Expert External Peer Review of the Development Support Document for Nickel Report of Conference Call

Peer Review organized by
Toxicology Excellence for Risk Assessment (*TERA*)
2300 Montana Ave, Suite 409, Cincinnati, OH 45211, 513-542-7475

December 2, 2009

PREPARED IN COOPERATION WITH THE
Texas Commission on Environmental Quality (TCEQ).
The preparation of this report was financed by Texas Commission
on Environmental Quality (TCEQ)

Table of Contents

This page intentionally left blank.

Introduction

The Toxicology Division of the Chief Engineer's Office, Texas Commission on Environmental Quality (TCEQ) has prepared a draft Development Support Document (DSD) that outlines the hazard assessment and dose-response processes used to derive Effects Screening Levels (ESLs), Reference Values (ReV), and a Unit Risk Factor (URF) for nickel. The toxicity values were developed using RG-442 *Guidelines to Develop Effects Screening Levels, Reference Values, and Unit Risk Factors* (TCEQ, 2006). ESLs are chemical-specific air concentrations set to protect human health and welfare. Short-term ESLs are based on data concerning acute health effects, odor potential, and vegetative effects, while long-term ESLs are generally based on data concerning chronic noncarcinogenic and/or carcinogenic health effects. ESLs are used in the evaluation of air permit applications as well as proposed rules and regulation (e.g. Permits by Rule). ReVs and URFs, used as the basis of ESLs, are used in the evaluation of air monitoring data and in the development of Protective Concentration Levels for remediation sites.

TERA is supporting the Texas Commission on Environmental Quality (TCEQ) in conducting an expert external peer review of the *Development Support Document for Nickel*. The review materials, including draft document, charge to reviewers, and key references (available at http://www.tera.org/Peer/nickel/nickel.html) were distributed to the panel in July 2009. Panel members reviewed the nickel DSD and submitted written comments that addressed the charge questions in August 2009. These written comments are presented in Part A of this report and the Peer Review Charge is attached as Appendix A.

On October 1, 2009, *TERA* facilitated a follow-up conference call between the panel and TCEQ. Conference call materials (available at the above website), including a focused charge, attached as Appendix B, and the reviewer comments in Part A, were distributed prior to the call; members of the public were allowed to listen to the call. The purpose of this call was to allow TCEQ to ask the panel questions regarding their written comments and to allow the panel members to discuss issues on which there were divergent opinions in the written comments. A *TERA* staff member took notes during the call to create a record of the panel's discussion and recommendation. This report of the conference call is presented in Part B of this report. Therefore, the written comments submitted by the panel and the report of the follow-up conference call comprises the complete peer review of the nickel DSD.

This page intentionally left blank.

Part A: Panel Written Comments

**Scientific Peer-Review of the
Noncarcinogenic and Carcinogenic Sections of the
Nickel Development Support Document (DSD)
Reviewers Comments**

General Issues

Please consider all aspects of the nickel DSD and evaluate strengths and weaknesses of the procedures used to develop acute and chronic toxicity values based on the specific questions described below. Where possible, try to put the strengths and weaknesses in perspective by indicating their relative magnitude. Please try to avoid emphasizing minor technical details or making tutorial comments. Reviewers should identify scientific uncertainties and suggest ways to reduce or eliminate those uncertainties.

- **Were procedures outlined in the ESL Guidelines followed by the TCEQ to perform nickel's toxicity assessment? If references to accepted procedures in federal, state, or other appropriate guidance documents were made in the ESL Guidelines, were those accepted procedures followed?**

Reviewer 1: *In general, yes, the procedures outlined in the ESL guidelines were followed in the DSD. An exception is the use of human data in the chronic toxicity and cancer assessment. On page 20, the ESL guidelines state: "When relevant human studies are not available, animal data are used to develop toxicity factors." Although the epidemiological data can be informative for supporting the development of ESLs, the studies relied upon for the chronic toxicity and cancer ESLs are not informative regarding the form of nickel which is associated with risk. In these cases, the animal data should be used instead.*

Reviewer 2: *Yes.*

Reviewer 3: *See question below.*

Reviewer 4: *I am not overly familiar with the procedures outlined in the ESL Guidelines, but examination of the overview of the steps presented in those guidelines suggests that they have been followed in the nickel assessment.*

Reviewer 5: *Yes, procedures outlined in the ESL Guidelines were followed by the TCEQ.*

- **Does the nickel DSD clearly describe the approaches used by TCEQ to perform the toxicity assessment (i.e., hazard identification and dose-response**

assessment)?

Reviewer 1: *Yes*

Reviewer 2: *Yes, except as noted below.*

Reviewer 3: *See question below.*

Reviewer 4: *I did not see in the DSD a description or overview of the approaches that are to be used by TCEQ doe such assessments. Nowhere was there a statement about any hazard identification steps. The document essentially starts with a review of the properties of nickel and nickel compounds and their main uses. Then, starting in Chapter 3, the remainder of the document is what I would characterize as dose-response assessment and risk characterization.*

Reviewer 5: *Most aspects of the toxicity assessment are clearly described. However, there are a few aspects that I think need clarification. For example, I believe that more discussion and justification are needed for characterizing the single exposure level in the Cirla study as a LOAEL, and I think the derivation of URFs and air concentrations for lung cancer incidence needs additional details. I will address those areas below in the proper context.*

- **Please identify any relevant studies or data that have not been cited. Explain how they may impact the assessment.**

Reviewer 1: *While the NTP (1996a, b, c) toxicity studies are all cited in the DSD, the data in these studies were not adequately reviewed. These studies provide data on the threshold for toxicity of the sulfidic, oxidic, and soluble nickel and can be used to calculate chronic ReV, URF, and ESL values.*

The DSD did not cite a recent study by Oller et al.: Oller, A. R.; Kirkpatrick, D. T.; Radovsky, A.; and Bates, H. K. (2008). Inhalation Carcinogenicity Study with Nickel Metal Powder in Wistar Rats. Toxicol Appl Pharmacol 233:262-275. This robust study in rats did not show evidence of carcinogenicity for metallic nickel

While the DSD does describe the toxicity of nickel, it does not fully explore the findings from the NTP and Oller et al. (2008) studies that soluble and metallic nickel were not carcinogenic in animals via inhalation, whereas insoluble forms were. Several studies (e.g., Benson et al., 1995a,b; Dunnick et al., 1995; Yu et al., 2001) describe differences between the forms of nickel in accumulation in and clearance from the lung, factors which help explain the carcinogenicity findings in these carcinogenicity studies. Citations for these studies are provided below, and are reviewed by Goodman et al. (2009).

Benson, J.M.; Chang, I.-Y.; Cheng, Y.S.; Hahn, F.F.; Kennedy, C.H.; Barr, E.B.; Maples, K.R.; Snipes, M.B. (1995a). Particle clearance and histopathology in lungs of F344/N rats and B6C3F1 mice inhaling nickel oxide or nickel sulfate. Fundam Appl Toxicol 28:232–244.

Benson, J.M.; Cheng, Y.S.; Eidson, A.F.; Hahn, F.F;, Henderson, R.F.; Pickrell, J. A. (1995b). Pulmonary toxicity of nickel subsulfide in F344/N rats exposed for 1–22 days. Toxicology 103(1):9–22.

Dunnick, J.K.; Elwell, M.R.; Radovsky, A.E.; Benson, J.M.; Hahn, F.F.; Nikula, K.J.; Barr, E.B.; Hobbs, C.H. (1995). "Comparative carcinogenic effects of nickel subsulfide, nickel oxide, or nickel sulfate hexahydrate chronic exposures in the lung." Cancer Res 55:5251-5256.

Yu, C.P.; Hsieh, T.H.; Oller, A.R.; Oberdorster, G. (2001). Evaluation of the human nickel retention model with workplace data. Regul Toxicol Pharmacol 33:165–172.

Reviewer 2: *A recent study of the inhalation carcinogenicity of nickel metal powder was not cited (Oller et al. 2008). As described in the context of the cancer weight of evidence, this animal study provides important information on sorting out the contribution of individual nickel species, in light of the coexposures seen with most of the epidemiology studies.*

.

***Oller AR**, **Kirkpatrick DT**, **Radovsky A**, **Bates HK**.*
Toxicol Appl Pharmacol. 2008 Dec 1;233(2):262-75.
Inhalation carcinogenicity study with nickel metal powder in Wistar rats.

Epidemiological studies of nickel refinery workers have demonstrated an association between increased respiratory cancer risk and exposure to certain nickel compounds (later confirmed in animal studies). However, the lack of an association found in epidemiological analyses for nickel metal remained unconfirmed for lack of robust animal inhalation studies. In the present study, Wistar rats were exposed by whole-body inhalation to 0, 0.1, 0.4, and 1.0 mg Ni/m(3) nickel metal powder (MMAD=1.8 microm, GSD=2.4 microm) for 6 h/day, 5 days/week for up to 24 months. A subsequent six-month period without exposures preceded the final euthanasia. High mortality among rats exposed to 1.0 mg Ni/m(3) nickel metal resulted in the earlier termination of exposures in this group. The exposure level of 0.4 mg Ni/m(3) was established as the MTD for the study. Lung alterations associated with nickel metal exposure included alveolar proteinosis, alveolar histiocytosis, chronic inflammation, and bronchiolar-alveolar hyperplasia. No increased incidence of neoplasm of the respiratory tract was observed. Adrenal gland pheochromocytomas (benign and malignant) in males and combined cortical adenomas/carcinomas in

females were induced in a dose-dependent manner by the nickel metal exposure. The incidence of pheochromocytomas was statistically increased in the 0.4 mg Ni/m(3) male group. Pheochromocytomas appear to be secondary to the lung toxicity associated with the exposure rather than being related to a direct nickel effect on the adrenal glands. The incidence of cortical tumors among 0.4 mg Ni/m(3) females, although statistically higher compared to the concurrent controls, falls within the historical control range; therefore, in the present study, this tumor is of uncertain relationship to nickel metal exposure. The lack of respiratory tumors in the present animal study is consistent with the findings of the epidemiological studies.

Reviewer 3: *The TCEQ appears to have followed the basic template set out in the ESL Guidelines. However, the Guideline are said to be based upon the traditional 4-step risk assessment process, which includes a thorough hazard assessment. The TCEQ has not performed a thorough hazard assessment and has instead focused primarily on key studies. This detracts from transparency and obscures the process by which the key studies were chosen.*

The TCEQ has relied heavily on material from secondary sources, especially the ATSDR and ICNCM documents. This reliance on authoritative secondary sources is understandable, in that it provides a cost-effective evaluation of the literature. However, this approach detracts from reader confidence in TCEQ expertise and is especially a problem when the TCEQ departs from the conclusions and recommendations of these secondary sources that are otherwise heavily relied upon. For example, the ATSDR determined that there was insufficient data to derive an acute MRL for inhalation, yet the TCEQ proceeded to derive one that was even lower than the conservative estimate derived by CalEPA. The authors of the TD need to make it clear that they are departing from federal precedent, and provide a rationale for this departure.

An almost universal aspect of the TD is that reference values developed by TCEQ are more conservative than those of otherwise conservative agencies, such as ATSDR, US EPA, and Cal EPA. The authors of the TD need to justify this extreme conservatism, especially given the fact that (as already stated) the TCEQ apparently places great confidence on secondary sources from these agencies.

The TCEQ has not attempted critical review of the epidemiologic literature, which suggests a lack of familiarity with the limitations of this discipline. Only limited critical assessment has been provided, and this is usually taken directly from the studies themselves or from the ATSDR and ICNCM documents, with the TCEQ accepting these arguments at face value. Lack of critical review is especially of concern for the cancer epidemiology, which consists of a large number of occupational studies with varying levels of quality and widely varying results.

Misuse of epidemiologic jargon also detracts from reader confidence. For example, page 26 of the TD document concludes that data are "confounded by poor Ni exposure," when this is actually a case of exposure misclassification rather than confounding. Similarly, page 37 of the TD indicates that use of Ni equivalents alleviates misclassification. However, combining all Ni exposure into one, without knowing the important species, does not remove misclassification, and actually enhances it to some degree. The authors of the TD also suggest that the "Grimsrud et al. (2003) cohort is more reliable because it includes greater than seven times more workers than the … case-control study" (p 33). However, this statement does not consider the inherent differences between the case-control and cohort approaches. Case-control studies rarely have the same number of individuals as cohort studies, because cohorts include all workers whereas case-control studies deal only with disease cases and a subset of noncases. The Grimsrud et al. (2002) study is a case-control approach nested within the larger cohort, and therefore has comparable numbers of cases and comparable power

The TCEQ has relied upon 'Ni equivalents' in all their assessments. However, I was somewhat confused about how these values were derived and whether or not they make the assumption that all nickel species as equivalent based on nickel content (even though soluble nickel is most toxic). The method and assumptions underlying Ni equivalents should be discussed, either within the body of the text or as an appendix.

Reviewer 4: *I know of no other data or studies that have not been cited.*

Reviewer 5: *I am not aware of any relevant studies or data that have not been cited.*

- **Other General Issues**

Reviewer 1: *NA*

Reviewer 2: *Overall, this was a very thorough assessment of a complex chemical and complex series of issues. The document overall was well-written. with clear presentation of the underlying rationales. The analyses were also generally thorough, taking into account the nature of the process, in building off of existing reports. I was pleased to see the level of sophistication in considering issues of dosimetry, and in the analysis of the epidemiology data. The document was well-edited. I have no editorial comments, and congratulate the authors on catching typographical errors.*

I have just one significant comment, related to the cancer assessment (along with some smaller comments on the acute ESL, as noted below). As note by the assessment authors, total nickel can be used as the dose measure ***if*** *the study population in the epidemiology study(s) used for the quantitative assessment was*

exposed to a mixture of nickel compounds similar to the mixture in Texas air. However, information presented in the DSD about the Enterline and Marsh (1982) study, and issues raised by Goodman et al. (2009) about methods for speciation of nickel (and implications of that analysis for understanding of the Grimsrud et al. (2003) study and other studies of the Kristiansand cohort) suggest that neither of the cohorts used for the quantitative assessment were exposed to mixtures comparable to that in Texan air. Instead, it is likely that both were exposed to substantially more nickel subsulfide than in Texas air, which would result in an over- estimation of cancer risk for the Texas population. I apologize that this issue is presented a bit repetitively in this response to the charge questions, since I wanted to be sure to answer each charge question completely.

