Texas Commission on Environmental Quality
Responses to Comments Received on the
Arsenic Development Support Document

Section 3 Acute Evaluation, Section 4.1 Chronic Noncarcinogenic Potential of the Development Support Document for Arsenic and Inorganic Arsenic Compounds, Draft, April 2009

and

Section 4.2, Carcinogenic Potential of the Development Support Document for Arsenic and Inorganic Arsenic Compounds, Draft, May 2010.

and
Public Comments on Arsenic Development Support Document and Summary of changes made to the Section 3.1.6.2

Prepared by:
Neeraja Erraguntla, Ph.D.
Roberta L. Grant, Ph.D.

Toxicology Division
Office of the Executive Director

TABLE OF CONTENTS

Introduction

Toxicology Excellence for Risk Assessment (TERA) supported the Texas Commission on Environmental Quality (TCEQ) by conducting two separate expert external peer reviews on different sections of the Development Support Document for Arsenic and Inorganic Arsenic Compounds.

TERA conducted an expert external peer review of acute and chronic noncarcinogenic sections of the *Development Support Document for Arsenic and Inorganic Arsenic Compounds, Preliminary Draft,* April 2009. The review materials, including draft document, charge to

reviewers, and key references (available at http://www.tera.org/Peer/arsenic/index.html) were distributed to the panel in July 2009. Reviewers submitted written comments that addressed the charge questions in August 2009. On August 25, 2009, *TERA* facilitated a follow-up conference call between the panel and TCEQ. The final peer review report of the follow-up conference call, along with the written comments submitted by the panel, comprise the complete peer review of the arsenic DSD (October 14, 2009 *Final Report for the Peer Review of the Acute and Chronic Noncarcinogenic Sections of the Arsenic Development Support Document and Follow up Conference Call*). These written comments represent the panel's review of the arsenic acute and chronic noncarcinogenic sections of the DSD and are available at http://www.tera.org/Peer/arsenic/index.html.

TERA conducted an expert external peer review of Section 4.2, *Carcinogenic Potential, of the Development Support Document for Arsenic and Inorganic Arsenic Compounds*. May 2010. The review materials, including draft document, charge to reviewers, and key references (available at http://www.tera.org/Peer/arseniccarc/index.html) were distributed to the panel in May 2010. Panel members reviewed Section 4.2 of the DSD and submitted written comments that addressed the charge questions in July 2010. On August 26, 2010, TERA facilitated a follow-up conference call between the panel and TCEQ. The final peer review report of the follow-up conference call, along with the written comments submitted by the panel, comprise the complete peer review of the arsenic DSD (September 24, 2010 *Expert External Peer Review of Section 4.2, Carcinogenic Potential, of the Development Support Document for Arsenic and Inorganic Arsenic* Compounds). These written comments represent the panel's review of the arsenic carcinogenic section of the DSD and are available at http://www.tera.org/Peer/arsenic/index.html.

The April 2009 and May 2010 versions of the DSD were revised based on panel member comments. TERA did not receive any public comments on the different sections of the arsenic DSD. The Toxicology Division (TD) of the TCEQ appreciates the effort put forth by the panel members to provide technical comments on the draft DSD for arsenic. The goal of the TCEQ is to protect human health and welfare based on the most scientifically-defensible approaches possible (as documented in the DSD), and evaluation of these comments furthered that goal. The panel members' comments on the acute and chronic noncarcinogenic sections are addressed first, followed by panel members' comments on the carcinogenic section. Comments on issues that suggest a change in the DSD are addressed whereas comments agreeing with TCEQ's approach are not. TCEQ responses indicate what changes, if any, were made to the DSD in response to the comment.

Panel Comments on Acute and Chronic Noncarcinogenic Assessment

Issue 1:

Are the peer reviewers aware of any information or references that would support (or refute) the position that data for arsenic trioxide should be considered relevant for arsenic pentoxide?

TCEQ Response: The panel was not aware of specific data on the comparative toxicity of arsenic trioxide and arsenic pentoxide. The panel agreed that arsenic pentoxide is converted to arsenic trioxide, and the use of arsenic trioxide data would be a health protective approach. One reviewer stated that arsenic pentoxide also has a separate mode of action, but would likely cause effects at higher doses than arsenic trioxide. So using arsenic trioxide as a surrogate may be conservative; it would not be predictive but it would be protective. This reviewer suggested that TCEQ add a discussion of this issue to the DSD.

Issue 2:

Does the panel have additional recommendations related to the rationale for using 10 μM as a size-selective cut point for the ESL?

TCEQ Response: One panel member asked TCEQ to clarify what they meant by "size-selective cut point." TCEQ clarified that the monitoring data was based on particles of 10 um or smaller because particles of that size are respirable. Therefore, TCEQ had originally intended to make the ESL applicable to certain size particles (those that are respirable). But based on the panel's initial written comments, TCEQ decided to use a total particulate matter approach rather than using the cut-off size of 10 μM. Another panel member suggested that a short discussion addressing the possibility of systemic effects from inadvertent exposure to larger particle sizes due to mucociliary clearance from inhalation exposure be added to the report as further justification of using total particulate for the assessment.

Issue 3:

The panel felt that the Nagymajtenyi et al. (1985) study is not an adequate study to develop an acute ReV and should be replaced by the Holson et al. (1999) study as the key study. One reviewer noted an alternative explanation (Reviewer #2) of the Nagymajtenyi et al. (1985) findings and indicated that Nagymajtenyi et al. (1985) study might be used to develop a supporting value. Further, the reviewer wanted to clarify whether such an approach for the derivation is reasonable. For the Nagymajtenyi et al. (1985) study, the critical effect is maternal toxicity, with the highest dose a LOAEL and the mid-dose of 2.9 mg/m^3 a minimal LOAEL or a NOAEL. The TCEQ agreed with the peer-reviewer's comments and did not include o BMCL modeling or RDDR calculation would be included due to weaknesses in the study. Is the use of the Nagymajtenyi et al. (1985) to derive only a supporting value to supplement Holson et al. (1999) scientifically defensible? The panel discussed the unreliability of the study design and the lack of adequate documentation for the Nagymajtenyi et al. (1985) study. One panel member stated that he did not have confidence in this study because the exposure could have been by the oral pathway in addition to inhalation. The panel member noted that animals in Nagymajtenyi et al (1985) were exposed to a "fog" that would have deposited chemical on their fur. Therefore, the animals could have had oral exposure in addition to inhalation exposure when they groomed.

In contrast, the animals in the Holson study were washed after exposure, to ensure exposure was limited to inhalation.

Thus, the Holson et al. (1999), study would be more reliable for developing an inhalation assessment. TCEQ stated they felt the Nagymajtenyi et al. (1985) study was more conservative, and was more of an acute duration study. Acute studies are preferred to drive 1-hr ESLs. However, repeat-dose, multi-day studies can be used to develop short-term ESLs if an acute study is not available, according to the TCEQ guidelines (TCEQ 2005). TCEQ asked the panel if it would be useful to still derive a value based on the Nagymajtenyi et al. (1985) study as a co-critical or supporting study if the Holson et al. (1999) study were selected as the key study for the revised assessment. One panel member responded that if the study is not credible enough to use as the critical study, then no value should be derived from it. The panel agreed that even though the Nagymajtenyi et al. (1985) study was qualitatively similar to the Holson et al (1999) study, the Nagymajtenyi et al. (1985) study is unreliable and should not be used for any quantitative estimates in support of the Holson et al (1999) study. Rather, the panel recommended that a paragraph be added that discusses the limitations of the study as well as the interpretations of the data and its qualitative consistency with the Holson et al. (1999) study.

One panel member discussed the issue of pretreatment of the rats in the Holson et al. (1999) study. Sprague-Dawley rats are resistant to arsenic due to the arsenic binding with the rat's hemoglobin. The rats were pretreated to ensure that all sites on the hemoglobin were occupied before assessing reproductive toxicity. The panel suggested that TCEQ should include a discussion on the pretreatment protocol to provide more confidence in the study. Additional information on the pretreatment of the rats in the Holson et al. (1999) study should also be included to provide more confidence in the study.

TCEQ Response: TCEQ has incorporated the peer reviewer's comments and did not select the Nagymajtenyi et al. (1985) study as the key study. TCEQ chose the Holson et al. (1999) study as the key study and included the respiratory effects and decrease in body weight gain as the critical effects (i.e., maternal effects).

Issue 4:

The critical effects observed in the Holson et al. study (1999) were respiratory effects, (i.e., point-of-entry effect) and decrease in body weight. If the Holson et al. (1999) study was used as the basis for the Acute ESL, would route-to-route extrapolation to derive values based on other critical endpoints observed in acute oral studies add significantly to confidence in the ESL derived from this study?

One panel member stated that in the range finding studies, many animals died early due to suffocation that resulted from the accumulation of dust in the lungs. The panel member added that the amount of arsenic in the lungs would probably not be enough to be designated a critical effect, but it should be noted that pulmonary irritation was seen in the animals and could be noted as a portal-of-entry effect. Another panel member agreed and added that route-to-route extrapolation would only be useful as support for a weak study; however the Holson et al. (1999) study is robust, so route-to-route would not be needed. TCEQ added that all the females died at the top three exposure concentrations due to suffocation and deaths were also seen in repeated-doses at the 25 mg/m3 exposure. The study authors did not do any pathology and it was reported

that the lungs did not show any pulmonary irritation. There were also signs of gastrointestinal tract lesions, which are consistent with arsenic exposure, and that decreased body weight was listed as a critical effect. One panel member replied that the study reported rales in the lungs which are a sign of lung inflammation. The panel member suggested that the assessment more clearly describe the observed effects and characterize those which are likely respiratory effects and those which may indicate systemic effects. The panel agreed that a quantitative route-to-route extrapolation approach was not necessary since a reliable inhalation study was available and would not add to the quantitative assessment. One panel member noted that comparisons to the oral database would help alleviate any concerns about whether the Holson et al. (1999) identified the likely critical effects.

