;90(1):188-97.

Raaijmakers and Pieters 1987. Measurement error and ANCOVA: Functional and structural relationship approaches. Psychometrika 2007 Dec;52(4):521-38.
http://raaijmakers.edu.fmg.uva.nl/PDFs/Raaijmakers%20and%20Pieters%20Psychometrika.pdf

Rämsch C, Pfahlberg AB, Gefeller O.Point and interval estimation of partial attributable risks from case-control data using the R-package 'pARccs'. Comput Methods Programs Biomed. 2009

Richardson DB. Occupational exposures and lung cancer: adjustment for unmeasured confounding by smoking. Epidemiology. 2010 Mar;21(2):181-6.

Schins RP, Borm PJ. Mechanisms and mediators in coal dust induced toxicity: a review. Ann Occup Hyg. 1999 Jan;43(1):7-33.

Steenland K, Mannetjé A, Boffetta P, Stayner L, Attfield M, Chen J, Dosemeci M, DeKlerk N, Hnizdo E, Koskela R, Checkoway H; International Agency for Research on Cancer. Pooled exposure-response analyses and risk assessment for lung cancer in 10 cohorts of silica-exposed workers: an IARC multicentre study. Cancer Causes Control. 2001 Nov;12(9):773-84. Review. Erratum in: Cancer Causes Control 2002 Oct;13(8):777.

Steenland K, Deddens JA, Zhao S. Biases in estimating the effect of cumulative exposure in log-linear models when estimated exposure levels are assigned. and J Work Environ Health. 2000 Feb;26(1):37-43.

Tuluce Y, Ozkol H, Koyuncu I, Ine H. Increased occupational coal dust toxicity in blood of central heating system workers. Toxicol Ind Health. 2010 Aug 19. [Epub ahead of print]

Veierød MB, Laake P. Exposure misclassification: bias in category specific Poisson regression coefficients. Stat Med. 2001 Mar 15;20(5):771-84.