Reviewer 3: *NA*

Reviewer 4: *NA*

Reviewer 5: *NA*

Health-Based Acute ReV and ^{acute}ESL

Chapter 3 of the nickel DSD describes the approaches used to derive the health-based acute ReV and acute ESL. The key decision points are listed below. For each decision, please consider the scientific defensibility of the decision and any additional approaches or analyses, or additional information that could improve that decision. Please indicate if there are other issues specific to developing acute toxicity factors that have not been adequately addressed in the document.

- **The choice of the critical study (Cirla et al. 1985):**
  - **Was the human study selected for the non-cancer estimates the most appropriate study? Was the form of nickel selected (nickel sulfate) for deriving the ReV/ESL appropriate given the purpose of these values?**

Reviewer 1: *Cirla et al. (1985) is the most appropriate study. Bronchial constriction after acute exposure has only been observed with soluble Ni. It is not clear whether this endpoint is applicable to other forms of Ni, although based on chronic studies, all forms of Ni have similar types of respiratory toxicity, with soluble nickel being the most toxic. Thus, using soluble nickel is a reasonable for a "worst-case" scenario.*

Reviewer 2: *Nickel sulfate is the correct form for deriving the ReV/ESL. However, I would recommend that the Cirla study be a co-principle study. As I understand the study and the method, bronchial provocation testing (BPT) determines the presence of allergic asthma; people who respond have already been sensitized. Further, as I understand the ESL guidelines, the acute ESL is designed to protect from becoming sensitized, but may not protect those who already have been sensitized. Thus, as I understand it, the Cirla study is actually evaluating an*

effect of chronic exposure (sensitization) using the acute challenge, rather than being a true acute toxicity test. Note that the study population worked with nickel and had already reported allergic/asthmatic symptoms. It is not clear, however, why 3/6 subjects who were positive in the BPT were negative for nickel IgE; I would be interested in the comments of those with more expertise in this area than I. Nonetheless, this study indicates that an acute ESL based on the Graham study would nearly be protective for sensitized individuals. Alternatively, TCEQ may make a policy decision that it is appropriate to base the acute ESL on protecting sensitized individuals. This is not standard practice for sensitizing agents under the ESL Guidelines (as I understand them), or general risk assessment practice, but may be an appropriate decision based on the relatively high percentage of the population that is dermally sensitized to nickel.

Reviewer 3: *The authors of the TD provide no comprehensive review or enumeration of the available literature for this endpoint. They cite several other papers dealing with occupational asthma from Ni exposure, as do both ATSDR and Fernandez-Nieto et al., but these papers are not discussed in any substantive manner. The TCEQ should summarize the strengths and limitations of these various papers and present specific criteria for why Circla et al. 1985 was selected as the critical study.*

Reviewer 4: *The study in question was limited in terms of sample size and because there was only one exposure level. It is difficult to identify a LOAEL (as purported in the DSD for Cirla et al.) when there are no other exposure levels to be used for comparison (e.g., how do we know this is a low adverse effect level when still lower exposures could have induced a significant response). This is the greatest concern about using this study, especially since the response rate was 6 out of 7 among asthmatics. That is, it is somewhat difficult to imagine that the single, arbitrarily chosen exposure level from this study really constitutes a low adverse effect level.*

On the other hand, I understand the desire to use a human study as the basis for the assessment, if at all possible. The use of uncertainty factors (and in particular the factor of 3 for data base deficiencies) may offset the limitations of the selected study.

I agree with the choice to use nickel sulfate as the basis for the derivation.

Reviewer 5: *Two human studies were described. Cirla et al. (1985) exposed 12 workers (eight men, four women) from a nickel plating operation to nickel sulfate hexahydrate at 300 $\mu g/m^3$ for 30 minutes. Seven of the twelve subjects were confirmed asthmatics. The investigators used a greater than 15% decrease in forced expiratory volume in one second (FEV_1) as a positive response for significant bronchoconstriction. Six of the twelve (all asthmatics) had positive responses (a 50% response rate). The other human study discussed was that of Fernando-Nieto et al. (2006), in which only 4 male workers were exposed to*

nickel sulfate AND potassium dichromate. Although several concentrations were used (increasing every 24 hours), exposure at each concentration was for only 2 minutes. Two of the four workers (50%) experienced a fall of 20% or more in FEV_1 at the highest exposure levels. Although this study supports the Cirla study, the DSD indicated that reliable airborne concentrations of nickel were not available. The Cirla et al. (1985) study appears to be the most appropriate study, even though it had only a single exposure level.

The form of nickel selected for deriving the Rev/ESL (nickel sulfate) is appropriate. As explained in the DSD, soluble forms of nickel (e.g., nickel sulfate) are more toxic than the insoluble forms so that nickel sulfate can serve as a conservative surrogate for all inorganic forms of nickel.

- **The choice of critical effect (greater than a 15% decrease in FEV_1 in six of seven asthmatics):**
  - **Was the most appropriate critical effect selected? If not, what would be a more appropriate critical effect?**

Reviewer 1: *Yes. This is consistent with US EPA's definition of a 15% decrease in FEV_1 as a cut-off for "adverse" effects (see, for example, US EPA. 2009. "Risk and Exposure Assessment to Support the Review of the SO_2 Primary National Ambient Air Quality Standard: Final Report. Second Draft." EPA-452/RP-09-007003, Office of Air Quality Planning and Standards (Research Triangle Park, NC). 896p., July).*

Reviewer 2: *This effect (greater than a 15% decrease in FEV_1) is generally relevant, with the caveat noted in the previous response that I believe that the observed effect is actually a measure of a response to chronic exposure. In the absence of good human data on acute toxicity, my initial thought was that the animal data (Graham et al. 1978) is a more appropriate basis for the acute ESL. However, that study is also not ideal, since it reflects a systemic effect, while the data indicate that the respiratory tract is the most sensitive target. Therefore, I agree with TCEQ that the best approach is to use the two studies together, but more discussion is needed of the caveats discussed here, and it may be appropriate to consider the studies co-critical.*

Reviewer 3: *The endpoint of bronchoconstriction among asthmatics is relevant to human risk assessment and could reasonably be considered adverse. The TCEQ has also appropriately considered the affected group to be a 'sensitive subgroup,' given that only asthmatics would be affected. However, it should also be noted that the affected workers in this study were sensitized under extremely harsh conditions, including high-level exposures to Ni, dust, combustion products, and other irritants. The implications of these coexposures need to be more fully discussed, especially as they contrast with low-level exposures to nickel alone.*

Reviewer 4: *The endpoint selected appears to be appropriate. Because it is a clinical indicator of respiratory deficiency, it appears to be a relevant and meaningful basis for setting regulatory limits intended to protect human health.*

Reviewer 5: *A decrease in FEV_1 (say, 15%-20%) is generally considered indicative of a significant asthmatic response. This effect is appropriate for setting ReV and ESL air concentrations.*

- **Is the endpoint relevant for human risk assessment?**

Reviewer 1: *Yes*

Reviewer 2: *Yes, in general, as a sensitive indicator of decreased pulmonary function.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *Yes, the endpoint is relevant for human risk assessment.*

- **The choice of point of departure**.

Reviewer 1: *The point of departure is appropriate and consistent with CalEPA and ATSDR.*

Reviewer 2: *This is appropriate, aside from the above caveat.*

Reviewer 3: *Seems appropriate based on Circla et al. data.*

Reviewer 4: *Given the choice of the critical study (Cirla et al., 1985), there is no choice with respect to the POD – there was only one exposure level in that study (an issue that was discussed above).*

Reviewer 5: *It is unfortunate that the study of Cirla et al. (1985) used only a single exposure level. This single level, 300 $\mu g/m^3$ (converted to 67 $\mu g/m^3$ divalent nickel equivalent), was considered a LOAEL by the TD. While technically this is correct, it is troublesome for several reasons. The response at that LOAEL was 50% which is far higher than the level of roughly 10% that is normally associated with a LOAEL in an experiment that also has a NOAEL. It seems quite possible that if a lower concentration had been included, a significant response might still have been observed. However, all things considered, it seems that using this purported LOAEL from Cirla et al. (1985) as a POD is preferable to choosing a*

POD from the alternative studies. Nevertheless, the discussion and application of uncertainty factors needs to address clearly how this LOAEL is appropriately reduced to derive the ReV and ESL.

- **The choice of dosimetric adjustments:**
  - **Was the most relevant, appropriate, and defensible dose metric selected?**

Reviewer 1: *Yes*

Reviewer 2: *Yes, with the caveat noted below for the RDDR.*

Reviewer 3: *Dosimetry adjustments appear appropriate based on underlying data. MMAD and geometric particle size distribution data were not available from Graham et al, so data from similar studies were substituted into the RDDR model. These simplifying assumptions should be satisfactory given that the mouse study by Graham et al is provided only as support to the human literature.*

Reviewer 4: *For effects at the site of contact (the lungs in this case) it is often the case that exposure concentration is used as the dose metric. I have no information that suggests that a different procedure should be used here, so I conclude that the most appropriate dose metric has been used in the nickel assessment.*

Although the standard exposure duration adjustment is to assume n=1 when no additional information is available, I question the statement (p. 13, line 12) that it is conservative to adjust from 30 minutes to 60 minutes by a simple ratio of 30/60. In fact, if n < 1, then that adjustment would not be conservative. I am not arguing that TCEQ adjust for duration in another way, I just think that (unless it can be strongly argued that n would not be less than 1) the document should just say that the policy decision is that the assumption of n=1 is the basis for the adjustment in the absence of additional information.

With respect to the animal dosimetric adjustment, the use of the RDDR methodology is a good choice. However, because of a lack of MMAD and σ_g data for the study in question (Graham et al., 1978), the document under-represents the uncertainties in the animal-based estimates. The document states (p. 13, lines 34-36) that this information is not particularly critical, but to the extent that the supporting animal-based estimates are used to corroborate the estimate from the primary study, there needs to be more characterization of that uncertainty. It might suffice to give a little bit more information about the other studies cited (from which MMAD and σ_g were obtained) with respect to how they compare to the Graham et al. study (similarity of exposure conditions, of the method of generating the test exposures, of the nickel compounds being tested). Moreover, a quick-and-dirty assessment of the impact of some other reasonable choices for MMAD and σ_g could show how much impact there was on the final adjusted animal POD estimate.

The fuller elaboration of the uncertainty associated with the animal-based POD is important because the POD derived from the primary, human study is supported by the claim that the latter is lower than the former (p. 16, line 4). As it stands now, it is not possible to know how seriously to take that statement.

Reviewer 5: *No adjustment of the LOAEL concentration was needed except for a duration adjustment.*

- **Were the appropriate default exposure duration adjustments conducted?**

Reviewer 1: *Yes*

Reviewer 2: *Yes.*

Reviewer 3: *See above***.**

Reviewer 4: *See above.*

Reviewer 5: *Yes, Haber's Rule was used to adjust the 30-minute exposure level of Cirla et al. (1985) to a 1-hour exposure level.*

- **For the supporting animal study (Graham et al. 1978), were the appropriate default dosimetry adjustments from animal-to-human exposure conducted? Specifically, were appropriate estimates (i.e. mass median aerodynamic diameter (MMAD) and geometric particle size distribution (σ_g)) for conducting the regional deposition dose ratio (RDDR) chosen when the supporting study did not report the required parameters?**

Reviewer 1: *TCEQ used an equation from US EPA (1994) to convert the adjusted POD to a human equivalent POD. This involves multiplying the adjusted POD by a dosimetric adjustment factor for the respiratory tract region, which is the regional deposited dose ratio ($RDDR_r$) for particles. TCEQ used a model to estimate $RDDR_r$ that used species-specific parameters (e.g., surface area, body weight, and ventilation rate) and particle parameters as inputs. TCEQ presents the model output, which includes all input values. It is not clear whether TCEQ chose the species-specific parameter inputs or whether they were built into the model. If the former, TCEQ should provide sources for transparency.*

Particle parameter values (MMAD and σ_g) were not provided in the key animal study, so TCEQ used input terms from other studies as surrogates. The values used (MMAD = 1.80 and σg = 1.60 μm) are on the lower side of the ranges

provided in NTP (1996a,b,c) and Oller et al. (2008). It is not clear how changes in particle input parameters affect the magnitude and direction of the model output. Model sensitivity to these parameters should be described for a greater understanding of their influence on the output.

Reviewer 2: *The approach used for estimating the MMAD and GSD appear reasonable. However, the authors state that the RDDR for the total respiratory tract was used because the critical effect is a systemic effect. If the effect is truly a systemic effect, with the appropriate internal dose measure being the amount of nickel that is absorbed from the respiratory tract and systemically available, then the RDDR should be calculated for the extrarespiratory region. (Relative deposition fractions are still used for the extrarespiratory RDDR, but the normalizing factor is body weight, rather than respiratory tract regional surface area.) The respiratory tract RDDR would be appropriate only if the effect on the immune system is believed to depend on the surface area dose to the lung.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *Yes, appropriate default dosimetry adjustments were conducted. Yes, appropriate estimates of MMAD and σ_g from other studies were chosen because Graham et al., (1978) did not provide those*

- **If the dosimetry adjustments were not acceptable, what would be a more appropriate dosimetric adjustment?**

Reviewer 1: *See previous question.*

Reviewer 2: *See previous question.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *NA*

- **The choice of uncertainty factors:**
  - **Have all of the appropriate uncertainty factors been considered and are the values assigned to the uncertainty factors clearly justified and defensible?**

Reviewer 1: *Yes. Although a LOAEL was used instead of a NOAEL, it was based on a 15% decrement in FEV_1. Table E-3 of TCEQ (2006) states that severe effects are*

>20% decrements in FEV_1. Because Cirla et al. (1985) don't actually state the change in FEV_1 for each individual, it cannot be known whether they were <20% or >20%, so it cannot be determined whether the effects were severe or not.