TCEQ Response: TCEQ did not conduct a route-route extrapolation from oral ingestion data to derive inhalation values because the Holson et al. (1999) study was a well designed inhalation study.

Issue 5:

Do the data support that concentration and duration both play a role in the toxicity of arsenic? If so, for the Holson et al (1999) study, rats were exposed to arsenic as a solid particulate. Should Haber's Rule as modified by ten Berge, with n = 3, be used to extrapolate the 6-hr per day exposure duration to a 1-hr exposure duration? One panel member noted that it is important to determine which dose-metric is relevant to toxicity in order to determine if a duration adjustment is appropriate. This panel member indicated that since toxicity of arsenic is related to area under the curve, then cumulative exposure is the important dose metric. Further, since cumulative exposure is the important dose metric, then it is appropriate to use Haber's Law. When Haber's law is appropriate, using the ten Berge modification is more health protective. Another panel member noted that the ten Berge modification would not be as important for exposure to particulates versus exposure to a vapor or gas. The panel agreed that the 6-hour concentration should not be used directly to represent the effects for a 1-hour exposure; rather, the 6-hr exposure from Holson et al. (1999) should be adjusted to a 1-hr equivalent using Haber's Law with the ten Berge modification since the critical effects may reflect the effects of cumulative dose and systemic effects are possible. TCEQ noted that their guidelines require the use of an "n" equal to a value of 3 as a default for the ten Berge modification unless other data suggest a better value. The panel agreed that there are no chemical-specific data to support a different choice of "n", so a default of 3 is appropriate.

TCEQ Response: For the Holson et al. (1999) study, the TCEQ concurred with the peer reviewers and conducted a duration adjustment from 6h to 1h based on the Haber's Law as modified by ten Berge with n = 3.

Issue 6:

If used as a supporting study, the Nagymajtenyi et al. (1985) study exposed mice to arsenic as a liquid aerosol. Should the Haber's Rule as modified by ten Berge, with n = 3, be used to extrapolate the 4-hr per day exposure duration to a 1-hr exposure duration? The panel agreed, based on previous discussions, that the Nagymajtenyi et al. (1985) study should not be used in any quantitative analysis, so no answer to this charge question was needed.

TCEQ Response: TCEQ agrees with the peer reviewers and did not derive acute inhalation toxicity factors based on the Nagymajtenyi et al. (1985) study as stated above in Issue 3. Therefore, this comment is not applicable.

Issue 7:

Are the oral studies conducted by Stump et al. (1999) and Baxley et al. (1981) (Reviewer #2) adequate justification for assuming that mice are not more sensitive than rats to developmental toxicity? Holson et al. (1999) states rats are more sensitive than mice based on Lewis and Sweet (1985). Would inclusion of a discussion of these studies enhance confidence in the selection of Holson et al. (1999) as the critical study?

One panel member stated that, based on the literature for arsenic, mice and rats appear to have similar sensitivity. The panel member noted that the overall database suggests that high exposure concentrations or doses could result in reduced birth weights; the other observed effects are consistent with maternal toxicity. This reviewer also stated that the available studies support the conclusion that arsenic does not cause malformations at environmentally relevant concentrations. The Lewis and Sweet (1985) reference is no longer in print and could not be reviewed. The panel concluded that it would benefit the assessment to add a discussion of the oral studies (Stump et al. 1999 and Baxley et al. 1981) which support the overall database, including the Holson et al. (1999) study.

A discussion should also be added to state that the general literature did not suggest sensitivity differences between the rat and mouse (which further supports the selection of the Holson et al. (1999) study over the poorly documented study by Nagymajtenyi et al. (1985)). TCEQ asked if the panel would comment on whether arsenic is a developmental toxicant. The answer to this question is not simple and requires some explanation. “Developmental toxicant” is a broad term that replaced “teratogen,” which refers to a substance that causes fetal malformations. Developmental toxicity encompasses multiple fetal effects that include alterations in fetal weight, death of offspring, and functional deficits in addition to malformations. The presence of any of these findings is sufficient to signal developmental toxicity. One panel member stated that arsenic does not cause malformations via typical routes of exposure because the dam cannot remain sufficiently healthy and take in enough arsenic to cause fetal malformations. However, the issue of decreased fetal weight from exposure to arsenic is more complicated. If dams receive a high enough concentration of arsenic to cause toxicity but not death, then reduced fetal weight could occur. This response could also possibly occur in humans. Consequently, based on the possibility that sufficiently high maternal arsenic exposure could result in fetal weight deficits, arsenic could be considered developmentally toxic at very high exposure concentrations. There is possible confounding in that the intoxicated pregnant animals will likely stop eating and reduced fetal weights could be an indirect effect. The panel member added that environmentally relevant exposures (low concentrations) would not cause any fetal effects, but concentrations high enough to cause severe toxicity in the pregnant woman could lead to fetal effects (likely reduced fetal weight). Importantly, such effects would occur in association with serious maternal toxicity. Because of the presence of concomitant maternal toxicity, arsenic should not be considered a selective developmental toxicant. (A selective developmental toxicant is one that causes effects in the offspring at exposures that are not harmful to the pregnant female and is very serious). The

panel did not have a discussion on uncertainty factors during the conference call, but did include discussion of uncertainty factors in their individual written comments (see Appendix B).

TCEQ Response: TCEQ reviewed relevant drinking water studies and agrees with the peer-reviewers that the Holson et al. (1999) study to be a better selection for key study.

Issue 8:

Based on the written comments, panel members had varying opinions as to whether the limitations in the Lagerkvist and Zeturland (1994) and Lagerkvist et al. (1986) studies precluded their use to develop a chronicReV. Nevertheless, all the panel members noted significant concerns with these studies. As an alternative approach, would a unit risk factor (URF) and chronicESL based on excess lung cancer mortality in four different cohorts of workers exposed to arsenic be a more robust basis for the chronicESL and would an ESL derived on this basis protect against sensitive noncancer effects?

TCEQ Response: To provide background information, TCEQ indicated that their guidelines do not necessarily require separate noncancer and cancer values. Rather, effects from long-term exposure should be evaluated. This is the first time moving away from deriving a separate chronic noncancer value for a known carcinogen has been considered. This approach is health protective because a quantitative assessment based on the carcinogenicity usually provides a lower toxicity value than one developed based on noncarcinogenic effects. A panel member asked if TCEQ would be using a linear extrapolation or a non-linear assessment approach (point of departure based on a cancer endpoint and application of uncertainty factors) for the carcinogenic assessment. The typical cancer assessment assuming a linear dose-response would be assumed by TCEQ because the arsenic mode of action for lung tumors is not well defined. Overall, the panel agreed that deriving only a cancer assessment would adequately characterize the long-term effects of arsenic and would likely be protective of potential noncancer effects of arsenic. However, in order to address this uncertainty, the panel suggested that TCEQ include a comparative analysis of the range of values that could be derived from the different endpoints and qualitatively discuss the results to support the selected value.

The panel discussed the most appropriate data set for developing a chronic inhalation assessment. One panel member stated that the most appropriate data to use for the chronic inhalation assessment is from the occupational epidemiology related to lung cancer. If chronic noncancer effects data were to be used, then the epidemiology literature related to cardiovascular or gastrointestinal effects should be considered as the basis for a chronic assessment. TCEQ should use the drinking water epidemiology studies to examine these endpoints. However, such an approach was not preferred because it would incorporate route-to-route extrapolation considerations and lower the confidence in the quantitative assessment. Another panel member asked if the use of the lung cancer studies would be protective of sensitive noncancer effects. Another panel member replied that most studies found gastrointestinal effects at high levels of exposure, which were higher than what you would get from a workplace inhalation exposure. However, the panel members agreed that a comparative analysis should be done to be to document this consideration.

TCEQ asked if it would be helpful to develop a reference value based on the Lagerkvist and Zetterland (1994) study and present it in the DSD to give a bound or range of potential values. A

panel member replied that presenting a range is a reasonable approach that has been used by other agencies, such as the USEPA. TCEQ clarified that they could develop a value for the chronic noncancer for the purpose of showing that the cancer value is protective. A panel member replied that the problem with that approach is it gives credibility to a value that is based on a single study that is not completely credible, whereas the value based on the cancer endpoint is based on an entire database and not a single study. The panel member discouraged this approach because a value derived from the Lagerkvist and Zeturland (1994) would give undue credibility to the Lagerkvist and Zeturland (1994). The panel concluded that the chronic value should be based on the cancer inhalation database. The noncancer studies are weak and do not provide adequate data for derivation on a chronic ESL with confidence. The use of the most appropriate noncancer effects data would be based on cardiovascular and gastrointestinal effects seen in drinking water studies, although this approach also was not preferred because the route-to-route extrapolation from oral to inhalation would introduce significant uncertainty. It was suggested that a qualitative discussion of the relative sensitivity of the cancer and noncancer endpoints would provide more support for the use of the cancer inhalation data as the final basis for the chronic assessment.

TCEQ Response: The TCEQ concurs with the peer-reviewers and has not derived chronic ReV for non-carcinogenic effects. The DSD includes a brief summary of the estimated LOAEL as described in the ATSDR (2007) toxicological profile. However, the TCEQ is of opinion that the estimated LOAELs include a large number of uncertainties because of inadequate exposure information. The TCEQ will not therefore not derive a chronic ReV and chronic ESL for non-carcinogenic effects ($^{chronic}ESLnonlinear_{(nc)}$. The TCEQ will however, derive quantitative estimates for carcinogenic effects ($^{chronic}ESL_{linear(c)}$). Based on the weight of evidence the TCEQ is of the opinion that the chronicESLlinear(c) will protect against both the carcinogenic and chronic non-carcinogenic effects too.