Reviewer 2: *Mostly. For the Acute ReV based on the human data, the documentation should state more explicitly that the study reports only that FEV_1 was decreased by >15%, but the actual magnitude of the decrease is not known. As written, my first impression was that the decrease was 15%, which would be considered a mild effect. Even with that uncertainty, however, one could make the argument that a reduced factor of 3 is adequate for LOAEL to NOAEL extrapolation, since the effect being measured appears to be a challenge response in a sensitized population, rather than an acute effect of nickel, as discussed above.*

While I agree that the acute data are very limited, I would suggest more careful consideration of the database for acute exposures before applying the database uncertainty factor. I agree that there are significant uncertainties regarding the magnitude of the threshold, but some bounding may be possible from the 12-day studies by NTP, which show more severe respiratory effects, but following exposure for durations much longer than that of interest for the acute ESL, and at concentrations not much higher than used in the Graham study.

Reviewer 3: *The authors of the TD addressed the appropriate areas of uncertainty and I agree with the UF of 1 for intrahuman variability, given that Circla et al. already deals with a sensitive subset of occupational asthmatics. However, the UF of 10 for extrapolation from LOAEL to NOAEL seems excessive. The TCEQ justified this value of 10 "because the severity of effects (mild or severe) could not be determined based on the FEV1 information presented" (p15). However, for diagnostic purposes, a 20% drop in FEV1 is considered to represent only mild obstruction, and a 30%-50% drop considered only moderate (not severe) obstruction (Rakel: Textbook of Family Medicine, 7th ed, 2007, Saunders – as cited on WebMD). These figures apply to chronic obstruction, but would seem to represent a conservative assumption for the acute situation as well. Furthermore, ATSDR lists acute/chronic lung inflammation (which is ostensibly more serious than subtle drop in FEV1) as a "less serious" effect in animal studies. Given this guidance, the 15% drop in FEV1 cited in the key study should be reasonably considered a less serious endpoint. Therefore, a UF of 3-6 would seem more appropriate.*

Reviewer 4: *The application of the uncertainty factors is the standard way of addressing extrapolation issues not directly quantifiable. All the standard factors are appropriately considered here.*

Because of the reliance on a human study that had only one exposure level (one that elicited a response in six of seven presumably sensitive individuals), I would consider increasing the UF_D value from 3 to 10. In this case, it is not only the limited number of studies but also the limitations within the critical study that is a

database deficiency. One additional point: it appears to be an overstatement to say that the critical study included a "significant number of occupational asthmatics" (p. 15, lines 14-15), when that number was only 7. Moreover, I wonder if occupational asthmatics are really a sensitive subpopulation – are they representative of (or maybe even less sensitive than) people who are asthmatic all the time and not only when associated with working conditions that entail nickel exposures? That is, have the individuals identified as being a sensitive subpopulation been selected only because they happen to be those who (among all the other exposed workers) were just that random set of workers who exhibited the asthmatic condition given their prior nickel exposures? In that case, does the Cirla et al. study merely indicate that those who have exhibited asthmatic symptoms because of nickel exposure will exhibit FEV effects if again challenged with nickel exposure? If that is the case, then I have to wonder whether the Cirla et al. results are telling us anything at all.

Reviewer 5: *All of the appropriate uncertainty factors have been considered. Some of the assigned values are clearly justified and defensible, but not all. The value $UF_H = 1$ is justified because the inclusion of seven asthmatics among the twelve subjects in the critical study should account for intrahuman variability including extreme differences in susceptibility due to supersensitive subpopulations. My main concern is whether the factor $UF_L = 10$ is large enough and whether its justification is clear enough. The value of 10 is the highest value normally applied for LOAEL-to-NOAEL extrapolation. However, LOAELs are not often associated with response rates as high as 50% and usually they are derived from studies with more than one exposure level. I think the TD expressed the situation well in the DSD in the statement from line 13, page 17 to line 4, page 18: "In other words, the potential magnitude of the difference between the NOAEL for respiratory effects and the single arbitrary concentration selected for use in the human study and later identified as the study LOAEL is unknown." I believe that this explanation ought to be included in Section 3.1.5.1.1 where the individual UFs are discussed and justified. Regarding the value $UF_D = 3$, I definitely agree that additional uncertainty needs to be reflected in the POD. Normally, I don't think that database uncertainty factors make sense for reducing PODs in a quantitative way. However, in this case an additional factor of 3 is definitely prudent. I suggest justifying $UF_D = 3$ as much in terms of the uncertainty in the LOAEL determined in the Cirla et al. study as in terms of limitations in the general acute study database.*

- **Would you make recommendations for a different approach to select uncertainty factors to calculate the acute ReV?**

Reviewer 1: *No.*

Reviewer 2: *I would suggest a bounding estimate using the 12-day study respiratory effects and appropriate dosimetry to evaluate the degree of concern for*

respiratory effects at lower concentrations, and the need for a database uncertainty factor.

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *In light of my above comments, I recommend that the discussions of the UF_L and UF_D be modified, to include acknowledging the considerable uncertainty in the LOAEL (and corresponding NOAEL, assuming one exists) so as to justify the total factor of 10 x 3 = 30. In other words, if the UFD is partially used to include uncertainty in the LOAEL, then I think the total factor of 30 can justify setting the acute ReV and ESL based on a study with only a single exposure level. I simply recommend more discussion of the uncertainty in the LOAEL used as a POD.*

Health-Based Chronic ReV and $^{chronic}ESL_{noncancer}$

Chapter 4 of the nickel DSD describes the approaches used to derive the health-based chronic ReV and chronic ESL for noncancer endpoints. The key decision points are listed below. For each decision, please consider the scientific defensibility of the decision and any additional approaches or analyses, or additional information that could improve that decision. Please indicate if there are other issues specific to developing chronic toxicity factors that have not been adequately addressed in the document.

- **The choice of the critical study (NTP 1996):**
  - **Was the animal study selected for the non-cancer estimates the most appropriate study? Was the form of nickel selected (nickel sulfate) for deriving the ReV/ESL appropriate given the purpose of these values?**

Reviewer 1: *The most appropriate animal study was selected regarding soluble nickel. Although soluble, oxidic, sulfidic, and metallic nickel have all been associated with pulmonary fibrosis and chronic inflammation in the lung (NTP, 1996a,b,c; Oller et al., 2008), the toxicity of each varies, with nickel sulfate being the most toxic. Each of these studies can and should be used to calculate a separate ReV/ESL for each form of nickel.*

Reviewer 2: *Yes, nickel sulfate is the most toxic form of nickel relevant to the ReV/ESL for noncancer endpoints.*

Reviewer 3: *The authors of the TD chose this large NTP bioassay as the key study because this was the study chosen by ATSDR and Haber et al. However, the ATSDR evaluated a large number of studies before making their decision. The choice of NTP (1996) appears reasonable, but it would seem appropriate to at*

least summarize the breadth of this literature and provide objective rationale for why NTP (1996) is most appropriate, rather than just relying on the judgment of ATSDR.

The choice of nickel sulfate also seems reasonable, given that this is among the most toxic forms. However, one needs to know how the reference values derived using the most toxic form will be applied to less toxic forms, such as metallic or insoluble nickel. This ultimately brings one back to the question of how 'Ni equivalents' were calculated and how they will be applied. For example, will metallic nickel be comparable to nickel sulfate on a weight to weight basis? This needs to be more fully elucidated.

Reviewer 4: *The NTP study was a well-conducted chronic bioassay entirely suitable for deriving chronic ESL values. The decision to base the ESL values on results from the testing of nickel sulfate is appropriate as a health protective policy decision.*

Reviewer 5: *Yes, the animal study selected for the non-cancer estimates was the most appropriate study. Two previous studies, Haber et al. (2000) and ATSDR (2005), selected the NTP (1996) animal study after reviewing human and animal data. The human database is very limited for evaluating the respiratory effects of soluble nickel because of there being only a small number of studies and because of uncertainties or deficiencies in the existing studies. The form of nickel selected (nickel sulfate) was appropriate for deriving the ReV and ESL values because soluble forms of nickel are at least as toxic as insoluble forms, so that nickel sulfate can serve as a conservative surrogate for all inorganic forms of nickel.*

- **The choice of critical effect:**
  - **Was the most appropriate critical effect selected (lung fibrosis and chronic inflammation)? If not, what would be a more appropriate critical effect?**

Reviewer 1: *Yes*

Reviewer 2: *Yes.*

Reviewer 3: *The authors of the TD have based their chronic assessment on the critical effect identified by ATSDR and Haber et al. This endpoint seems reasonable, given that airborne Ni is known to have effects on the respiratory system in people, but it would seem appropriate to at least summarize the breadth of the available effects and provide objective rationale for why lung inflammation is most appropriate, rather than just relying on the judgment of ATSDR and Haber et al. The TCEQ should also note that epidemiologic cohort studies have not typically identified significant associations between relatively high occupational exposures to nickel and non-malignant respiratory disease, which detracts from the finding in rats.*

Reviewer 4: *It is difficult to know if the most appropriate critical effect was selected because no data were presented that compared the rates or response levels for various possible alternative choices. This is a major deficiency of the document.*

I know of no reason why that selected endpoint should not be relevant for human risk assessment.

Reviewer 5: *Lung fibrosis and chronic inflammation in rats were the most appropriate critical effects because the lung is the most sensitive target of nickel toxicity.*

- **Is the endpoint relevant for human risk assessment?**

Reviewer 1: *Yes. The lung is the target organ in humans.*

Reviewer 2: *Yes.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *Yes, the endpoint is relevant for human risk assessment*

- **The choice of dosimetric adjustments:**
  - **Was the most relevant, appropriate, and defensible dose metric selected?**

Reviewer 1: *Yes.*

Reviewer 2: *I agree with the use of MPPD to for the dosimetric adjustment. I do recommend a clarification to Section 4.1.3, which says that the dose metric was exposure concentration. It would be more appropriate to state that the dose metric was regional deposited dose, calculated using MPPD (using a default categorical model). (Other metrics are also possible with MPPD.) This is an appropriate internal dose metric for the subject endpoints.*

Reviewer 3: *The dose metric seemed appropriate. I am not familiar with the specifics of the MPPD and RDDR models, but they seemed appropriate given their use in other toxicological profiles (eg, ATSDR 2005). The authors of the NTP study provided necessary inputs to these models, so there was no need to read across from other studies.*

Reviewer 4: *The MPPD model is a good basis for determining the HEC value. The choice of the deposition in the pulmonary region as the basis for the metric appears to be reasonable given the endpoints selected for analysis.*

Reviewer 5: *Exposure concentration from the key study was used as the default dose metric because the MOA for the critical effect is not fully known. This was relevant, appropriate and defensible.*

- **Were the appropriate default dosimetry adjustments from animal-to-human exposure conducted? Specifically, was the Multiple Pass Particle Dosimetry (MPPD) Model used appropriately and is the (RDDR) appropriate? Were the parameters used scientifically defensible?**

Reviewer 1: *The citation for the MPPD model (i.e., CIIT, 2004) is not included in the reference section. The full cite and a short description of the model should be added. The species-specific parameters used to calculate $RDDR_r$ should be referenced.*

Reviewer 2: *Mostly correct adjustments were made. The one caveat is that there was an inconsistency in the calculation of the RDDR from the deposition fraction. For the human breathing rate, the DSD used the EPA default of 13,800 mL/min. However, the deposition fraction was calculated with MPPD using the default scenario of light activity, which results in a different minute volume. The minute volume used by MPPD can be calculated as follows: (1) the output provides human tidal volume (volume/breath) of 625 mL and a breathing frequency of 12/min. (2) the product of tidal volume and breathing frequency is the volume/minute = 7500 mL/min. This human minute volume would then be used to calculate the RDDR based on the MPPD default values. Similarly, the rat minute volume used in the deposition calculation is the product of 2.1 mL x 102/min = 214.2 mL/min. Alternatively, the authors could calculate the deposition with MPPD using the human tidal volume and breathing frequency corresponding to the default EPA parameters. The choice of which human minute volume to use is a science policy decision, but internal consistency in the calculation is needed.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *Yes, the appropriate default dosimetry adjustments from animal-to-human exposure were conducted, including standard duration adjustments. The MPPD Model was used appropriately to convert the animal POD to a human-equivalent POD, and the RDDR of the pulmonary region was selected as the appropriate output because of the critical effect(s) of lung fibrosis and chronic inflammation. Parameters from the NTP study were used along with default parameters from the MPPD program, which is scientifically defensible.*

- **If the dosimetry adjustments were not acceptable, what would be a more appropriate dosimetric adjustment?**

Reviewer 1: *NA*

Reviewer 2: *Using minute volume based on the MPPD default minute volume, I calculate an RDDR of 1.97 (unrounded). Of course, a different value would be calculated using the EPA defaults.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *NA*

- **The choice of point of departure**.

Reviewer 1: *The DSD states that data from the NTP (1996a) soluble nickel study were not amenable to standard benchmark concentration modeling. More information should be provided regarding why this is the case. If it is true that BMC modeling is inappropriate, then the point of departure chosen in appropriate.*

Reviewer 2: *Mostly, with the caveat regarding the dosimetric adjustment noted above. In addition, the authors should explain further why the data were considered not amenable to standard BMC modeling. I presume that the issue is the difficulty of obtaining adequate fit, even after the high concentration is dropped, but the authors should explain what is "not amenable." "Not amenable" could also mean that the data are not available in an appropriate form for input into a model, which is not the case here.*

Reviewer 3: *Seems reasonable.*

Reviewer 4: *It is not clear why the assessment concluded that benchmark dose analysis could not be done for the endpoint(s) selected (p. 21, lines 10-11). Again, without the data being reproduced in the document, there is no direct way to corroborate that claim. Given the many known issues with NOAELs and LOAELs as the basis for regulatory decision making, I believe that there needs to be much stronger support for using a NOAEL over a BMDL. Thus, at this time, I have serious reservations about the choice of the POD.*

Reviewer 5: *The DSD states that because the MOA for the adverse respiratory effects of nickel has not been fully elucidated, the default threshold, nonlinear dose-response is used. However, the DSD also states that the data were not amenable to standard benchmark concentration modeling (page 21, lines 10-11), although no further discussion is provided. That being the case, then the choice of the*

NOAEL from the chronic rat study as the POD is appropriate, given that rats were more susceptible to the effects of nickel than mice. I think the DSD should indicate why the dose-response data were not amenable to benchmark concentration modeling

- **The choice of uncertainty factors.**
  - **Have all of the appropriate uncertainty factors been considered and are the values assigned to the uncertainty factors clearly justified and defensible?**

Reviewer 1: *Yes*

Reviewer 2: *Yes.*

Reviewer 3: *The rationale given for the various UFs seem reasonable, as does the combined UF of 30*

Reviewer 4: *The uncertainty factor selection appears to be entirely appropriate.*

Reviewer 5: *The DSD correctly states that a UF_L is not applicable, because the animal POD is a NOAEL. A factor of $UF_A = 3$ is clearly justified and defensible to account for toxicokinetic differences between animals and humans not accounted for by the MPPD model. A value of $UF_H = 10$ is clearly justified and defensible to account for intrahuman variability, including sensitive subpopulations. A $UF_D = 1$ was clearly and justifiably applied because of the strong database.*

  - **Would you make recommendations for a different approach to select uncertainty factors to calculate the chronic ReV?**

Reviewer 1: *No*

Reviewer 2: *The approach is appropriate.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *No*

Welfare-Based Acute and Chronic ESLs

The TD did not find any data to allow the derivation of welfare-based acute or chronic ESLs. Please indicate if there are other issues specific to developing welfare-based ESLs that have not been adequately addressed in the document.