Issue 9:

Do the overall data support a causal relationship between arsenic exposure and cardiovascular effects? One panel member stated that heart disease was seen in arsenic exposures from drinking water, specifically data from Taiwan. Cardiovascular effects can lead to heart disease, so it should be considered. Another panel member noted that studies showed confounding exposure to lead; it is biologically plausible that the observed effects were related to arsenic exposure. However, it is not possible to rule out the effects due to co-exposure with lead. The panel concluded that based on the drinking water data, cardiovascular effects due to arsenic exposure appear to be biologically plausible, but additional examination of the overall data would be needed to provide a firmer conclusion.

TCEQ Response: TCEQ did not derive a chronic ReV and Chronic ESL for non carcinogenic effects. TCEQ has derived a chronic ESL for carcinogenic effects and based on the WOE is of the opinion that the ESL developed for carcinogenic effects will be protective of non-carcinogenic effects including cardiovascular effects. As part of a sensitivity analysis to better inform the weight of evidence, the TD is considering supplementing the analysis of the chronic ESL based on the epidemiology studies by using experimental studies and toxicity values from oral routes of exposure to calculate an inhalation toxicity value: (1) USEPA's RfD (Tseng et al. 1977), (2) ATSDR's MRL (Tseng et al. 1977). This would involve route-to route extrapolation

or the use of PBPK models to derive an inhalation ReV, for comparison purposes only. These values will be compared to the air concentration corresponding to a 1 in 100,000 excess risk for lung cancer mortality using the URF derived by the TD. Given the uncertainties associated with route-to-route extrapolation (see the ESL methodology document for the current policy of using route-to-route extrapolation), would such an approach be sufficiently robust to inform the selection or evaluation of the chronic ESL? Do any data limitations as highlighted in the ESL methodology document preclude the meaningful use of route-to-route extrapolation for arsenic with the above mentioned oral toxicity studies?

One panel member stated that route-to-route extrapolation is not necessary because of the availability of robust lung cancer mortality data, and it may cause confusion if used. Another panel member added that, generally, route-to-route extrapolation can be a useful approach, but agreed it is not needed for this assessment. The panel agreed the Wasserman et al. (2004) study needs further review and has not yet been supported with studies by other investigators. The panel member stated that TCEQ should only use the verified endpoints, such as cardiovascular and gastrointestinal effects as seen in the drinking water studies. Other potential effects, such as intellectual development, incorporate many additional variables and are difficult to effectively measure. However, the panel did suggest that route-to-route extrapolation would be a useful tool as part of the sensitivity analysis to evaluate the protectiveness of a cancer value for noncancer effects. If done as part of the sensitivity analysis, the extrapolation should use the identified endpoints for cardiovascular or gastrointestinal effects as the primary non-cancer effects of concern.

TCEQ Response: TCEQ reviewed the Tseng et al. (1977) and conducted a sensitivity analysis (Appendix A3) to evaluate deriving an inhalation reference value based on the RfD. The estimated Reference value based on the RfD is greater (that is less conservative and less health protective) than the chronic ESL for carcinogenic effects that the TCEQ has derived. The TCEQ will therefore use the chronic ESL for carcinogenic effects as the WOE indicates that this value will also protect against non-carcinogenic effects.

Issue 10:

If the TD were to develop a route-to-route extrapolation approach to supplement the ESL derived based on the occupational epidemiology study, what studies should be used? Three agencies have used Tseng et al. (1977) and Wasserman et al. (2004) as key studies in the development of chronic values (RfD, MRL, and chronic REL). Are these the appropriate studies to use as the key studies in the development of chronic values based on route-to-route extrapolation? A panel member began the discussion stating that the Wasserman et al. (2004) study would not be appropriate for use due to the many uncertainties. Another panel member agreed and added that the Tseng et al. (1977) study should be used along with oral animal data to determine if the cancer value is health protective.

TCEQ clarified with the panel member that they were recommending the use of the oral animal data. TCEQ then asked if they should use the human oral data which would include less uncertainty that using animal oral data. One panel member replied that the animal data may give support for the human data. Another panel member added that the animal data is not predictive of human response, and they should review epidemiological studies to determine which studies would be the most appropriate to use in the sensitivity analysis. A panel member added that there

are some epidemiological studies from Taiwan and should also review the epidemiological literature for more studies, such as the Tseng et al. (1977) study. They also recommended review articles, such as the ATSDR toxicological profile on chronic health effects of arsenic to find useful human studies. The panel agreed that the Wasserman et al. (2004) study had too many uncertainties and would not be appropriate to use in a quantitative route-to-route extrapolation. If oral studies were to be examined as a supplement to the inhalation studies, the epidemiology studies, rather than toxicology studies should be the focus. Before proceeding along this course the panel felt that the pool of available epidemiological studies should be critically reviewed prior to selecting any studies or endpoints for further quantitative assessment.

TCEQ Response: TCEQ concurs with the peer reviewers that the Wasserman et al. (2004) study has too many uncertainties and would not be appropriate to use in a quantitative route- to-route extrapolation. Therefore, the TCEQ will not use the Wasserman et al. (2004) study to derive the chronic Rev and chronic ESL. The TCEQ conducted a sensitivity analysis using the RfD derived for the Tseng et al (1977) study. Please see Appendix A3 for a detailed explanation. The RfC in this example is the estimated $^{chronic}ESL_{nonlinear(nc)}$. The $^{chronic}ESL_{nonlinear(nc)}$ based on the RfD (1.05 µg/m^3) is greater (less conservative and less health protective) than the TCEQ's chronic ESL for carcinogenic effects $^{chronic}ESL_{linear(c)}$ of 0.067 µg/m^3 that is based on respiratory and lung cancer in occupational workers. The TCEQ's $^{chronic}ESL_{linear(c)}$ for carcinogenic effects will be used as the final ESL because it will also protect against chronic non-carcinogenic effects. Please refer to Section 4.2.5 for additional information on the $^{chronic}ESL_{linear(c)}$.

Issue 11:

The preferred method for route-to-route extrapolation is the use of PBPK modeling, which provides the best estimate of a toxicant's internal and biologically effective dose as a function of exposure. The peer reviewers suggested that the TD determine if available PBPK models (e.g., Mann et al. 1996) are adequate to conduct route-to route extrapolation. Has the Mann et al. (1966) model been accepted by the scientific community as a valid model that could be used for this purpose? No federal or state agency has used this model to develop inhalation toxicity values based on oral studies. In fact, in their June 2008 Draft TSD for Noncancer RELs, Cal EPA states the following "while some PBPK modeling has been applied to inorganic arsenic and its methyl metabolites, the modes of toxic action and relevant internal dosimetry are not sufficiently understood at present to use this modeling directly in REL development." Is the panel aware of other models or other route-to-route extrapolation approaches that should be considered?

The panel discussed the use of PBPK modeling to enhance the route-to-route extrapolation. One reviewer stated that he had a PBPK modeler review the current status of available models for arsenic. A written summary statement was provided for inclusion into this report (see Appendix F). Panel members noted that the use of PBPK models does not involve all the key uncertainties. TCEQ asked if enough is known about the mode of action to pick the appropriate dose metric panel member replied that the metabolites of interest would need to be known. TCEQ added that the USEPA states that the mode of action for arsenic is unknown. Another panel member agreed and recommended precluding it for use quantitatively, but indicated that it might be used qualitatively as part of the overall sensitivity analysis.

TCEQ Response: The TCEQ agrees with the peer-reviewers. While, the TCEQ reviewed the PBPK models from a qualitative arsenic perspective and will not use the PBPK models to

quantitatively estimate the toxicity factors. The MOA for arsenic is largely unknown and therefore the TCEQ will conduct a linear extrapolation which will result in conservative and health protective screening values.

Panel Comments on Section 4.2 Carcinogenic Section

TELECONFERENCE MAIN COMMENTS

Issue 1 (page 33):

One reviewer opened by commenting that the document relies too heavily on the human data when animal data also support the weight of evidence conclusions. The panel reached consensus that they agreed with overall WOE as developed by TCEQ, but recommended that TCEQ add more discussion regarding the support that the animal studies provide to the WOE statement.

TCEQ Response: TCEQ agrees with the peer-reviewers' comments that there is limited information from animal models in regards to lung cancer. The USEPA's IRIS document indicates that there is inadequate information on animal carcinogenicity on exposure to inorganic arsenic. The IRIS document further states that the "meaning of non-positive data for carcinogenicity of inorganic arsenic is uncertain, the mechanism of action in causing human cancer is not known, and rodents may not be a good model for evaluating lung tumor formation in humans. However, there is some data to indicate that arsenic may produce lung tumors in hamsters if retention time in the lung can be increased (Pershagen et al., 1982, 1984). The following sentence was added to the DSD:

The TCEQ is of the opinion that animal models are significantly limited in their usefulness in regards to arsenic inhalation toxicity and lung cancer. Animal models used to investigate skin and bladder cancers for drinking water exposure are useful, but are weaker for lung cancers since animal studies appear to be negative for lung cancer. The WOE statement for lung cancer is based on the human data.