Reviewer 1: *NA*

Reviewer 2: *I am not aware of any.*

Reviewer 3: *NA*

Reviewer 4: *I know of no issues related to the derivation of a welfare-based ESL*

Reviewer 5: *I am not aware of any other issues specific to developing welfare-based ESLs that have not been adequately addressed in the document.*

Cancer Weight of Evidence and Unit Risk Factor (URF)

The nickel DSD describes the approaches used to evaluate carcinogenicity and derive the URF and chronic ESL for cancer. Please review the key decisions made by TCEQ in deriving these values. For each decision, please comment on the consistency of the decision with TCEQ's ESL guidelines, the scientific appropriateness of the decision, and any additional approaches or additional information that would improve that decision. The key decisions and some specific issues to consider are listed below. Please discuss other issues specific to developing URFs for carcinogenic effects that have not been adequately addressed in the document.

Reviewer 1: *NA*

Reviewer 2: *A critical consideration in the development of the relevant carcinogenicity assessment for the DSD is the nature of the exposure (i.e., form[s] of nickel) in Texas. Section 4.2.4 does a nice job of presenting information relevant to this consideration and comparing emissions in Texas with the nature of exposure in the epidemiology data. This is a very important comparison, but some additional information is needed.*

Much of Section 4.2.4 focuses on emissions in Texas. While emissions information may be the data that are most readily available, the metric of interest is of course exposure levels in ambient air (and in air near emitters). There may be little or no information on actual exposures, but this should be addressed explicitly. In particular, the text indicates that emissions are primarily metallic nickel, along with soluble nickel and nickel oxides, but then later states that the major nickel species in ambient air is soluble nickel. Please address/explain this

apparent discrepancy. Is the latter statement based on measurements, or inferred from emissions data? Might it be that the metallic nickel emissions tend to be larger particles that rapidly precipitate out of the air, while the soluble nickel species are smaller and have longer residence times in the air? Is it known to what degree oxidic nickel contributes to nickel exposure in ambient air?

Recognizing that some data may not be available, it is very important that TCEQ provide what is known and what is not known with regard to nickel speciation in ambient air. Is there any air monitoring data that includes speciation? If not, and industry emissions provide the best estimate of the composition of the nickel, that should be stated. Can any estimate be made based on the emissions data on the percent of nickel in ambient air that is present as nickel subsulfide?

As noted by the authors, a critical issue in determining whether there are appropriate epidemiology data to use to estimate the risk of exposure to nickel in Texas air is the relative proportion (not absolute amount) of nickel subsulfide in the Texas air and under the worker exposure conditions in the epidemiology studies. This helps in the determination of whether the two mixtures are sufficiently similar that the epidemiology data can be used to estimate the risk from exposure to total nickel in the ambient air. (The proportion is the key metric, rather than absolute amount, because risk is expressed per amount of total nickel, and the proportion is assumed constant as total dose decreases.) I focus on nickel subsulfide levels in this discussion, since the authors have already addressed metallic nickel and nickel sulfate levels (aside from the issues noted in the previous paragraph). Thus, it would be useful to have additional data on the best estimate for the proportion present as nickel subsulfide, the most potent form of inorganic nickel for carcinogenic potential. The initial text gives the impression that there is minimal subsulfide in the air. Later on the text states that one "cannot exclude the possibility of some nickel subsulfide," which also sounds like minimal subsulfide. Chapter 2 mentions metallic nickel, nickel sulfate, and nickel oxide in Texas air, with no indication of anything above de minimis nickel subsulfide levels. Together, these characterizations of Texas air would suggest that the appropriate mixture for worker exposures in the epidemiology studies used for risk calculation would also include only de minimis levels of nickel subsulfide.

This has significant implications with regard to the appropriateness of using total nickel as the exposure measure from the epidemiology studies. If nickel subsulfide constitutes a substantially higher percentage of the total nickel in the epidemiology studies than in Texas air, this would result in a substantial over-estimate of cancer risk to the population of Texas. Marked discrepancies would suggest either (1) the use of other epidemiology data, if available; (2) extrapolation from animal data; or (3) use of the approach as in the draft as a very conservative estimate, with clear language on the highly conservative nature of the estimate. Alternatively, if it is consistent with the ESL guidelines, TCEQ might consider using some adjustment to the current risk estimate to account for

differences in the composition of the mixtures. Conversely, if the risk estimate is based on total nickel in an epidemiology study, the closer the composition of the nickel species in the Texas air and the worker exposure in the epidemiology studies, the greater the confidence in using the risk estimates (because the mixtures are more similar).

I recognize that this evaluation is by no means an easy data set. I commend TCEQ for the attempt to address the mixtures issue, which is needed in order to consider the contribution of soluble nickel to the total risk. There is little or no precedent on how to quantitatively address mixtures of this nature, with some components being likely promoters, and others more likely to act as complete carcinogens.

Reviewer 3: *NA*

Reviewer 4: *NA*

Reviewer 5: *NA*

- **Was the proper weight of evidence (WOE) classification using the new USEPA carcinogenic guidelines given to nickel compounds? If not, what WOE classification should be given to nickel compounds, specifically metallic nickel?**

Reviewer 1: *The DSD states: "According to guidance in the new cancer guidelines (USEPA 2005a), the TD considers nickel compounds as a group to be 'Carcinogenic to Humans' via inhalation." If the TD decides to consider nickel classes as a group, this conclusion would be appropriate. Because data clearly show differences in the carcinogenic potential among sulfidic, oxidic, soluble, and metallic nickel, a WOE analysis should consider each form separately, and should consider animal and mode-of-action (MOA) data. Currently, the WOE analysis relies almost solely on human data, and largely ignores animal and MOA data, some of which have become available since the reviews referenced in the DSD (i.e., ATSDR, 2005; IARC, 1990; ICNCM, 1990). There is little mention of MOA data, and the discussion of animal data is limited to the following (pg 29):*

> *"See Section 3.2.1.7 of ATSDR (2005) for a discussion of the inhalation animal studies which have examined the potential of various forms of nickel (i.e., nickel subsulfide, nickel oxide, and nickel sulfate) to increase lung tumors. In general, only chronic inhalation exposure to nickel subsulfide and nickel oxide resulted in lung tumors (adenocarcinomas, squamous cell carcinomas, and fibrosarcoma) in Fisher 344 rats, but no significant alteration in tumor incidences were observed in mice. Considering available study data for both inhalation and other exposure routes (e.g.,*

intraperitoneal, intrapleural, intramuscular, subcutaneous), the ICNCM (1990) indicates there is: sufficient data in experimental animals for the carcinogenicity of metallic nickel, nickel monoxides, nickel hydroxides, and crystalline nickel sulfides; and limited animal evidence for the carcinogenicity of nickel alloys, nickelocene, nickel carbonyl, nickel salts (e.g., nickel sulfate), nickel arsenides, nickel antimonide, nickel selenides, and nickel telluride; and inadequate animal carcinogenic evidence for nickel trioxide, amorphous nickel sulfide, and nickel titanate."

The ICNCM (1990) conclusions are outdated, as several robust studies have been conducted in the 19 years since it was published, many of which do not support the findings of earlier studies. Similarly, the ATSDR (2005) review precedes the Oller et al. (2008) study of metallic nickel, which clearly does not support an association between metallic nickel and cancer risk. To rely so heavily on the ICNCM (1990) report is not conducting a complete WOE analysis. In addition, the DSD does not appear to consider the conclusion of the ATSDR (2005) analysis, based on the NTP (1996a,b,c) carcinogenicity studies, which demonstrates positive, equivocal, and no evidence of carcinogenicity for sulfidic, oxidic, and soluble nickel, respectively. There is also little discussion of the MOA and how it may differ for different forms of nickel.

Were the TD analysis to consider the full WOE, including all of the animal and MOA data, the most appropriate classifications would be "not likely to be carcinogenic to humans" for soluble and metallic nickel, "likely to be carcinogenic to humans" for oxidic nickel, and "carcinogenic to humans" for sulfidic nickel.

Incidentally, IARC (1990) did not conclude that there was "sufficient evidence in humans that nickel sulfate (soluble nickel) is carcinogenic" as stated on page 30 of the DSD. Rather, IARC concluded that "there is sufficient evidence in humans for the carcinogenicity of nickel sulfate, and the combinations of nickel sulfides and oxides encountered in the nickel refining industry." There is not sufficient evidence for nickel sulfate alone, but only in the presence of other nickel compounds. IARC defines this group of nickel forms (soluble and insoluble forms) as carcinogenic to humans (Group 1), and metallic nickel as possibly carcinogenic to humans (Group 2B). IARC made no conclusions regarding soluble nickel alone.

Reviewer 2: *The authors appear to have missed a critical study in the evaluation of the WOE of carcinogenicity of metallic nickel. Oller et al. (2008) conducted a recent inhalation study in rats with metallic nickel powder and reported toxicity and lung lesions, but no tumors. Inhalation studies with other nickel compounds show respiratory tract tumors in rats (or not at all) but not in mice; therefore lack of a mouse study with nickel metal powder is not a critical data gap, This study should be considered in the weight of evidence. Based on the animal data alone,*

this study would suggest that metallic nickel is not likely to be carcinogenic to humans. This conclusion is supported by the epidemiology data on exposures to metallic alone (Oak Ridge Gaseous Diffusion Plant - Godbold et al., 1979; nickel alloy plant workers - Cox et al., 1981, Sorahan 2004; and workers in a French factory producing stainless and alloyed steel - Moulin et al., 2000), which have not observed increased respiratory cancer risks.

The WOE classification for the nickel mixture in Texas air needs to take into account the composition of the nickel species in that air. Based on the description of the mixture as containing minimal levels of nickel subsulfide, compared with the relatively higher proportion of nickel subsulfide in the Enterline and Marsh (1982) refinery workers, and the likely under-estimation of nickel subsulfide in the Kristiansand cohort (see comments below), I do not believe "carcinogenic to humans" is appropriate for Texas air. I would lean towards "likely to be carcinogenic to humans" for Texas air, based on the animal data for oxidic nickel and the potential for soluble nickel to enhance the carcinogenicity of other forms of nickel, although a biphasic approach ("suggestive evidence of carcinogenicity" under certain conditions and "likely evidence under others"). This analysis takes into account the analysis of ICNCM (1990) of "low" vs. "high" levels of metallic, sulfidic, oxidic, and soluble nickel in various combinations for the Clydach cohort, but that even "low" subsulfide in that cohort was relatively high by modern standards (and potentially higher than "low" soluble nickel in the same study). Additional information on the composition of Texas air would be useful for this determination, as noted above.

Reviewer 3: *The TCEQ did not present any independent weight of evidence or classification assessment to individual nickel compounds. Instead, they quote material from the ATSDR document and present the cancer classifications derived by USEPA, ATSDR, IARC, and DHHS. The TCEQ then proceeds to treat all nickel species as comparable carcinogens, based on the fact that there is some uncertainty as to the most important species. This is somewhat problematic given that all the agencies cited in the TD have given metallic nickel a markedly lower classification (eg, possible, not suspected, no evidence) for carcinogenicity, suggesting that it is inappropriate to classify metallic Ni as carcinogenic with this same potency. This issue is further complicated by the fact that most Ni emissions in TX consist of metallic Ni. A classification for metallic nickel based on the weight of evidence should more reasonably be 'suspect' or 'unlikely' human carcinogen.*

Reviewer 4: *Despite the uncertainties associated with the particular forms of nickel that may or may not be carcinogenic in humans, or may differ in their potency with respect to human cancer induction, the presentation of the various considerations appears to be consistent with the TD's conclusion that nickel compounds as a group should be considered carcinogenic to humans. The presentation of any animal data that show different cancer results by nickel compound (or, the*

existence of cancer risks for various forms of nickel) would be an appropriate addition here to support the conclusion.

Reviewer 5: *The proper WOE classification was given to nickel compounds. The TD's consideration of nickel compounds as a group to be "Carcinogenic to Humans" via inhalation is justified. Because the carcinogenic MOA for nickel is yet to be fully elucidated, the TD used linear low-dose extrapolation to calculate URFs as a conservative default assumption.*

- **The cancer assessment relied upon human epidemiological studies. There are also animal studies; were the animal data used appropriately to support the weight of evidence conclusions?**

Reviewer 1: *The animal data were not given enough weight in the analysis. Although the epidemiology studies were reviewed at great length, there was very little information provided on the animal data. As discussed in the DSD, because of co-exposures to several forms of nickel in the epidemiology studies, one cannot determine which form is associated with lung or nasal cancer risk. The animal studies provide clear information on which forms are carcinogenic (and this is supported by MOA studies and studies of lung accumulation and clearance). Across studies, insoluble nickel compounds have been shown to be carcinogenic to different degrees while soluble and metallic nickel have not. In the case of metallic nickel, no carcinogenicity in animals or humans has been observed, possibly because of factors leading to an overall low nuclear nickel bioavailability, as described in Oller et al. (2008). The most robust animal study of soluble nickel, conducted by NTP (1996c) using an inhalation exposure route, did not report an association with respiratory cancer risk, although risks in this animal model were noted with nickel oxide and nickel subsulfide (NTP, 1996a,b). These models are more appropriate than the human data for calculating cancer risks, and risks should be calculated separately for sulfidic, oxidic, soluble, and metallic nickel.*

Reviewer 2: *I would recommend that additional weight be given to the animal studies, particularly the impact of the negative metallic nickel study, and the substantially higher potency of nickel subsulfide compared with other forms.*

In addition, in referencing previous assessments, it is important to note key studies that have been published since the conduct of certain assessments (and to recognize that the assessments were not able to take newer data into account). In particular, the NTP studies of soluble nickel, nickel subsulfide and nickel oxide were conducted in part to address the speciation issues raised by ICNCM (1990).