Issues 2 and 3 combined (pages 33-34):

One reviewer expressed concerns about including the data from the study by Jones et al (2007) in the calculations for the URF. Therefore, this reviewer is still not convinced that the Jones et al (2007) study is comparable with others because it appears to be an outlier. Therefore, inclusion of Jones et al (2007) is still questionable. Other reviewers agreed that Jones appears to be an outlier. The panel reached consensus that Enterline et al. (1995), Lubin et al. (2000 and 2008), Jarup et al. (1989) and Viren and Silvers (1994) are all appropriate to use in calculating the URF and are all properly characterized. The panel agreed that using multiple studies provides a better characterization of the data than any single study.

TCEQ Response: The TCEQ concurs with the peer reviewers that multiple studies are more appropriate for calculating the URF and have included epidemiological evidence from Enterline et al. (1995), Lubin et al. (2000 and 2008), Jarup et al. (1989) and Viren and Silvers (1994) studies. Based on the peer-reviewer's comments the TCEQ did not include the Jones et al. (2007) study in the final URF calculations, although an analysis of the Jones et al. (2007) study was provided in an appendix to the DSD for information purposes only.

Issue 4 (Page 34-35):

In summary, the panel agreed that a linear approach is appropriate given the lack of information on MOA. The panel suggested that TCEQ consider conducting metaanalysis to improve the quantitative estimate if it can be done with minimal effort. However, the panel suggested that a meta analysis should be presented in addition to TCEQ's existing analysis, but should not replace what has been done already.

TCEQ Response: The TCEQ considered a linear approach to be appropriate given the lack of information on a definitive MOA. As the peer reviewers recommended, the TCEQ also conducted different meta-analysis (Section 4.2.4.6 Sensitivity Analysis with Various Meta-Analysis Procedures) in addition to the combined meta–analysis using the inverse weighting existing analysis (Section 4.2.4.5 Combined – Analysis Using Inverse Variance of the URFs to weigh Individual URFs).

Issue 5 (Page 36):

Please discuss the rationale for making restrictions (i.e., fixing the background hazard rates for unexposed workers) to dose-response models.

TCEQ Response: The DSD did not provide the rationale for making restrictions. However, two different sensitivity analyses were evaluated to address this issue. The first model had one intercept and one slope fit to the SMRs of the Tacoma, Montana and Sweden cohorts combined (refer to Figure 4 and related discussion) and the second model had five intercepts and one slope fit to the SMRs of the Tacoma, Montana and Sweden cohorts combined (refer to Figure 5 and related discussion). Refer to Table 21. The range of the MLE URFs is 1.48E-04 to 2.07E-04 per μg/m3 for the analyses shown in Table 20. The URF and 95% UCL on the URF corresponding to the second model are more reliable than the estimates based on the first model because the second model fits the data statistically significantly better than the first model. In addition, the first model is rejected because of lack of fit whereas the second model is not rejected as a model fitting the data.

Issue 6 (Pages 38-39):

Please discuss how a meta-analysis that combines the data before modeling can reflect heterogeneity in data sets due to different epidemiological settings, differences in co-exposures, differences in dose characterization, and differences in dose-metric, differences in follow up. The panel recommended that TCEQ conduct a simple meta-analysis first, and noted that some of these issues have already been addressed in the process of combining URFs from these studies.

TCEQ Response: Please see above response. Conducting the different sensitivity meta analyses addressed this issue. The meta-analysis that combines the arsenic epidemiological data before modeling can better reflect heterogeneity in the data sets due to differences in co-exposures and differences in epidemiological settings by allowing different intercepts for cohorts or sub-cohorts with different characteristics. In addition, the model that fits the combined arsenic epidemiological data with different intercepts and one slope has several statistical characteristics that indicate the homogeneity of the data; a) the model with multiple intercepts and one slope is

not rejected for lack of fit, b) the model with one intercept and one slope is rejected for lack of fit, c) the model with multiple intercepts and one slope fits the data statistically significantly better than the model with one intercept and one slope, e) the model with multiple intercepts and multiple slopes does not fit the data statistically significantly better than the model with multiple intercepts and one slope. Thus, the heterogeneity in the background rates of the combined data sets was addressed by fitting a model with multiple intercepts and one slope. The arsenic epidemiological data sets all had the same cumulative exposure and there was no need of dealing with heterogeneity caused by different dose metrics.

Issue 7 (Page 39):

Generally, the panel agreed that the approach and the URF value rounded to 2.2E-04 per $\mu g/m^3$ used by TCEQ was reasonable, because the value was not inconsistent with other studies and it was developed as an average among different dose groups. As a caveat, one panel member stated that since Lubin (2008) shows that the slope in the low dose group is clearly different, then perhaps combining slopes from different doses is not appropriate, especially since it is the slope in the low dose region that is of interest. However, the panel did not recommend that TCEQ take a different approach.

TCEQ Response: The DSD was not revised based on this comment. The reason for the choice of slope (beta) from the Lubin (2008) study has already been discussed in Section 4.1.10.2.4 Preferred β and URF Potency Estimate (Lubin et al. 2000, 2008), as follows:

> "However, the URF calculated using the full cohort (2.18E-04 per $\mu g/m^3$) is preferred because it includes more deaths and person years in the analyses and is slightly more conservative. There was approximately a 1.5-fold difference between the URF (MLE) compared to the URF (95% UCL) from the Lubin et al. (2008) full cohort study as compared to a 2.2-fold ratio difference for the full cohort with average arsenic concentrations of 0.29 mg/m^3, which indicates that the estimate based on the full cohort is better because it is less uncertain."

Issue 8 (Pages 39-40):

In general, the panel suggested that TCEQ should present an upper bound estimate in addition to the central estimate of URFs. While the panel all agreed that TCEQ should include upper bound estimate calculations in the DSD, two panel members felt strongly that the upper bound value should be the basis of the regulatory value.

TCEQ Response:. After careful consideration of reviewer comments, the TD decided to continue to use the central estimate for the reasons cited in the DSD. Two reviewers agreed, and two disagreed with this approach. Under the TCEQ guidelines (2006), an important consideration in determining the need to use upper bounds is, "when estimates of mortality are available rather than incidence because survival rates for different cancers vary." Lung cancer incidence and mortality rates are sufficiently similar to the respiratory cancer mortality rates as to be comparable for purposes of the TD's assessment (see revised Figure 2 of the DSD). The guidelines also add support to using central estimates, "when well-conducted meta-analysis based on several epidemiologic studies are performed, the risk calculation can be done with greater precision thus decreasing uncertainty." The final URF is derived using a meta-analysis

approach that combines URFs based on the preferred individual epidemiological studies. Though meta-analyses usually combine results of primary research, herein the meta-analysis combines URFs estimated from published data of primary epidemiological research studies.

Issue 9 (Pages 40-41):

The panel reached consensus that TCEQ should weight their studies by the inverse of study variance, since the measurement is more standard and is appropriate for this dataset. They determined that no additional weighting approaches were necessary.

TCEQ Response: Based on the peer reviewer's suggestions, the TCEQ used the inverse of study variance of the risk factor estimates for the weighting procedure to calculate the final URF. The individual URF's were weighted based on inverse variance rather than person years. The individual weighted URFs were then combined together to calculate a final URF. Simulation experiments have shown that the inverse-variance weighting results in minimum variance estimates while sample-size or person-years weighting results in less biased estimates (Sanchez-Meca and Marin-Martinez, 1998). Although, the number of person-years plays a role on the size of the estimated variance of the URF, the inverse-variance weighting is a more standard statistical procedure used in meta-analyses. The inverse-variance weighting meta-analyses require the estimated variance of the individual URFs. In addition, the TCEQ performed a sensitivity analysis with various meta-analysis procedures for comparison purposes (Section 4.1.10.6).

Issue 10 (Page 40):

One reviewer suggested in the uncertainty discussion, TCEQ should address uncertainty in exposure measurement more completely. Another reviewer suggested that TCEQ expand the comparison of its URF with USEPA's value cancer assessment. Specifically, this reviewer suggested that TCEQ should add specific discussion on the approaches that TCEQ used that differ from USEPA's approach and describe the quantitative effect these different approaches have.

TCEQ Response: The TCEQ has enhanced the uncertainty section of the DSD in regards to exposure assessment error and has also included a more comprehensive comparison of the TCEQ risk estimate to that of the USEPA's risk estimate. Because the TCEQ used published summary data for the arsenic risk assessment, evaluation of the uncertainty in exposure estimation is limited. However, the TCEQ made every effort to account and discuss any sources of uncertainties in the summary exposure data. For example, Appendix H evaluates and discusses the effect that the average exposure estimates assigned to the highest exposure group in the Swedish cohort has on the estimated model parameters. Also, Appendix I discusses the UK study (not included in the estimation of the final URF) and presents the results for three different exposure extrapolation procedures.

WRITTEN MAIN COMMENTS THAT HAVE NOT BEEN ADDRESSED IN THE TELECONFERENCE REPORT

Comment Page 7-8:

Clarity could be improved in the text. For example, TCEQ could put summary language in front of each study description to allow non-math-oriented folks to get a sense of what they are about to read. Appendix A includes an example of how TCEQ might do this.

TCEQ Response: The DSD was revised. The summary language provided by the peer reviewer has been included as Appendix D in the revised DSD.

Comment Page 8:

Much of the literature you cite here is very old. You have omitted lots of good studies and reviews. In addition; there are so many good books and review articles on arsenic carcinogenesis that it is surprising that some of them are barely or not at all mentioned in your document.

TCEQ Response: The TCEQ appreciates the reviewer's suggestion on additional articles and books. The DSD was revised to reference some of the suggested articles, but the DSD is a summary document and is not intended to include or cite reference to all articles on arsenic. We typically have two reference sections: References cited in the DSD and a list of references that were reviewed, but not cited in the DSD. The draft DSD did not contain these two references sections so the reviewers were not aware that the TCEQ reviewed numerous articles. The proposed DSD now includes both reference sections.