Reviewer 3: *Results from animal studies are only mentioned in passing within a paragraph that refers to the ATSDR and ICNCM documents (p29). The TCEQ does not mention that the NTP (1996) study, which was relied upon heavily as the*

critical study for chronic, noncancer assessment, reported no significant lung cancer increase in rats exposed to nickel sulfate. This is inconsistent with those human studies that reported increased risk from exposure to soluble nickel, and adds to the uncertainty regarding carcinogenic nickel species.

Reviewer 4: *As stated at the end of the previous comment, I believe the animal data could be used to support a conclusion about weight of evidence. But I saw no such presentation – I did not see discussion of any animal results whatsoever, in relation to any cancer endpoint or with respect to the weight of evidence as a whole.*

Reviewer 5: *I saw only one reference to an animal carcinogenesis study (NTP, 2002) which was cited on page 30 but was not listed in the reference list. The NTP study used to establish the chronic ReV and ESL for noncarcinogenic effects was not mentioned in the cancer discussion.*

- **Is the epidemiological evidence in Grimsrud et al. (2003) and Enterline and Marsh (1982) properly used in the characterization of chronic cancer risks? Is use of these two studies for calculating URFs justified?**

Reviewer 1: *The DSD states: "Available information indicates that Texas nickel emissions would predominantly be metallic (e.g., railroad equipment, steel foundries, aircraft engines, metal forging, oil/gas field machinery, plate work), along with soluble nickel (e.g., electric utilities) and nickel oxides (e.g., electric utilities, steel foundries and works, aircraft engines), and would therefore be low in sulfidic nickel" (pg 32). The TD states that they chose studies with low sulfidic nickel exposures because it felt this would be more comparable to exposures in Texas. In reality, exposures in the Grimsrud et al. (2003) and Enterline and Marsh (1982) study are also quite different than those in Texas, as individuals in these studies are exposed to several forms of nickel while individuals in most Texas industries are not (and even if they are, the ratio of forms of nickel differs from those in these two studies). For example, Texas industries primarily emitting metallic nickel should not be considered to be exposed to a carcinogenic form of nickel, based on strong evidence in animals and in vitro. The Grimsrud et al. (2003) and Enterline and Marsh (1982) studies could be used as support for a URF, but the URF should be specific for each form of nickel and based primarily on animal data.*

In addition, while the DSD acknowledges the overall lack of statistical significance in the Enterline and Marsh (1982) study, this does not seem to play a role in the derivation of the URF. The DSD should not calculate a risk value based on a study for which there were very few statistically significant risks.

Reviewer 2: *As noted above, a critical consideration in the determination of the studies to use as the basis for the quantitative assessment was that the epidemiology*

studies involved subjects exposed to the same (primary) forms of nickel as those relevant to population exposure in Texas. According to the DSD, the major ambient form of nickel is nickel sulfate, with also some (not clear from the relative proportion) exposure to metallic nickel. Sulfidic nickel exposure is low. On this basis, the authors appropriately looked for studies involving predominantly exposure to soluble nickel and metallic nickel, with low exposure to sulfidic nickel, and chose the two specified cohorts.

A critical issue regarding the appropriateness of the cohorts is that exposure in those studies should be predominantly to soluble nickel, not sulfidic nickel. The Goodman et al. (2009) work can help us better understand the specific nickel species in these key studies. According to Goodman and colleagues, due to the analytical methods used, the material that was measured and described as soluble nickel at the Kristiansand refinery (e.g., Grimsrud et al., 2003) may have actually been predominantly nickel subsulfide. There are two important implications of this statement: (1) There are a number of observations from this facility that are inconsistent with the animal data. For example, the NTP (1996b) study concluded that there was "clear evidence" of nickel subsulfide carcinogenicity in rats, but NTP (1996a) found "no evidence" of nickel sulfate carcinogenicity in rats. In contrast, Grimsrud and colleagues studying the Kristiansand cohort have consistently interpreted their data as showing that soluble nickel is much more potent than sulfidic nickel. If the "soluble nickel" measurements at that facility reflected a systematic misclassification of sulfidic nickel, this may explain the presumed association of cancer risk with soluble nickel. In addition, the dose-response developed by the Kristiansand group (Grimsrud et al. 2002) exhibits a very unusual shape, with a very sharp increase at low doses, followed by rapid leveling of the curve. This unusual shape may also reflect misclassification of the exposure. (2) The DSD appropriately emphasized the importance of choosing epidemiological studies for the dose-response assessment that matches as closely as possible the nature of the exposure experienced by the Texas population. If the exposure at Kristiansand was actually primarily to nickel subsulfide, this would raise questions about the appropriateness of using that cohort to describe the risk to total nickel in Texas air, since the observed dose-response would be driven primarily by the sulfidic nickel, which is described as being present at low levels in Texas air.

Based on these considerations, I conclude that there are too many uncertainties regarding the actual nature of the exposure at Kristiansand to use that study as a basis for the quantitative assessment. Using total nickel exposure (instead of speciated nickel) does not resolve the issue, in light of the animal data showing high potency for nickel subsulfide, and the likely/potential differences in the proportion of nickel subsulfide at Kristiansand and in Texas.

Only total nickel exposures (not speciated values) are available for the Enterline and Marsh (1982) study, making it harder to compare that mixture with the composition of Texas air. The URF calculated in the DSD was based on all

workers, including both refinery workers exposed to nickel subsulfide, and workers in positions with "no" exposure to nickel subsulfide. The rationale provided is that some nickel subsulfide exposure in Texas air "cannot be excluded." While this may be the case, it is also possible that some exposure to nickel subsulfide in the non-refinery workers cannot be excluded (depending on the physical plant setup and potential for carryover between operations, as well as the potential for a worker to have been at the refinery for <1 year). The more relevant question is which group and exposure is better representative of the composition of Texas air. If a policy decision is made to use the "all worker" data as a health-protective approach due to inadequate speciation data for Texas air, this needs to be clearly stated, along with a characterization of the uncertainties and likely over-protectiveness of this approach.

Reviewer 3: *The TCEQ provide no substantive or critical review of the epidemiologic literature, as was done in the ICNCM document. The TD contains no clear inclusion criteria, except to say that "not all these studies are adequate to define the dose–response relationship" (p33) and that the 1986 assessment by US EPA relied on 4 cohort studies (Copper Cliff, Clydach, Huntington, and Kristiansand) (p33). This is not an objective or transparent approach, and detracts from reader confidence as to whether the epidemiological evidence has been fully or properly evaluated. Most importantly, the rationale for selection of the Grimsrud et al. (2003) and Enterline and Marsh (1982) studies appears to be flawed.*

The authors of the TD indicate that they chose these two studies because (1) studies other than refinery studies have not shown carcinogenic effects from Ni exposure, (2) refineries are usually associated with high levels of sulfidic nickel, and (3) TX has mostly low sulfidic emissions. Therefore, they excluded two studies (Copper Cliff and Clydach) that had higher levels of sulfidic nickel, leaving Grimsrud et al. (which is an update of Magnus) and Enterline and Marsh. In my opinion, this rationale is flawed on several levels:

1) It sounds like circular reasoning that uses loosely related arguments to reach a predetermined goal,

2) It runs directly contrary to their previous assertion that all species of Ni are potentially carcinogenic, so that the species of nickel is unimportant, and

3) It ignores issues of data quality, such as sample size and lack of bias, and selects studies based solely on generalizability of exposure.

In general, one should not select epidemiologic studies based on issues of generalizability to a specific population. Rather, one needs to consider all the epidemiologic evidence, judge the quality of each study, and use the totality of the evidence to estimate the most likely association between an exposure and outcome. Generalizability is a secondary concern that should be addressed after an accurate estimate of association has been determined using the best available evidence.

The best approach to study selection would have probably been to include all studies with suitable exposure estimates (based on objective criteria), summarize the strengths and limitations of each study, and calculate a meta-summary or meta-regression using appropriate weighting factors (eg, sample size and study quality). A meta-summary ends up combining relatively heterogeneous SMR/SIR that range from 1.0 to >3, but would still generate a summary estimate that is more consistent with the TCEQ assumption that all species of Ni are potentially carcinogenic.

Reviewer 4: *I have considered these two questions together, because the answer to the second one has some bearing on the answer to the first one. The fundamental reason reported in the nickel document for excluding the Copper Cliff and Clydach cohorts relates to the dissimilarity of the nickel exposures to the emissions expected to be seen in Texas. The document states that the Texas emissions would be expected to be low in sulfidic compounds. It also states, somewhat confusingly, that those emissions would be predominantly metallic (p. 32, line 14), but then says (p. 32, line 24) that the major nickel species in ambient air would be in a soluble form.*

With respect to the exclusion of the Copper Cliffs and Clydach cohorts, I can understand the exclusion of Copper Cliffs because it had sulfidic species in very high concentrations. But the Clydach cohort after 1930 only had ">1 mg Ni/m^3" (Table 6), which was not the highest in that cohort (oxidic was ">5") and which was not that much greater than the Kristiansand cohort sulfidic concentrations ("> 0.5"). Moreover, the ratio of oxidic/sulfidic in the Clydach (after 1930) was something around 5 (judging just based on the typical exposure ranges shown in Table 6) whereas for Kristiansand it was only 4, so that latter actually had more sulfidic relative to oxidic. Furthermore, the Huntington cohort exposures apparently had no soluble nickel forms ("soluble" is not listed for that cohort in Table 6) despite that being the form most expected in Texas ambient air. A similar reservation could be made about the Kristiansand cohort where soluble and sulfidic were in the same range of concentrations (">0.5") unlike what was reported to be expected for Texas emissions.

I would conclude from Table 6 and the discussion about Texas conditions that the only clear exclusions would be Clydach before 1930 and Copper Cliff, where sulfidic concentrations were much greater than soluble and metallic concentrations.

Therefore, I would conclude that while the use of the data from Enterline and March and Grimsrud has been appropriate, and that they can be used to derive URFs, there may be one additional cohort (or subcohort of those employed at Clydach after 1930) that could also have been used.

Reviewer 5: *Yes, the epidemiological evidence in the two selected studies is properly used in the characterization of chronic cancer risks. Yes, the TD explains that the*

lower levels of sulfidic nickel to which workers were exposed in the Grimsrud et al. (2003) and Enterline and Marsh (1982) studies reflect more closely the exposure profiles more relevant to nickel emissions in Texas.

- **Were the reasons for not using the following epidemiological studies to develop URFs clearly described and justified: Copper Cliff, Ontario (Chovil et al. 1981) and Clydach, Wales (Peto et al. 1984)?**

Reviewer 1: *The TD's explanation for excluding these studies was clearly described, but its reasoning is not adequate. The reason given is that the nickel forms in these studies are dissimilar to those forms released in Texas (there are low amounts of sulfidic nickel released in Texas). Yet the epidemiology studies upon which the TD relied are in sulfidic ore refineries, and these exposures also vastly differ from those in Texas with regard to the ratio of the different forms of nickel and the exposure levels. In addition, the DSD currently relies on total nickel for the ReV/ESL calculations, and the forms of nickel that make up the total vary between the two studies on which the TD relies and among all the nickel releasers in Texas. If the TD is going to apply the results of calculations of total nickel to industries with just one form of nickel, then it is inconsistent to exclude studies based on the forms of nickels to which subjects are exposed. It is more appropriate to use animal data as the primary data for the ReV/ESL calculations and the epidemiological data as supportive.*

Reviewer 2: *Yes. However, it is not clear whether the authors considered using the data from Harjavalta, Finland (Karjalainen et al., 1992; Anttila et al., 1998). Those data would need to be evaluated for nature of the exposure, appropriateness of measurement technique, and availability of adequate dose-response information.*

Reviewer 3: *See above.*

Reviewer 4: *See above.*

Reviewer 5: *Yes, it was clearly explained that workers in the two studies in question were exposed to relatively high levels of sulfidic nickel relative to other species, which are not reflective of the lower levels of sulfidic nickel expected to be emitted by nickel sources in Texas.*

- **Were the statistical and modeling approaches used for calculating URFs appropriate?**

Reviewer 1: *Yes, given than they were based on human data.*

Reviewer 2: *This is beyond my area of expertise.*

Reviewer 3: *In my opinion, the modeling approaches used by the TCEQ were too complicated given the uncertainty and variability inherent in the underlying data. The ICNCM document (Doll et al. 1990) identified 10 cohorts with lung-cancer SMR ranging from approximately 1.0 to 4, suggesting substantial between-study variability (ie, heterogeneity). Results vary considerably even within cohorts. For example, the Grimsrud et al. cohort reports many 95% confidence intervals that range from <2 to 7-9. It should also be remembered that there was substantial potential for confounding that adds to the uncertainty of results reported in these occupational cohorts.*

The refinery process encompasses more than Ni exposure, with workers coexposed to other metals, combustion products, asbestos, acid mists, and irritant dusts. Respiratory cancer is often driven by chronic irritation, so any of these irritating exposures could have played a substantial part in the excess lung cancers identified in these cohorts, especially during the earlier years of high exposure and lax protective practices. Grimsrud et al. (2005) adjusted for some (but not all) of these exposures in their case-control study, but it should be noted that the smoking-adjusted relative risk for any Ni exposure in that study is only a non-significant 1.4, compared with smoking-adjusted RR of approximately 2-5 in the cohort study. In the cohort study, Grimsrud et al. (2003) adjusted for cigarette smoking using internal RR analyses (although not in the SIR), but this produced more variable results than the SIR analyses because of the smaller reference population used for comparison. Therefore, given the substantial variability and uncertainty suggested by the above, it is difficult to say with any confidence where the true relative risk lies, especially at low levels. In fact, the ICNCM concluded that "risk to the general population from exposure to the extremely small concentrations (less than 1 ug Ni /m3) to which it is exposed in the ambient air is minute, if indeed there is any risk at all."

The TCEQ has applied a complex series of models consisting of Poisson regression, least-squares regression, BEIR IV models, and piecewise linear distribution functions. Applying such a complex approach to highly variable and uncertain epidemiologic data is akin to measuring something to within an inch, then cutting it out with a laser that can accurately measure to within 0.01 mm. The sophisticated modeling adds an illusion of certainty, precision, and objectivity, but this is not borne out by the underlying data. The "simplistic" approach used by EPA would seem to be more appropriate. The most appropriate approach is probably meta-regression using inputs from all available data sets.

The TECQ also used a complicated system of weighted averages to generate summary URFs, first averaging smoking-adjusted RR and SIR from Grimsrud et al., then averaging this summary estimate with the Enterline and Marsh results. This approach begins by combining two dependent results, producing a summary that is similar to either input alone, and then combines this average with independent results produced via post-hoc manipulation of Enterline and Marsh.

This messy and somewhat convoluted process produces a combined summary ($\sim 2x10^{-4}$) that is not very different from the results from Grimsrud et al. alone ($\sim 3x10^{-4}$). As already mentioned in regards to modeling, this complex combinatorial process may add the appearance of certainty, precision, and objectivity, but this is not borne out by the underlying data. As mentioned previously, either the simple approach used by EPA or a meta-summary would probably have been preferable.

The modeling discussion on pages 39 and 46 of the TD is convoluted and difficult to follow. Furthermore, the authors of the TD provide no details on how the various modeling exercises were performed. Appendix B provides general guidance on how the linear multiplicative risk model should be performed, but does not appear to describe the actual process and specific calculations performed by the TCEQ. Descriptions or summaries of the actual modeling exercises should be included either in the text or with an appendix. Perhaps most importantly, the TCEQ provides no information regarding whether regression diagnostics were performed, or the results of those diagnostics. Regression diagnostics provide important information on model fit, the integrity of modeling assumptions, collinearity of variables, and the influence of outlying observations. In the absence of diagnostics, one is forced to blindly believe in modeling results that may not accurately reflect reality.

Reviewer 4: *In general, I would say that the statistical and modeling approaches used were appropriate. The methods used here are a definite improvement over the average relative risk calculations from earlier assessments.*

I have one specific comment about portions of the analysis. When using the maximum likelihood approach for estimation of parameters, I think it would be much better to derive bounds on those parameters (in particular on β which is used as the basis for URF determination) through the use of the profile likelihood method. Bounds that are based on asymptotic approximations to the variance, i.e., the ones reported in the nickel DSD, are known to be problematic in certain cases. The profile likelihood method is one that avoids the asymptotic approximation of the variance and which finds bounds while simultaneously considering the variability of the other parameters.

Reviewer 5: *The multiplicative relative risk model is appropriate. Where data on observed and expected numbers of cases were available, Poisson regression was appropriately used with maximum likelihood estimation of the parameters of the model. Although detailed information on the BEIR IV approach to calculating URFs is given in Appendix E, the DSD needs to spell out in greater detail in the body of the document (e.g., page 42, lines 19-27) exactly how air concentrations were calculated using the BEIR methodology. For example, where does $\beta x d_j$ in the relative risk model fit in the BEIR formulas? Is $\beta x d_j$ in the relative risk model equal to e(i) in the BEIR formulas? If so, don't the BEIR formulas simplify considerably? With the BEIR approach, isn't the air concentration calculated*

first, and then the URF calculated as the ratio of the extra risk (10^{-5}) to the air concentration, rather than the other way around?

- **Was the dose metric selected (mg/m^3-months or -years) the most relevant and appropriate choice?**

Reviewer 1: *It is the most appropriate choice given the URF is based on human studies. If it were based on animal studies, exposure concentration (vs. cumulative exposure) would be more relevant.*

Reviewer 2: *This is a reasonable default in the absence of knowledge on dose metric or more specific data.*

Reviewer 3: *This measure of cumulative exposure seems appropriate given the nature of the carcinogenic process and the underlying data.*

Reviewer 4: *A cumulative exposure metric is the typical metric used with cohort studies, so it appears to be relevant here. On the other hand, one often will see examinations that explore a lag between "effective" cumulative exposure and cancer risk (e.g., a lag of 5 or 10 years to account for the time between induction of the cancer and the appearance on a physical exam, for incident cases, or death, for mortality). Examination of such lags is probably not possible in this case because the analyses were based on summary data.*

Reviewer 5: *The dose metric, mg/m^3–year, is relevant and appropriate.*

- **Is use of total nickel for both studies, and all workers for Enterline and Marsh (1982), justified given the purpose of the URF and carcinogenic ESL and in light of the recent work by Goodman et al. (2009)?**

Reviewer 1: *Because the animal and MOA data clearly show that only insoluble nickel is likely to be carcinogenic, only sulfidic and oxidic nickel should be used for calculating URFs and ESLs. Because these data are not provided in the Enterline and Marsh (1982) study, this study should not be used for this analysis (making the question regarding the use all workers moot).*

Reviewer 2: *I do not believe so. Use of total nickel would be appropriate if the quantitative composition of the different nickel species in the studies used for the quantification were similar to that in Texas air. In that case, one would be comparing two similar mixtures, and it would be appropriate to base the risk on epidemiology data from a well-studied population exposed to the mixture. However, the Goodman et al. (2009) study suggests that the Kristiansand exposures may have had much higher levels of sulfidic nickel than what was reported (and lower levels of nickel sulfate), due to oxidation of the nickel*

particles after deposition on the filter. This would mean that the two mixtures are substantially different, and specifically that the Kristiansand cohort was exposed to a much more potent mixture than that present in Texas air. Using the quantification from that data would be over-conservative. Furthermore, a key reason for choosing the Kristiansand cohort (and the West Virginia one studied by Enterline and Marsh) and not the other cohorts was based on the desire for a cohort with low sulfidic nickel exposure. If exposure at Kristiansand was primarily to sulfidic nickel, a key reason for using that cohort would be removed. Thus, while use of total nickel removes issues related to exposure misclassification, it does not address the issue of whether the worker exposure was to a mixture relevant to evaluation of ambient Texas air.

Similarly, as noted above, inclusion of the refinery workers in the Enterline and Marsh (1992) study means that a group with relatively high subsulfide exposure is included. Use of total nickel may be appropriate for the non-refinery workers and/or the workers hired after 1946, ***because*** *they had low subsulfide exposures – a key criterion in the choice of cohort to use for the analysis.*

Reviewer 3: *The authors of the TD point out that there is substantial uncertainty within the epidemiologic literature regarding the carcinogenic potential the various species that make up soluble and insoluble nickel compounds. Goodman et al. add to the discussion on this topic, but do not definitively resolve the issue. Therefore, it would seem appropriate to treat soluble and insoluble Ni as equivalent exposures when generating conservative and health-protective quantitative risk assessment. However, there appears to be little compelling evidence of a substantive carcinogenic hazard from metallic nickel, so it is not appropriate to apply the derived potency to metallic exposure, which is the predominant exposure in TX.*

Reviewer 4: *I believe the use of all workers in the Enterline and Marsh study was appropriate, given the uncertainty about what forms of nickel might be responsible for carcinogenic responses. If the metric of choice is total nickel exposure, then there is no basis for selecting from and differentiating among the different subcohorts of that study. Furthermore, a reason not given in the DSD is that the inclusion of all workers gives a more "robust" basis for estimating the dose-response pattern, especially in the important low-dose region, because those workers employed after 1946 or who did not work in the refinery had much lower cumulative exposure estimates. Thus, they provide more information about the responses likely to occur outside the occupational setting.*

And I do believe that the use of total nickel is the best supported choice at this time. The evidence cited in the DSD and the conflicting results and conclusions from the various studies do not allow me to decide that one form is more carcinogenic than any other form.

However, I question the statement on p. 37 (lines 34-35) about use of total nickel alleviating "the significant uncertainty associated with attempting to attribute cancer risk to a particular form or forms of nickel." That uncertainty has not magically gone away; rather it is transferred to an equally significant uncertainty associated with the assumption that is being made in the reported assessment: that all forms for nickel are equally relevant for determining cancer risk. I think that that assumption is supportable for the purposes of the risk calculations, but that does not mean that the uncertainty can be ignored or has somehow disappeared because that assumption has been made.

As a follow-on comment, I believe the document is lacking with respect to presentation of sources of uncertainty and their consequences on the risk estimates. Uncertainty considerations, and their implications for a risk assessment, are by no means easy to convey, but some attempt to itemize and "address" the major uncertainty issues should be added to this document.

Reviewer 5: *Because the purpose of the URF and carcinogenic ESL is to address potential cancer risk from all forms of inorganic nickel, the use of total nickel is justified even though soluble nickel has been shown not likely to be a complete carcinogen and may be carcinogenic only in the presence of insoluble nickel compounds. The TD justifies the use of all workers from the Enterline and Marsh (1982) study, including refinery workers hired before 1947 who were exposed to higher levels of nickel subsulfide, because the possibility of some nickel subsulfide exposure due to emissions from facilities in Texas cannot be excluded, This rationale seems justified.*

- **Are the most appropriate URFs from each study used to calculate the final URF?**

Reviewer 1: *The smoking-unadjusted URF in the Grimsrud et al. (2003) study should not be included in the calculation of the final URF. When smoking data are unavailable, one is forced to use unadjusted risk estimates if basing a URF on epidemiological data. Because smoking data are provided in the Grimsrud et al. (2003) study, it is inappropriate to use the smoking-unadjusted values for calculating the final URF. In addition, the URF should be based on insoluble nickel (sulfidic and oxidic), not total nickel, in the Grimsrud study.*

Reviewer 2: *Based on the previous comments, it appears that the URF based on Enterline and Marsh (1982) for "hired after 1946+ non-refinery hired before 1947," corrected for Texas background rates, may be the most appropriate from that study. I do not agree with the use of the Grimsrud study, as noted above, but if it is used, the appropriate URFs appear to have been chosen.*

Reviewer 3: *See statistical discussion above.*

Reviewer 4: *The choice to use URFs calculated using Texas rates (mortality or incidence) is consistent with the desire to characterize the risks associated with emissions typical of or relevant to the Texas population. Also, it is correct to choose the smoking adjusted estimates, when available.*

Reviewer 5: *The URF based on smoking-adjusted RRs for lung cancer incidence from Grimsrud et al. (2003) was chosen over the smoking-unadjusted RRs from that study, which is appropriate. The smoking-unadjusted SIRs from the same study are also used, because, apparently, it is not possible to derive smoking-adjusted SIRs. The final URF for the Grimsrud et al. study is a pooled URF based on the two separate URFs, each weighted by the number of person-years times the reciprocal of the variance of the slope from the multiplicative relative risk model. This is appropriate. However, I do not follow the logic of the statement in lines 8-10, page 44, which seems to justify choosing the smoking-unadjusted SIR over the smoking-unadjusted RR. The smoking-adjusted RR has already been selected over the smoking-unadjusted RR. Why is the smoking-unadjusted RR relevant to the smoking-unadjusted SIR? The RR and SIR are two different measures. The URF based on excess respiratory cancer mortality from the Enterline and Marsh (1982) study is appropriate. In all cases, the URFs are based on Texas-specific mortality and survival rates, which are the most appropriate. In all cases a central estimate of the slope of the multiplicative relative risk model has been used to derive the URFs. The appropriateness of that choice is addressed in the next question.*

- **Is use of the central estimate URFs justified for reasons discussed in the DSD?**

Reviewer 1: *It appears that the justification for the use of the central estimate is that it is consistent with TCEQ (2006).*

Reviewer 2: *Yes, the rationale follows the ESL guidance.*

Reviewer 3: *Central estimates are appropriate in that they provide the best guess (maximum likelihood estimate) for a value.*

Reviewer 4: *The use of the central estimate was justified in part because incidence data were available (as opposed to mortality data) in the analysis of the Grimsrud data. But that is not the case for the Enterline and Marsh data analysis, which used mortality for all respiratory cancers. This is stated (p. 51, line 9) to be essentially equivalent to lung cancer incidence. Admittedly, the mortality rates for all respiratory tract cancers would be greater than those for lung cancers, and lung cancer incidence rates would be greater than lung cancer mortality rates, but concluding that those two things are essentially equivalent just because they have rates that are greater than lung cancer mortality rates does not appear to be enough of a reason. For one thing, how do the age-related rates compare to one*

another? At the very least, the document should show some additional supporting evidence (like Figure 3, but plotting respiratory cancer mortality and lung cancer incidence).

In addition, if the goal is to protect public health (not restricted to public mortality but extending to morbidity as well), then it is questionable whether or not use of predicted effects on lung cancer incidence should be used to support use of the central estimates rather than the upper bound estimates. The truly health protective choice would be to use the upper bound β estimates to derive the URFs. This is even more the case because many significant uncertainties not associated merely with model fit and uncertainty about model parameter values have not been considered in this document at all. Uncertainties about exposure reconstruction, misclassification (even for total nickel exposure), choice of reference rates, use of summary data, selection of study cohorts, and many others are likely to contribute much more uncertainty than that associated with the estimation of β in the dose-response models. Until and unless those potentially substantial uncertainties are shown to be not important (or to be such that risk would only, or predominantly, be overestimated by the choices made in the current analysis), I would recommend the use of the upper bound β estimates (improved, where possible, by using the profile likelihood approach to deriving those bounds as mentioned in a response to an earlier question).