Comment Page 10:

It is worth mentioning that there is no doubt that inorganic arsenic is carcinogenic to humans if the exposure is via drinking water. Lungs respond to arsenic exposure with carcinogenesis, both oral and respiratory routes are carcinogenic

TCEQ Response: The DSD was not revised based on this comment. The DSD is designed to evaluate the inhalation route of exposure. It is well known that arsenic is carcinogenic via the oral route of exposure.

Multiple Comments Pages 11-13:

Q1: The use of an excess relative risk model (ERR), or multiplicative linear RR model, seems appropriate but it is not clear why it is not consistently tested against the common log linear model (also referred to in the text as a linear exponential model), where the log RR is a linear function of exposure.

Q2: A second question regarding modeling is why only models linear in exposure are used. For data where risk attenuates at higher exposures a two piece linear model is a nice alternative for example.

A simple graph of the categorical points might help the reader be more convinced that a linear model is a good one.

Q3: Adjusting for time at hire, as is done in 2/4 studies, seems questionable to me for two reasons. First, I can't see any data to indicate the time since hire is significantly associated with outcome (is the coefficient significant, is it a potential confounder?), although I suspect it is, as lung cancer rates have changed over time. Second, time at hire may often serve as a surrogate for exposure and hence should not be controlled a priori. Time at hire may be associated with lung cancer rates because of changing smoking rates over time, but the use of an intercept term to adjust background rates is supposedly capturing differences between working cohorts and general population referent groups.

Q4: While there appears to be some smoking data in at least two studies (Lubin et al.2006 Jarup et al.1991), it is not used. Further discussion of the data on this key potential confounder should be provided. Why did Welch et al. and Higgins et al. in the early1980s conclude that smoking was not likely to be a confounder in the Montana cohort? In the Jarup et al. (1991) data it is said there is an interaction between arsenic and smoking. Generally such interactions are hard to ascertain and usually nonsignificant due to the rarity of lung cancer among non smokers. On the other hand, smoking (as noted above) is often a confounder between worker and general populations, as workers smoke more. If data were available on smoking rates in the exposed and referent populations, even if not known for the entire cohort, an adjustment to the SMRs can be made (eg, Steenland and Greenland AJE 2006).

TCEQ Response to Q1: The linear-exponential model is not the same as the log-linear model. The linear-exponential model defined by Lubin et al. (2008) refers to a model where $RR=1+\beta\times d\times\theta_f\times\exp\{\phi_f\times\ln(c)\}$ where d is cumulative exposure, c is arsenic concentration, θ_f is the effect of time since last exposure, and ϕ_f is exponent of the concentration (ϕ_f can be a constant or dependent on the time since last exposure), where the subscript f denotes the category. That is, the model is linear with cumulative exposures but the slope depends on the concentration of arsenic raised to a power dependent on the time since last exposure. The TCEQ used only linear RR models to fit the epidemiological data.

The question could also be referring to the analyses for the Jones et al. data. (Although this study was not used in the derivation of the final URF, the analyses based on this study are in Appendix I). There, the model used by the TCEQ is also a linear RR model. However, the dose metric includes some exponential functions. That is, the weighting factors of the concentrations used to calculate the weighted cumulative exposure are quadratic functions of age and the time since exposure.

The TD did not fit models other than the linear rate ratio multiplicative model.

For the Lubin et al. data, observed and expected number of respiratory cancer deaths were given for different categories of cumulative exposure to arsenic (mg/m^3-years). Thus, the TCEQ's only option was to fit a linear RR model to the Lubin et al. data using cumulative arsenic exposure as a dose metric. The data to fit the model using the weighted cumulative exposure to arsenic were not available to the TCEQ.

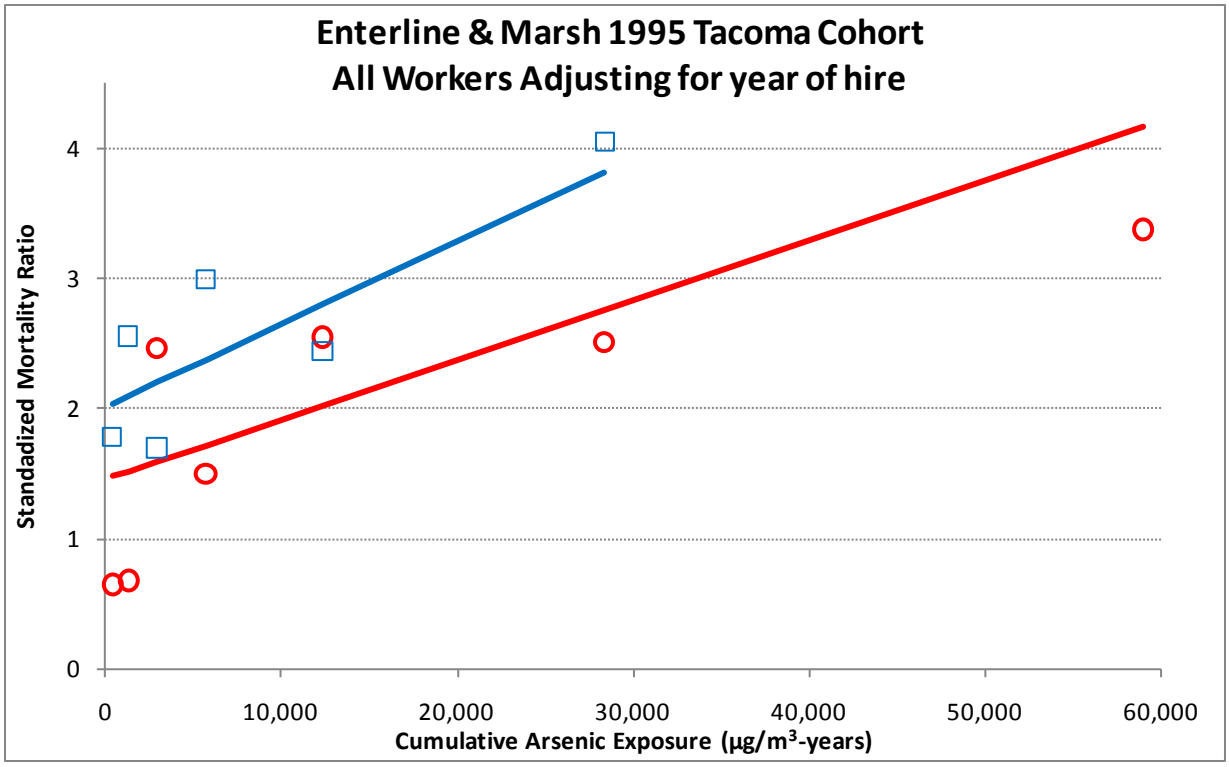
For the Jones et al. data observed and expected number of lung cancer deaths were given for different categories of weighted cumulative exposure to arsenic (mg/m^3-years) only. Thus, the TCEQ's only option was to fit a linear RR model to the weighted cumulative exposure to

arsenic. The data to fit the model using the un-weighted cumulative exposure to arsenic were not available to the TCEQ.

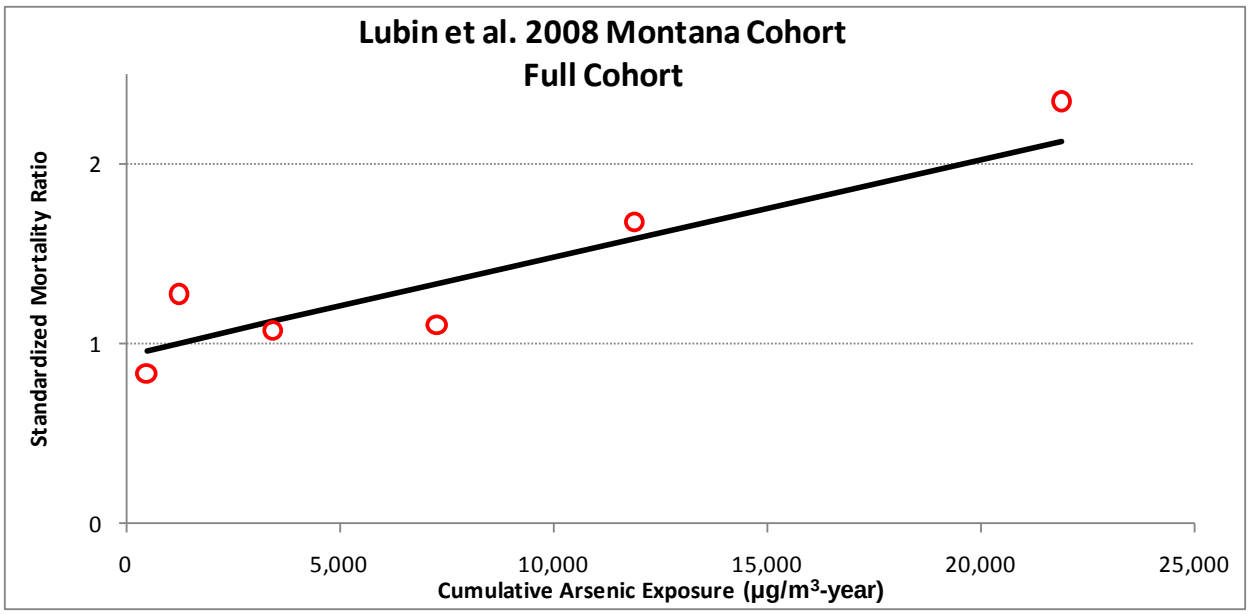
Q2: A second question regarding modeling is; why are only models linear in exposure used? For data where risk attenuates at higher exposures, a two piece linear model is a nice alternative for example. Granted only summary data (with a few categorical points) are available, limiting model possibilities, but nonetheless it might be possible to try other models (square root transformation, two piece linear, etc). A simple graph of the categorical points might help the reader be more convinced that a linear model is a good one.