Reviewer 5: *The use of the central estimate URFs might be justified, but additional clarification and justification are needed. The apparent justification is given in bullet point two on page 44 of the DSD: "As incidence data were available and utilized as opposed to mortality data, use of the URF (MLE) based on the central β estimate was preferred over use of the β (95%UCL), consistent with TCEQ (2006)." The justification is also stated in lines 8-9 of page 51. By themselves, these statements don't sufficiently justify the use of the central estimate. The only place in the Guidelines document (TECQ, 2006) where I found any mention of using a central estimate (EC) versus a confidence limit (LEC) was on page 68, where it is stated: "When TS staff develops a toxicity value for a carcinogen based on a human epidemiology study, the following types of uncertainties of the human study are considered in order to determine whether to use the EC or the LEC as the POD (Section 3.2.12 of USEPA 2005a):*

- *when estimates of mortality are available rather than incidence because survival rates for different cancers vary;"*

This bullet does not actually state what the rationale is, or even which estimate is used with which type of data. There are other bullets that follow this one in the Guidelines document, but this is the only one that mentions mortality versus incidence. Thus, the justification in the DSD is not clear, and the guidance itself is not clear. The rationale for using a central estimate needs to be clarified in the DSD.

- **Was cancer endpoint selected as the basis of the potency estimates (lung**

cancer) the most appropriate and relevant choice?

Reviewer 1: *Yes. The vast majority of respiratory cancers associated with inhalation of nickel dust is lung cancer. Calculations based on this cancer type will account for the majority of cancers and result in the most stable risk estimates.*

Reviewer 2: *Yes.*

Reviewer 3: *There is substantial epidemiologic evidence for an association between lung cancer and exposure to nickel refining, but little consistent evidence for cancer at sites other than the respiratory system. Therefore, lung cancer would seem to be a relevant and appropriate choice.*

Reviewer 4: *I consider these two questions together, since the answer to the second question bears on the answer to the first one. I do believe that using respiratory tract cancer rates (observed and expected) is a reasonable choice. I believe that not (or not only) because I think that lung cancer dominates the respiratory tract cancers. Rather, it should be acknowledged that the full carcinogenic effect of nickel is not known; it may have an effect on the respiratory tract more generally, so it is appropriate to examine the observed and expected results for the full respiratory tract. As stated above, I cannot accept that respiratory tract mortality is a surrogate for lung cancer incidence.*

Thus, I must say in response to the first question that looking at lung cancer incidence in the Grimsrud cohort appears to be more of a case of "we'll analyze what we have available" rather than a conscious choice to pick the most appropriate and relevant endpoint. That being said, I believe that the lung cancer incidence analysis for Grimsrud is acceptable; there was an effect and it was exposure related. It is just that the attempt to justify the choice of lung cancer incidence in one case and respiratory tract cancer mortality in the other case appears to be unnecessary and susceptible to potentially problematic claims (i.e., that respiratory tract cancer mortality is essentially the same as lung cancer incidence).

Reviewer 5: *Yes, lung cancer has consistently been found to be associated with airborne nickel exposure in many epidemiological studies. While other types of cancer have been found in nickel workers in some studies, no consistent patterns of increased nonrespiratory tract cancer risks have been found.*

- **Are respiratory cancer data from Enterline and Marsh (1982) a reasonable surrogate for lung cancer for reasons discussed in the DSD?**

Reviewer 1: *Yes, because the majority of respiratory cancers were lung cancers.*

Reviewer 2: *Yes.*

Reviewer 3: *Yes.*

Reviewer 4: *See above.*

Reviewer 5: *The data from Enterline and Marsh (1982) are respiratory cancer mortality. The DSD states that more than 93% of the observed and expected respiratory cancers were lung cancers. Thus, the respiratory cancer mortality data from Enterline and Marsh are a reasonable surrogate for lung cancer mortality.*

- **Are lung cancer incidence and mortality sufficiently similar as to be comparable for purposes of this assessment for reasons discussed in the DSD?**

Reviewer 1: *Yes*

Reviewer 2: *This is beyond my independent knowledge base, but based on Figure 3, they do appear to be sufficiently similar.*

Reviewer 3: *Lung cancer is generally highly aggressive, and often leads to death within 6-12 months. Therefore, incidence and mortality data should provide comparable results.*

Reviewer 4: *As mentioned to in the previous response, I see no reason to worry about their comparability. The life table approach can work as well with incidence data as with mortality data. I would be inclined just to consider the two estimates as separate and independent. In which case, one would probably want to use the more health protective estimate as the basis for the ESL for cancer.*

Reviewer 5: *The DSD states that lung cancer mortality is reasonably predictive of lung cancer incidence because the five-year survival is only about 15%. Thus, it seems reasonable to consider lung cancer incidence and mortality sufficiently similar as to be comparable for purposes of this assessment.*

- **Is the URF weighting procedure used to calculate the final URF reasonable and justified?**

Reviewer 1: *When smoking data are unavailable, one must use unadjusted risk estimates if basing a URF on epidemiological data. Because smoking data are provided in the Grimsrud et al. (2003) study, it is inappropriate to give any weight to the smoking-unadjusted values for calculating the final URF. The weighting procedure for the Grimsrud et al. (2003) study vs. the Enterline and Marsh (1982) study is appropriate.*

Reviewer 2: *Given the assumptions and "sufficiently similar" considerations (e.g., predictivity within 90-95% for incidence vs. mortality being appropriately considered sufficient) noted in my comments and in the charge questions, the weighting procedure seems overly precise, in light of the quality of the data. The weighting procedure assumes that the studies are entirely equivalent and the only differences of note are the number of person-years of follow-up and the variance. As noted in these comments, there are critical issues of exposure measurement and exposure composition. The impact of these latter issues would be expected to greater than (or at least equal to) that of the two weighting factors. This is particularly true in light of the potential differences in potency of the mixtures to which the workers were exposed (differences between the two studies, and differences between the study population and ambient exposure conditions for Texas air). In light of this uncertainty, I would recommend merely identifying the best overall URF. Alternatively, TCEQ could take the weighted average of the two Grimsrud estimates (since they are from the same cohort), and then take the mean of that estimate and the one from Enterline and Marsh (1982). (This suggestion assumes that the three estimates noted are still the preferred ones in a revised assessment.)*

Reviewer 3: *See statistical discussion above.*

Reviewer 4: *I liked the idea of the weighting procedure that was described here. In particular, I was favorably impressed by the consideration of both the sample size (as in the number of person-years) as well as the precision (variance) of the parameter estimate to derive the weights. I agree that, if this is going to be done, the combination of the two Grimsrud-based estimates was best done in one step and then that average could be averaged with the value from Enterline and Marsh.*

On the other hand, if one is skeptical about the claimed equivalence of respiratory tract cancer mortality (from Enterline and Marsh) and lung cancer incidence (from Grimsrud), then the motivation for combining them is weak. Thus, while I like the approach as it is presented in the DSD, I am not certain that is should have been used here.

Reviewer 5: *Yes. The two URFs from Grimsrud et al. (2003) were first pooled by weighting each one by the number of person-years times the reciprocal of the variance of the slope. Then, that resulting URF was pooled with the URF derived from Enterline and Marsh (1982) by the same weighting scheme to calculate the final URF.*

- **Is the decision not to apply age-dependent adjustment factors (ADAFs) to the URF to account for potential increased sensitivity of children justified and**

properly considered?

Reviewer 1: *Yes*

Reviewer 2: *Yes, I agree with the conclusion, and believe it was properly justified.*

Reviewer 3: *I agree with the decision of the TCEQ in this regard. Nickel is not thought to be mutagenic and most probably acts through epigenetic mechanisms. Therefore, there is no clear a priori reason to believe that low-level exposure would represent a disproportionate lung cancer risk to children, who are exposed primarily via diet.*

Reviewer 4: *I have no particular expertise in relation to factors that do or do not make children more sensitive. Therefore I do not have any comment to make in relation to the selection of ADAFs, except to say that the decision not to apply ADAFs appears to be consistent with the other decisions that related to lack of knowledge about mechanisms of action.*

Reviewer 5: *Yes, the decision not to apply age-dependent adjustment factors to the URF to account for potential increased sensitivity of children is properly considered and justified because the MOA for nickel-induced lung cancer has not been determined to be mutagenic.*

Other comments on the assessment:

Reviewer 1: *None*

Reviewer 2: *I commend TCEQ for presenting the MLE as well as both upper and lower bounds for β. It would be useful to present some additional context. The LCL for the RRs has a negative slope, i.e., exposure decreases risk. This is a meaningful difference from the MLE. Are such results common in epidemiology studies, and what does that mean with regard to confidence in the conclusion? TCEQ did appropriately recognize the uncertainty with the bounds, using that as part of the rationale to use the SIRs for the final β, rather than the RRs.*

P. 53, line 41 appears to be missing the notation to divide the sum by 2. Otherwise, the described pooled URF is the sum of the individual URFs, rather than the weighted average.

Reviewer 3: *None.*

Reviewer 4: *None.*

Reviewer 5: *There are some statements in the DSD regarding the carcinogenicity of soluble nickel that seem contradictory. On page 27, lines 42-44, it is stated that*

the available weight of evidence does not suggest that exposure to soluble nickel, in the absence of carcinogenic compounds, will increase the risk of cancer. However, the paragraph that starts on page 28, line 40 and ends on page 29, line 11 seems to try to make the case that soluble nickel is a carcinogen by itself, although there are some statements in that paragraph that seem to contradict that notion as well. I also found the paragraph on page 30, lines 3-25 to be confusing regarding the carcinogenicity of soluble nickel. I recommend that all the statements in the DSD regarding the carcinogenicity of soluble nickel be reviewed for consistency.

An argument is made on page 42, lines 23-27 about needing to adjust the BEIR IV methodology to reflect the use of lung cancer incidence instead of lung cancer mortality. Detailed mathematical formulas are presented in Appendix E, and the adjusted BEIR methodology is use for the Grimsrud study. For the Enterline and Marsh study, a rationale is given on page 34, lines 16-20 for treating the respiratory cancer mortality as lung cancer incidence. On page 49, lines 6-9, it is stated that the BEIR IV methodology was used for the Enterline and Marsh study. I believe that it needs to be clarified whether the unadjusted or adjusted BEIR methodology was use for Enterline and Marsh, and why.

This page intentionally left blank.

Part B: Report from Follow-Up Conference Call

On October 1, 2009, *TERA* facilitated a follow-up conference call between the panel and TCEQ. Conference call materials (available at the above website), including a focused charge, attached as Appendix B, and the reviewer comments in Part A, were distributed prior to the call; members of the public were allowed to listen to the call. The purpose of this call was to allow TCEQ to ask the panel questions regarding their written comments and to allow the panel members to discuss issues on which there were divergent opinions in the written comments.

Cancer Weight of Evidence and Unit Risk Factor (URF)

Hazard Characterization of Nickel Mixtures

The toxicity values generated in the Nickel DSD will be used to evaluate emissions from facilities during the air permit review process, for evaluation of ambient air monitoring data, and as toxicity factors in the Texas Risk Reduction Program remediation program. Because TCEQ typically evaluates total nickel and does not have nickel compound-specific air data, TCEQ used a mixtures approach for the derivation of URFs. The challenge is to select an appropriate mixture, which will be health protective without being unduly conservative, on which to base the cancer assessment.

> **Given the wide spectrum of toxicity of the different nickel species, the needs that the TCEQ faces in dealing with mixtures, and the fact that the TCEQ only receives total nickel data, is using a mixtures approach appropriate for a nickel cancer assessment?**

The issue under discussion is the appropriateness of using a "total nickel" approach compared to conducting a separate assessment for each nickel species. In addition, the panel was asked to evaluate whether the data were sufficient to develop an adequate description of the nickel mixture likely to be found in Texas air. One panel member suggested that the only justifiable way to use a mixtures approach for the ESL development would be if there are human data with exposure to a similar mixture. However, this reviewer did not think that the mixtures in the human studies had been sufficiently characterized to evaluate the similarity with Texas air. Therefore, this panelist suggested that a better approach would be to understand the composition of nickel species in Texas air and then compare the concentrations for each form of nickel to the cancer potency estimates for each nickel species derived from the animal cancer bioassay data.

TCEQ agreed that having a better characterization of the composition of nickel in Texas air would be useful. However, the air monitoring data that are currently collected only

give measurements of total nickel, so TCEQ cannot determine whether any particular nickel species dominates in any particular site in Texas. Based on general information about nickel levels in ambient air that is in the literature and knowledge of the industrial processes and operations that affect nickel emissions in Texas, TCEQ expects that nickel mixtures in Texas air include a combination of soluble, oxidic, and metallic nickel, with little, if any sulfidic species. However, because monitoring data only report total nickel, it will be highly unlikely that the State will know what form of nickel is present in any given sample. TCEQ also indicated that it can request information from permitted facilities about their processes and so can estimate on a case-by-case what species may be contributing to the total nickel measurements. Such data may allow TCEQ to make adjustments to permit requirements, different than would be allowed solely on the basis of risk estimated from the total nickel data and the ESL.

One panel member noted that there will be uncertainties in both a mixtures approach and an individual species approach. However, this reviewer felt that using animal data to derive species-specific ESLs avoids the limitation of basing ESLs on human studies where the exposures are predominantly to sulfidic species with no clear way of demonstrating that such mixtures are representative of Texas air. Therefore, there would be less uncertainty overall if ESLs were based on the animal data.

Another reviewer stated that, generally, use of a mixtures approach is appropriate. This reviewer discussed U.S. EPA guidance on mixtures risk assessment and noted that EPA established the following priority for choice of data to assess a mixture: data on the same mixture, data on a similar mixture, and then data on the components of the mixture. Therefore, the key question to be addressed in the risk assessment is how similar are the mixtures in Texas air compared to the mixtures in the epidemiology studies.

The panel then considered the issue of whether the data from the epidemiology studies characterized exposure sufficiently to select a study (or studies) that evaluated a "similar mixture" to that in Texas air. One panel member did not believe that it was realistic to select with any degree of confidence only those epidemiology studies where nickel speciation was similar to that in Texas air. Moreover, this reviewer felt that it was not possible to make a clear determination from epidemiology data which nickel species are more carcinogenic. Because there is co-exposure to a variety of nickel species as well as other compounds, it is not possible to determine what was causing the effects in each of the epidemiology studies. Therefore, based on these two considerations this reviewer concluded that a meta-analysis of all well-conducted epidemiology studies would be a better approach given the uncertainties, since this avoids making *a priori* choices of studies to include in the analysis. However, other reviewers disagreed, stating that it is reasonable to select studies that are more relevant to the types of exposure that are anticipated in Texas air based on TCEQ's assumptions about relative prevalence of nickel forms.