TCEQ Response to Q2: None of the epidemiological data analyzed for arsenic seem to suggest any specific nonlinear behavior that warrants a model more sophisticated than a linear model. There are no known biological bases (MOA) that suggest any specific form of the relationship between exposure to arsenic and lung cancer mortality. The TCEQ guidelines (2006) on Page 3 state that "…when MOA information indicates that carcinogenic effects may not follow a linear pattern below the dose range of the observed data, nonlinear methods for determining risk at low dose may be justified." On page 34, TCEQ guidelines state "For carcinogens, when the MOA information supports linearity (i.e., the case for a carcinogen operating via a mutagenic MOA or when MOA is not understood), a linear default approach is used. (Figure 2-5)." In addition the 2005 USEPA guidelines state on page 3-11 (emphasis added) "For epidemiologic studies, including those with grouped data, **analysis by linear models in the range of observation is generally appropriate unless the fit is poor**. The relatively small exposure range observed in many epidemiologic studies, for example, makes it difficult to discern the shape of the exposure- or dose-response curve. Exposure misclassification and errors in exposure estimation also obscure the shape of the dose-response curve. When these errors are unsystematic or random, the result is frequently to bias the risk estimates toward zero. When a linear model fits poorly, more flexible models that allow for low-dose linearity, for example, a linear-quadratic model or a Hill model (Murrell et al., 1998), are often considered next." The TCEQ tested for lack of fit the linear model fit to the three epidemiological data sets used for the final URF – that is, the TCEQ tested the decrease in the likelihood from the full model to the linear model. The model fits to the Tacoma cohort and the Montana cohorts were not rejected at the 5% significance level but the model fit to the Swedish cohort had a slightly significant lack of fit with a p-value of 0.037. The following graphs show the SMRs and the models fit to the SMRs for the three studies included in the derivation of the final URF.

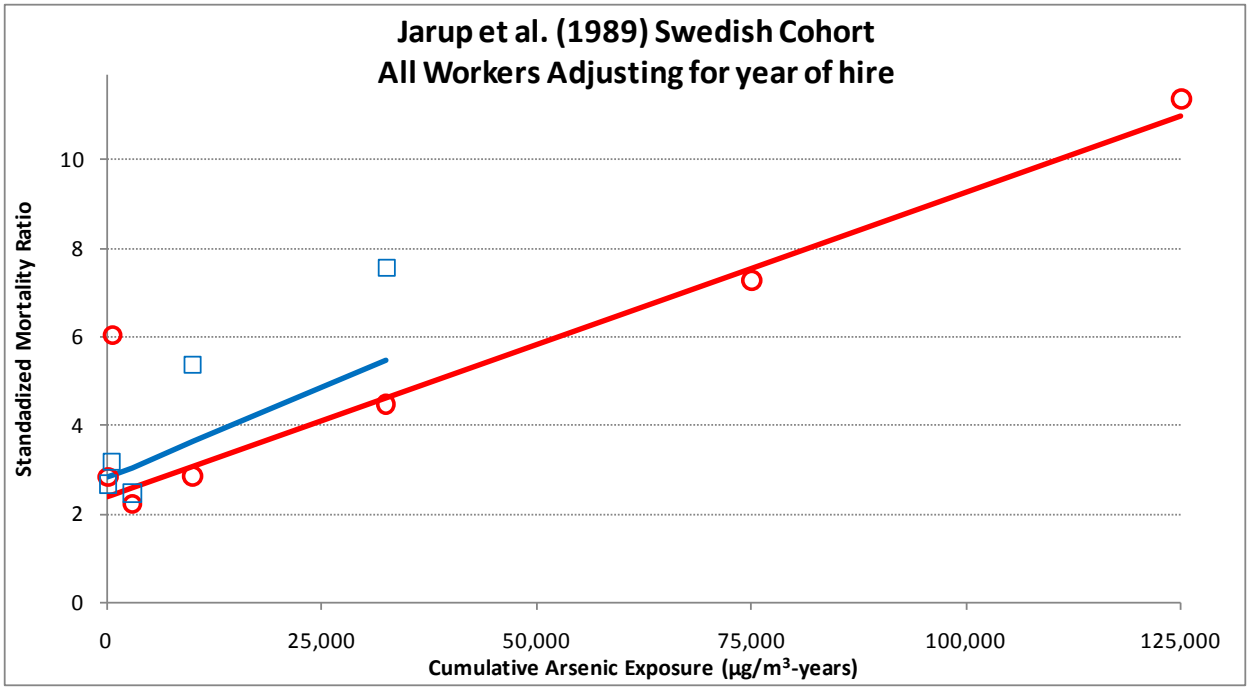
Enterline & Marsh 1995 – All workers adjusting for year of hire



Lubin et al. 2008



Jarup et al. (1989) and Viren & Silvers (1994)



Q3: Adjusting for time at hire, as is done in 2/4 studies, seems questionable to me for two reasons. First, I can't see any data to indicate the time since hire is significantly associated with outcome (is the coefficient significant, is it a potential confounder?), although I suspect it is, as lung cancer rates have changed over time. Second, time at hire may often serve as a surrogate for exposure and hence should not be controlled a priori. Time at hire may be associated with lung cancer rates because of changing smoking rates over time, but the use of an intercept term to adjust background rates is supposedly capturing differences between working cohorts and general population referent groups (see below).

TCEQ Response to Q3: The TCEQ did not adjust for the "time at hire" in any of the studies. Rather, the studies by Lubin et al. (2008) and Jones et al. (2007) included an element of the "time since exposure" as part of the dose metric. A dose metric similar that include "time since exposure" has been previously used also and published in BEIR VI (NRC 1999). The TCEQ used the best available data and published assessments that had been peer reviewed by experts in the corresponding fields.

The intercept term accounts for discrepancies in background rates between the general population and the workers in the cohort. The "time since exposure," however, is not another adjustment on the different background rates but, rather, an adjustment on the dose metric.

Or, the reviewer may be talking about the year of hire adjustment in the Enterline & Marsh (1995) Tacoma cohort and in the Jarup et al. (1989) Swedish cohort. Enterline & Marsh (1995) indicate that "The reason for this stratification is that for the cohort hired before 1940 only the person-years accumulated from 1941 were followed up for deaths, whereas for the cohort hired in 1940 and later all the person-years, to the end of the follow up period, were assessed. The stratification to some extent also separates workers before 1940 with relatively high exposure and with poor respiratory protection from workers with lower exposure, but with better exposure data and perhaps better respiratory protection." Thus, there are reasons to believe that the SMRs for workers hired before 1940 could be different than the SMRs for workers hired in 1940 or later. There may be other factors that Enterline & Marsh do not cite such as co-exposures to different chemicals before 1940, the lack of protection before 1940 did not protect from exposures to other potential contributors to lung carcinogenesis. Furthermore, the TCEQ does present the estimates for both; models adjusting for the year of hire and models not adjusted for the year of hire. The slopes (betas) of the adjusted models are between 1.2- and 1.5-fold greater than the slopes (betas) of the unadjusted models.

Q4: There seems to be some problem in the modeling of differences between background referent rates and cohort rates independent of an exposure effect, for example in the estimation of 'alpha'. We know that smoking differences between working cohorts and external referents can produce SMRs between 1.1 and 1.4, with the more common situation being about 1.2. Other factors which might drive up cohort rates would be use of inappropriate national rates rather than state or local rates which might be higher, although such differences are usually not great, on the order also of 20%. (no data are given on what referent rates are used in these studies). Yet here we find 'alphas' of 1.5-2.5 in Tables 8 and 14, much higher than can be explained by smoking differences (no 'alphas' are presented for Montana cohort analyzed by Lubin et all; why - it would appear to be small or even negative as per data in Table 10? The alpha for the Jones et al cohort is a more reasonable 1.2-1.3. These estimates of alpha are likely driven by the lowest exposure category, which have high SMRs in the Tacoma and Swedish cohorts (see Tables 7 and 13, SMRs of 1.54 and 2.71, compared to SMRs of 0.84 and 1.25 for Montana and English cohorts).

The use of these high (and ostensibly implausible) 'alphas' for two cohorts biases exposure-response coefficients downward, I believe. I can think of two possible solutions. First, use a priori an alpha of 1.2 (based on known likely effects of confounding by smoking) and perhaps take a Bayesian approach where this is an a priori parameter with a distribution. Second, and the course I recommend, is to do a meta-analysis rather than analyze each study separately. This will provide a different a more reasonable estimated 'alpha', given that the Montana (Lubin et al.) study is the largest study and likely has the lowest alpha.

TCEQ Response to Q4: The TCEQ makes the best use of the published data and makes every effort to obtain the best unbiased estimates of risk. The models fit to the data were estimated without forcing any preconceived ideas of what the values of the estimates should be or should not have been. The model parameters were estimated using maximum likelihood estimation. The different background rates between workers in a cohort and the reference population may be due to several reasons, as indicted by the reviewer. The reviewer, however, missed some other reasons that are important to consider; e.g., co-exposures to other potential lung carcinogens

(known or unknown) in the workplace. The solution proposed by the reviewer calls for the introduction of bias in the model to satisfy preconceived background rates. This strategy will violate the TCEQ goal of unbiased risk assessments. In addition, the reviewer's claim that for two cohorts the values of the estimated intercept "bias" the "exposure-response coefficients downward" is incorrect and unfounded. The parameters (both the intercept and the exposure-response coefficient) are unbiased estimates, but introducing any of the reviewer's proposed "possible solutions" do introduce bias in the exposure-response coefficients. Any manipulation of models or model parameters to accommodate preconceived background rates is bound to introduce bias. The high background rates had already been discussed by the authors of the original studies. The ratio between the cohorts's observed cancer rate in unexposed workers and the reference population's background cancer rates ranged between 1.43 and 2.05 for the Asarco smelter workers and between 2.37 to 2.67 for the Swedish workers. Viren and Silvers (1999) obtained similar results for the Asarco cohort and concluded that the model was "strictly linear with evidence of highly significant background lung cancer risk (fitted intercept), suggesting that other factors unrelated to arsenic effected the lung cancer outcome among these workers." Järup et al. (1989) also noted that the background rates in the unexposed workers of the Swedish cohort were substantially higher than the background rates in the reference population. They concluded that the differences could be due to "other occupational exposures to carcinogenic substances" or to different personal lifestyles (Doll 1985).

Q5: In general I believe a meta-analysis would be preferable to the current approach. Note the meta-analysis done by Crump et al. (2003) in a similar analysis of SMR data for dioxin studies, which incidentally estimated alpha as 1.17. Instead of weighting the individual slopes at the end, instead one estimates a common slope in the meta-analysis, potentially avoiding the problem of the aberrant alphas. Of course one would also be able to test heterogeneity in the different slopes of each study, and various functional forms besides linear, with potentially more power due to more data.

TCEQ Response to Q5: The TCEQ has derived the final URF performing a meta-analysis based on the three studies with cumulative exposure data (See section 4.1.10.5 Combined – Analysis Using Inverse Variance of the URFs to weigh Individual URFs). In addition, the DSD includes a section (4.1.10.6 Sensitivity Analysis with Various Meta-Analysis Procedures) that presents the results of other meta-analyses; including a meta-analysis where a model is fit to the combined data as in Crump et al. (2003). This meta-analysis includes tests for homogeneity of intercepts and slopes. The likelihood-based statistical tests indicated that a model with multiple intercepts (an intercept for each cohort or sub-cohort) and a single slope fit the data adequately.

Q6: While there appears to be some smoking data in at least two studies (Lubin et al, Jarup et al), it is not used. Further discussion of the data on this key potential confounder should be provided. Why did Welch et al. and Higgins et al. in the early 1980s conclude that smoking was not likely to be a confounder in the Montana cohort? In the Jarup et al. data it is said there is an interaction between arsenic and smoking. Generally such interactions are hard to ascertain and usually non-significant due to the rarity of lung cancer among non smokers. On the other hand, smoking (as noted above) is often a confounder between worker and general populations, as workers smoke more. If data were available on smoking rates in the exposed and referent

populations, even if not known for the entire cohort, an adjustment to the SMRs can be made (e.g., Steenland and Greenland AJE 2006).

TCEQ Response to Q6: Jarup et al. (1991) state "In conclusion, the present study confirms the strong dose-response relation between arsenic exposure and lung cancer found in our previous cohort study also after control of smoking. There was suggestive evidence of negative confounding by smoking in the highest exposure group. Our study also indicates that the interaction between arsenic exposure and smoking in relation to lung cancer is less pronounced among heavy smokers." The TCEQ considered the data in Jarup et al. (1989) more appropriate for dose-response modeling than the data in Jarup et al. (1991). The data in the 1989 article includes the whole cohort (3916 workers) (and is not adjusted for smoking for the obvious reason that there are no reference population background rates for smokers and non-smokers) while the 1991 article includes a subset of the cohort (321 workers) and odds ratios that adjust for smoking.

Lubin et al. (2000 and 2008) do not have any data on smoking that can be used to adjust the dose-response modeling of arsenic exposures.

Minor points Page 13:

Page 20. The author should explain the difference between added risk and extra risk (perhaps the latter is (R1-R0)/R0?). P 22. "supports the good fit of the Enterline data"? Narrow confidence intervals say the exposure term is important and well estimated, but say little about how well that particular model fits the data, compared to other possible models. I would modify this language..

TCEQ Response: Air concentrations based on extra risk were calculated as opposed to added risk. The relationship between lung cancer mortality and exposure to arsenic was evaluated based on healthy male workers employed in smelters. Although these workers were often healthier than the general population, the approach used by TCEQ estimates how the risk of lung cancer mortality changes with exposure to arsenic after adjusting for the differences between the workers and the general population background lung cancer mortality rates. The estimates of excess risks based on the derived models apply to the target population (e.g., Texas all sexes and all races, Texas white males, U.S. black females, etc.) whose background lung mortality cancer rates and survival probabilities are used in the estimation of the extra risks. The assumption being made in the calculation of the URFs is that the increase in the excess risk per a unit increase in the dose metric (i.e., cumulative exposure or weighted cumulative exposure to arsenic) is the same for the workers and for the target population. Subpopulations with higher background lung cancer mortality rates will have higher estimated URFs.

Comment Page 13:

Jarup et al found an effect of exposure intensity but not duration. Some further explanation should be provided why Texas should calculate an exposure-response using cumulative exposure (i.e, it is the metric used in the other studies and it is significantly related to lung cancer, presumably because it is correlated with average intensity).

TCEQ Response: Arsenic concentration influenced the outcome more than duration in the combined metric of per years. This is in line with Lubin's idea that concentration is an effect

modifier. Cumulative exposure is readily available for three out of the four epidemiological studies with sufficient data for dose response modeling. Cumulative exposure is the most commonly used dose metric in epidemiological studies and it captures the exposure intensity as well as the exposure duration. Jarup et al. for the Swedish cohort, Lubin et al. for the Montana cohort and Jones et al. for the UK cohort suggested an important effect of arsenic concentrations. Jones et al. used a weighted cumulative exposure concentration where the concentration was weighted by time of exposure and age. Lubin et al. also suggested a cumulative exposure where the cumulative exposure was weighted by the average arsenic concentration. These dose metrics, however, were considered inappropriate and not generally accepted the peer reviewers. The TCEQ, however, included the discussion of these alternative dose metrics in the appendices.

Comment Page 14:

Reviewer 1: TCEQ appeared to make the correct choice URFs from each study. However, in at least one study, the slopes are different among concentrations over time. For example, see Lubin et al. (2008) Figure 1 (page 31 of appendices). If this is true of other epidemiology studies, why should TCEQ settle for one slope per study? Alternatively, why not pick the slope associated with the lowest concentration, since this is the one most likely to be appropriate for low dose extrapolation?

TCEQ Response: The DSD was not revised based on this comment. The reason for the choice of slope (beta) from the Lubin (2008) study has already been discussed in Section 4.2.4.2.1 Preferred β and URF Potency Estimate (Lubin et al. 2000, 2008), as follows:

> "However, the URF calculated using the full cohort (2.18E-04 per $\mu g/m^3$) is preferred because it includes more deaths and person years in the analyses and is slightly more conservative. There was approximately a 1.5-fold difference between the URF (MLE) compared to the URF (95% UCL) from the Lubin et al. (2008) full cohort study as compared to a 2.2-fold ratio difference for the full cohort with average arsenic concentrations of 0.29 mg/m^3, which indicates that the estimate based on the full cohort is better because it is less uncertain."

The TCEQ strives to protect the general population from health hazards. The Lubin et al. URF based on the full cohort was selected as the URF instead of the URF based on the sub-cohort exposed to 290 $\mu g/m^3$ because 1) it is based on more data and, therefore, reduces uncertainty, 2) it is more-health protective, 3) it is more in line with the URFs estimated from the other studies, 4) it is between the URF based on low concentrations and the URFS based on high concentrations.

Reviewer 4: I don't think the correct slope factor (beta) was used for the Lubin study. Lubin et al. found an interaction with, or effect modification by concentration, the beta for cumulative exposure being much higher for higher average exposures. Since the environmental exposures of interest to Texas are clearly in this low range, it would seem appropriate to use the coefficient for the low average exposure group.

TCEQ Response: See Section 4.2.4.2.1.2.1 in the Arsenic DSD-Concentration as an Effect-Modification Factor: The dose-response relationship used by Lubin et al. (2008) uses concentration as an effect-modification factor rather than as a covariate. A covariate effect is

generally used to account for differences in background hazard rates of different groups of person years. An effect-modification factor, on the other hand, is used to model how the excess hazard rate changes due to the effect-modification factor. The covariate effects are normally excluded in the estimation of excess risks and the background risks of a target population are used instead. The effect-modification factors, on the other hand, are kept in the estimation of excess risks because they describe how the risk changes with these factors. One can think of these effect-modifying factors as part of the dose metric. The usual dose metric in dose-response models for epidemiological data is cumulative exposure. Lubin et al. (2008), however, used a dose metric that is equal to the cumulative exposure multiplied by the average concentration over the exposure period raised to a power.

The TCEQ strives to protect the general population from health hazards. The Lubin et al. URF based on the full cohort was selected as the URF instead of the URF based on the sub-cohort exposed to 290 $\mu g/m^3$ because 1) it is based on more data and, therefore, reduces uncertainty, 2) it is more-health protective, 3) it is more in line with the URFs estimated from the other studies, 4) it is between the URF based on low concentrations and the URFS based on high concentrations.

Multiple Comments Page 18:

Thus, TCEQ is not limited to the consideration of only one MOA and may elect to consider different MOAs at different parts of the dose response curve. An example of this for acrylamide is found in Dourson M., Hertzberg, R., Allen, B., Haber, L., Parker, A., Kroner, O., Maier, A. and Kohrman, M. 2008. Evidence- Based Dose Response Assessment for Thyroid Tumorigenesis from Acrylamide. Regulatory Toxicology and Pharmacology 52 (2008) 264–289.

In addition, perhaps I am confused but I was perplexed by the authors citing older USEPA (1984) text on MOA (page 10). USEPA (2005) would let TCEQ approach the dose response assessment in several ways, including linear, or linear with an upper bound determined by a threshold. I can step TCEQ through the latter approach if this is desired. Or you could refer to USEPA (1998) where this is described for thyroid tumors. See: USEPA. 1998. Assessment of thyroid follicular cell tumors. EPA/630/R-97/002. March.

TCEQ Response: TCEQ agrees with the reviewers and the appropriate changes were made in the DSD.Older version of EPA.

Multiple Comments Page 18 Contd…:

Reviewer : I would be especially interested in the last difference, “the use of 2005 Texas-specific lung cancer mortality rates and survival probabilities compared to USEPA (1984) use of US 1976 lung cancer mortality rates and survival probabilities,” since other folks might want to use more generalized data, or data from their location.

TCEQ Response: USEPA developed a URF of 4.3E-03 per $\mu g/m^3$ in 1984 (USEPA 1984) which was reviewed by IRIS in 2007. IRIS did not recommend any changes to the URF value. The URF was based on excess lung cancer mortality in workers at only two smelters and is the final estimated geometric mean of the risk estimates from the Asarco smelter in Tacoma, Washington

(Enterline and Marsh 1982) and the Anaconda smelter in Montana (Brown and Chu 1982, 1983a, 1983b; Lee-Feldstein 1983; and an unpublished paper by Higgins and associates).

The TCEQ developed a URF of 1.5E-04 per $\mu g/m^3$ based on three cohorts that included updated estimates from Tacoma smelter (i.e. Enterline et al. 1987 and 1995 updates) and the Montana smelter (i.e. Lubin et al. 2000 and 2008) as well as estimates from the Ronnskar Copper Smelter cohort study in Sweden (Järup et al. 1989; Viren and Silvers 1994). The TCEQ is of the opinion that these three human epidemiological studies contain adequate dose-response data for an updated assessment of the carcinogenic potential of arsenic and the development of new inhalation unit risk factor (URF). The TCEQ's preferred URF of 1.5E-04 per $\mu g/m^3$ is less conservative than the USEPA's URF of 4.3E-03 per $\mu g/m^3$. Consequently, the resulting 10^{-5} risk air concentration for excess lung cancer mortality of 0.067 $\mu g/m^3$ based on the selected URF by TCEQ is approximately 29 times higher than the 10^{-5} risk air concentration of 0.0023 $\mu g/m^3$ based on the USEPA's URF of 4.3E-03 per $\mu g/m^3$ (USEPA 1984).

The difference in the URFs calculated by the TCEQ and USEPA are mainly due to the fact that TCEQ used updated and more accurate exposure estimates for the Tacoma and the Montana cohorts. The USEPA developed a URF from the Tacoma smelter based on the Enterline and Marsh (1982) study and did not use the updated estimates with additional years of follow-up (Enterline et al. 1987 and 1995 updates) when they developed and updated their inorganic arsenic assessment in IRIS in 1984 and 2007 respectively. The USEPA relied on the first Enterline and Marsh (1982) study that was itself based on the results from the Pinto et al. (1976) study. The Pinto et al. (1976) study reported an association between airborne arsenic concentrations and urinary arsenic concentrations. The basis of this relationship was that the urinary arsenic concentration could be used as a biomarker for airborne exposure, and the dose response for arsenic-related lung cancer mortality could be expressed in terms of cumulative urinary arsenic exposure (µg/As/liter urine years). This relationship was expressed with the following formula:

$$As_{air} = 0.304\ As_{urine}$$

where As_{air} is measured as $\mu g/m^3$ and As_{urine} is measured as µg/liter

Enterline and Marsh (1982) used this relation and estimated cumulative air exposure by multiplying the 1982 cumulative urinary arsenic exposure by 0.304. However, in 1987, Enterline and associates indicated limitations in the Pinto et al. (1976) study, conducted a re-analysis, and reported an updated relationship between airborne arsenic exposure and respiratory cancer mortality among workers from the Tacoma smelter in the following formula:

$$As_{air} = 0.0064\ (As_{urine})^{1.942}$$

where As_{air} is measured as $\mu g/m^3$ and AS^{urine} is measured as µg/liter.

Thereafter, Enterline et al. (1987) reported that in the Pinto et al. (1976) study the authors did not take into account prior arsenic exposure through diet. This resulted in high baseline levels of urinary arsenic (about 150 µg/liter). As such, Enterline et al. (1987) indicated that the Pinto et al.

(1976) study did not depict the true relationship between urinary arsenic measurements and airborne arsenic levels. The USEPA, however, did not use the updated results from Enterline et al. (1987) study. Therefore there is uncertainty associated with the USEPA's exposure assessment.

Similarly, for the Montana smelter, the TCEQ also used the updated exposure estimates that included additional years of follow-up with more person years and deaths (Lubin et al. 2000; 2008). The cumulative exposure estimates calculated by Lubin et al. (2000) based on duration in jobs with low and medium exposure concentrations and time of exposure in areas of heavy exposure were re-calculated using a weighting factor (γ) of 0.1 to take into account the reduction in exposure due to the use of respiratory or air filtration masks in heavy-exposure jobs. This resulted in more representative arsenic cumulative exposure estimates that were lower than the estimates using a weighting factor (γ) of 1.0 used previously, particularly at the highest cumulative exposures. Furthermore, using the weight of 0.1 on high-exposure jobs resulted in rate ratios that conformed to a linear dose-response relationship with cumulative exposure to arsenic.

The TCEQ used these updated estimates in the final URF determination because they represent more realistic exposure estimates for cumulative exposure for workers especially in heavy exposure areas where respirators were used. The USEPA URF did not include these updated exposure estimates, thus their URF is more conservative than the TCEQ URF.

In addition to differences in exposure estimates, there are other factors that could contribute to differences in the TCEQ and USEPA URFs including the following:

- the availability of the estimates from the Ronnskar Cooper smelter in Sweden
- the use of a 70-year default lifetime exposure (TCEQ 2006) versus 76.5 years (USEPA 1984).
- the use of 2005 Texas-specific lung cancer mortality rates and survival probabilities compared to USEPA (1984) use of US 1976 lung cancer mortality rates and survival probabilities.
- the TCEQ also used the Lubin et al. (2008) data and conducted standard Poisson regression analysis, a multiplicative model, and SMR data adjusted for calendar period and country of birth as opposed to USEPA (1986) who used an absolute risk model. The absolute risk is referred to as a "crude" risk and is not very useful to compare two populations.

Public Comments on Arsenic DSD

Dear Drs. Erraguntla, Grant and Lee,

I respectfully submit the following comments on your respective DSDs.

Overall

I applaud your use of reasonable, conservative hazard assessment methods in the development of the subject DSDs.

Arsenic DSD

Page 6 Line 29: "Arsenic naturally occurs in the earth's crust and is present in pesticides." Since most pesticides are not arsenicals, it would be appropriate to insert a qualifier, e.g., "…and is present in some pesticides."

- Page 7 Lines 8-13. A 2006 reference regarding widespread use of organic arsenicals as pesticides is provided. However, since 2006 significant and relevant changes have been made regarding authorized use of inorganic arsenical herbicides in the US. The EPA has determined that MSMA use in cotton is eligible for reregistration, but all other uses will be (or have already been) phased out and canceled. For more information please see http://www.epa.gov/oppsrrd1/reregistration/organic_arsenicals_fs.html.
- Page 22 Line 6: "Minimal Risk Value" should be "Minimal Risk Level"
- Pages 27, 71-74: Some information (equations?) is not visible.

TCEQ Response: The TCEQ appreciates the public comments received on the arsenic DSD and has made all the appropriate changes as suggested. The TCEQ has also resolved the issue with the equations not being visible.

Summary of the changes in the 3.1.6.2 section of the arsenic DSD.

1) Use of the latest version of the MPPD (version 2.1) software.
2) The TD also checked the "Inhalability Adjustment" box while running the MPPD version 2.1 software as per communications with Dr. Fredrick Miller. See Appendix A2
3) The key study (Holson et al. 1999) still remained the same as before. However, the TD made changes in the selection of the target regions to calculate the RDDR. The RDDR of the "pulmonary region" was selected as the appropriate output instead of the "total region" to develop a POD_{HEC} because the adverse effect noted in the key animal study is rales (i.e., labored respiration and gasping) and because the average particle size used in the study was 2.07 µm, that is generally associated with alveolar region and airways. Dr. Fredrick Miller an expert in dosimetry agreed with the TD's approach. (Per communications with Dr. Fredrick Miller)
4) Use of 214.2 ml/min for VE for animal instead of 137.3 ml/min. This value was adopted based on a peer reviewer's comment in the Nickel DSD. The reviewer had suggested using this value instead of 137.3 ml/min. The arsenic acute toxicity study was similar to the nickel acute study in that it was also conducted in rats and the TD was using the MPPD software for dosimetric adjustments. Further, based on the nickel DSD peer reviewer's comments the TD changed the inputs for human tidal volume and breathing frequency. Appendix A2 has detailed information on the reasons for the updated inputs and links to the nickel DSD.

Per Communications:

Dr. Fred Miller Response: "The source of the sound associated with rales clearly reflects that the problem is with the alveolar region tissue and airways. Thus, focusing on the pulmonary region for the study under consideration is quite appropriate".

So my opinion is that the alveolar/pulmonary region doses are the ones to use in this instance.

Best regards,

Fred

Frederick J. Miller, Ph.D., Fellow ATS
Fred J. Miller & Associates LLC
911 Queensferry Road
Cary, NC 27511-6422
phone: 919-467-3194
e-mail: fjmiller@nc.rr.com