Generally, the panel agreed that there are not sufficient data to characterize with certainty the mixture of nickel species likely to be found in Texas air. One reviewer stated that the document's description of the composition of the nickel mixture in Texas air was

confusing. For example, the document states that nickel emissions in Texas will be primarily metallic nickel, but that ambient air is expected to contain primarily soluble nickel. Therefore, this reviewer found it difficult to address the issue of which epidemiology studies had exposures similar to Texas air.

TCEQ responded that the information presented in the document came from its knowledge of the process used by the organizations that are the top emitters on the Toxic Release Inventory and based on a publication (Schaumlöffel 2005) indicating that ambient air generally contains primarily soluble nickel. TCEQ also indicated that its intent was not necessarily to characterize the nickel mixture in Texas air, but to provide support for the conclusion that little or no sulfidic nickel is likely to be found in Texas air. The reviewer suggested that this purpose is appropriate, but that the document should clearly state that purpose.

Another panel member suggested that there is not really a discrepancy between emissions of metallic nickel and the presence of soluble nickel in ambient air. If metallic nickel is present in facility emissions, it will be in the form of large particles that are deposited from air easily. Soluble nickel is present as smaller particles and so would remain in air and be available for exposure. This reviewer agreed that TCEQ should add all of these explanations to the document, and also asked about TCEQ's process if it found an exceedance of the nickel ESL for ambient air. TCEQ responded that it would take action to reduce the levels, but would also look at the potential source(s) and evaluate what species was likely to have been emitted from that source based on its process. TCEQ would then decide how to adjust or apply the ESL, if necessary.

One reviewer suggested that TCEQ extrapolate from other data to estimate the nickel species that are likely to be in ambient air. For example, this panelist noted that a study cited in the public comments submitted by the Nickel Producers Environmental Research Association (NiPERA) provides data on nickel species in Florida ambient air. Galbreath et al. (2003) reported that the biggest contributors of nickel to the ambient air are power plants and natural releases. This study also concluded that sulfidic species are minimal in ambient air. Another panel member agreed that TCEQ should attempt to get an upper bound estimate of the amount of each species likely to be present in ambient air. Because each of the different nickel species has such a different risk, using this approach with information on relative potency from the animal data could help provide a sensitivity analysis regarding the degree to which an ESL based on the epidemiology study(s) provides a conservative estimate of risk.

Summary of Discussion on Characterization of Nickel Mixtures:

In summary, the panel agreed that if TCEQ were able to adequately characterize the nickel species in Texas air and if the nature of the exposures in the epidemiology studies were adequately characterized, then using a mixtures approach to develop the ESLs would be appropriate. However, the panel agreed that the data are not sufficient to accomplish this degree of characterization. Therefore, as an alternative the panel unanimously agreed to the following conclusions and recommendations:

- Because of TCEQ's regulatory process and the fact that only data on total nickel are available from air monitoring, the panel agreed that the derived ESL should be based on total nickel.
- The panel agreed that there are insufficient data to accurately describe the actual nickel speciation in Texas air. However, available data on air composition of nickel species in other states as well as any data and assumptions that TCEQ used in making conclusions about relative prevalence of nickel species in Texas air should be added to the DSD document to enhance the discussion.
- TCEQ should also add a discussion in the document of its purpose for describing the data regarding the air composition of nickel species and how such data would be used for the ESL program.
- If the ESL is based on data sets with significant exposures to sulfidic nickel, then the risk characterization should note the resulting ESL value as likely to be conservative. This is because, overall, the data appear to support the conclusion that in ambient air sulfidic forms are minimally present as compared to the occupational epidemiology studies.
- The document should more fully separate the concepts of how the range of data was used to support the hazard characterization versus the dose-response assessment. A brief discussion could be presented of the range of suggestions from panel members, including using all the human data, using a subset of ,most representative epidemiology studies, and using animal bioassays to help characterize the uncertainty in the epidemiology approach.

Weight of Evidence

In the written peer review comments there was no consensus among the panel members on the appropriateness of the weight of evidence statement. In particular, there was disagreement on whether it was appropriate to assess WOE for nickel compounds as a group compared to WOE for the individual nickel species, or whether it is feasible to assess WOE for the mixture of nickel species that is typical of Texas air. In addition, Reviewer #1 indicated that there is not sufficient evidence for carcinogenicity of nickel sulfate alone, but only in the presence of other nickel compounds.

> **If a mixtures approach to nickel assessment is appropriate, would a WOE for mixtures of nickel be defensible? If so, how should data for the different nickel species be taken into account and how should the WOE narrative address the different forms of nickel likely to be present in the mixture?**

One panel member started this discussion by noting that it is important to use the animal data to help frame the WOE discussion for nickel. The animal data are critical for understanding the carcinogenicity of the individual nickel species since epidemiology

data do not provide this type of information. This reviewer noted that the NTP studies provided summary conclusions of carcinogenicity for soluble nickel, nickel oxide, and nickel subsulfide, and the recent Oller et al. (2008) study provides a summary conclusion regarding the carcinogenicity of metallic nickel. The NTP studies concluded that there was no evidence of carcinogenicity for soluble nickel and clear evidence of toxicity for nickel subsulfide, and this information must be included in the WOE narrative. Other panel members agreed with this conclusion, noting that in order to develop a WOE narrative for nickel compounds as a group, it is critical to understand and describe the potential for carcinogenicity of the individual compounds. In particular, the panel was concerned about how to frame the WOE narrative when sulfidic nickel species are the most probable carcinogenic species, yet the smallest contributor to exposure via ambient air. One suggestion was that TCEQ should estimate the relative proportions of nickel species that are likely to be present in the air and modify the WOE statement based on the relative proportions.

One reviewer noted that TCEQ's nickel document concluded that nickel compounds, as a group, are carcinogenic to humans. However, according to standard risk assessment practices, this WOE descriptor requires clear evidence of carcinogenicity in humans. This reviewer indicated that, in the reviewer's opinion, this WOE descriptor is only appropriate for nickel refinery dust. The data are not as clear for any other nickel species. The reviewer stated that if no nickel subsulfide is likely to be present in Texas air, then a WOE descriptor of "carcinogenic to humans" is not appropriate. This reviewer stated that even a WOE descriptor of "likely to be carcinogenic" would be conservative for the remaining nickel species that are expected to be found in ambient air. Another panel member agreed, noting that if the nickel mixture in Texas air includes even some metallic nickel, then that suggests a lesser descriptor.

Summary of Discussion on Nickel Weight of Evidence

The panel agreed to the following conclusions and recommendations regarding the weight of evidence statement for nickel:

- The panel recommended that TCEQ incorporate more of the available data on the carcinogenicity and mode of action for each nickel species into the weight of evidence discussion. It is important to better integrate the epidemiology, animal data, and mode of action studies in developing a weight of evidence statement. In addition, the document should discuss the weight of evidence of each nickel species and how each species contributes to the overall weight of evidence for nickel compounds as a group.
- The panel recommended that TCEQ estimate the composition of nickel species in Texas air and use this to weight the overall weight of evidence and clearly describe the assumptions and conditions under which the descriptor applies, including key uncertainties and their impacts on the interpretation of the weight of evidence descriptor.

Selection of Critical Study

The written peer reviewer comments expressed a broad range of opinions on which studies would be appropriate as the critical study. Reviewer recommendations included using the animal studies to derive the URF, adding some of the other human studies in addition to Grimsrud and Enterline and Marsh to the analysis, and using all of the human studies to derive the URF via a meta-analysis. Specifically, Reviewer #1 suggested not using the Enterline and Marsh study because there were no statistically significantly increased risks. Reviewer #2 suggested not using the Kristiansand study due to uncertainties in the exact nature of the exposure. Reviewer #3 suggested conducting a meta-analysis using data from all human studies with suitable exposure estimates. For the URF based on the Enterline and Marsh study, the reviewers' comments were divided on the appropriateness of using total nickel and all workers. Two reviewers commented that only insoluble forms of nickel are carcinogenic and so only exposure estimates from these species should be used. The remaining reviewers suggested that the approach was appropriate and conservative given the uncertainties in the literature regarding the carcinogenic potential of the various species that make up soluble and insoluble nickel compounds.

Given the issues discussed above for hazard characterization, and the preference of TCEQ for human data, which studies are considered superior by the reviewers for deriving a somewhat conservative generalization of risk to the population of Texas and why?

What specific URF analyses do the reviewers suggest as more applicable for the evaluation of total nickel in ambient air data? What epidemiology studies would be appropriate for TCEQ to use to develop a URF? The following studies are suggestions from either reviewers or the public for alternate studies other than Enterline and Marsh and Grimsrud:

- **The Clydach cohort after 1930**
- **Harjavalta, Finland (Karjalainen et al., 1992 and Anttila et al. 1998)**
- **Public Comments (Nickel Producers Environmental Research Association): Arena et al. (1998) and Sivulka and Seilkop (2009)**

If nickel compound-specific URFs were derived based on animal data, how should they be applied to a nickel mixture so as to not grossly over- or under-estimate risk?

Given the cohort in Enterline and Marsh 1982 appears to have been exposed to a nickel mixture appropriate for Texas air, is it reasonable to exclude this study from use in URF derivation because it does not have statistical significance? Could the Enterline and Marsh study be used as a supporting study and how?

What is the panel's opinion of the Grimsrud study given that is appears to contradict findings in animals regarding carcinogenicity of soluble nickel compounds?

The panel started with a general discussion about their ideas on choice of critical study. One reviewer expressed the opinion that no single study should be "the" critical study. This panelist stated that Grimsrud is a good study, but is not necessarily the only one with the right answer. Since there is no "right" study, then this reviewer recommended that all of the studies be analyzed in a meta-analysis, because this provides a better use of the entire body of the literature. Other reviewers agreed that no single study should be selected as a critical study, noting that there should be a more robust justification for the selection of studies to support the quantitative analysis and weighting of studies because other studies that were not selected may be reasonable choices with exposures just as relevant. For example, one reviewer commented that the Arena et al. (1998) study that was suggested in the submitted public comments appears to support the Grimsrud study and may be appropriate.

However, another panel member suggested that it is only appropriate to do a meta-analysis if the available studies have homogeneous exposures. In the studies available for nickel, there are no homogenous exposures or risks, so a meta-analysis does not seem appropriate. The first reviewer explained that a meta-regression (as opposed to a simple meta-analysis) is actually designed for situations in which there are heterogeneous data, because that data heterogeneity can be addressed by including appropriate parameters in the regression. Therefore, it would be possible to do an analysis that addresses the characteristics that contribute to risks. This reviewer felt that it was more important to include some assessment of all the epidemiologic data, precisely because there was heterogeneity that could not be readily explained. The first reviewer did not necessarily agree that a meta-regression could completely deal with heterogeneous data; although it is better than a meta-analysis. This reviewer stated that there would still be a need to decide which studies are most similar to include in the meta-analysis, but there are not enough data on nickel species in Texas air to make that decision. This reviewer stated that the speciation measurement in Grimsrud et al. (2002) was not accurate. The authors used an extraction method that removes increasingly less soluble forms, but if the collected nickel particles had become oxidized during the collection and extraction process, nickel subsulfide (for example) could be erroneously identified as nickel sulfate. Also, exposure was estimated based on job classification, which is less accurate than actual personal exposure measurements. Therefore, including the Grimsrud study in a meta-analysis will skew the results to a high risk, but including the Arena study will skew the results to no risk because the study did not have enough power to detect a response.

Another reviewer disagreed, noting that power is irrelevant in this context of adding the exposure and incidence data into a meta-regression; the data can still be used to derive a risk estimate because the regression weights studies based on sample size. This reviewer did not believe that the lack of a statistically significant increase should limit use of data if the study was well-conducted and included appropriate estimates of exposure and response. Another reviewer agreed, noting that the result (absence of a statistically significant increase) could be related to variability in the data.

Another panel member agreed that basing the URF on human data is the best approach, as long as appropriate data are available. This reviewer also considered the meta regression approach to be worth considering. However, for this approach to be appropriate, there must be studies with relevant types of exposure that account for speciation. In determining which studies are relevant, TCEQ needs to look at the ratio of different forms of nickel, not just the absolute amounts of nickel species when deciding which exposure is the most similar. This reviewer agreed that there is utility in including studies that are not statistically significant because they inform the overall picture.

Dr. Robert Sielken, a TCEQ consultant, replied that although the available epidemiology studies all provide useful information on the hazard characterization for nickel, most of the studies were not appropriate for doing quantitative estimates. The only studies that have quantified exposures are Enterline and Marsh, Grimsrud, and Copper Cliff. But Copper Cliff has very high levels of sulfidic nickel, and, therefore, is not relevant to Texas exposures. Therefore, TCEQ used the only two available studies for its URF.

One reviewer questioned this statement, noting that all of the epidemiology studies have some estimate of exposure. For example, this reviewer stated that on page 909 of the study, the Arena study gives a range and average of nickel concentrations. Dr. Sielken explained that neither the Arena study nor the Sivulka study relate the SMRs to measures of exposure and that several studies, including Clydach cohort after 1930, Harjavalta, Finland studies (Karjalainen et al., 1992 and Anttila et al. 1998), Arena et al. (1998) and Sivulka and Seilkop (2009), have absent or insufficient dose-response data for dose-response modeling. He noted that a minimum amount of quantitative epidemiological data is necessary for a scientifically defensible, quantitative dose-response model of how the probability of a specified response changes with exposure. Certainly, the better the quality of the data and the greater the amount of epidemiological data (e.g., individual jobs histories, individual time-dependent exposure estimates, information on confounding factors, etc.) the better it is for quantitative dose-response modeling. At the opposite extreme from quality data is one response frequency and some composite measure of exposure (e.g., a rough average exposure among exposed individuals). Dr. Sielken stated that this minimal information is insufficient to scientifically defensibly model how the response changes with the exposure. Multiple data points are needed to even have a chance of evaluating the assumptions made during the dose-response modeling. Included in these assumptions that need to be evaluated are assumptions about the shape of the dose-response relationship (e.g., linearity), assumptions about the response rate in unexposed individuals, etc. For nickel, the Enterline and Marsh and Grimsrud et al. studies provide far better data for dose-response modeling than the rest of the studies

where only one value for the overall SMR was given. The dose-response models fit to the Enterline and Marsh and the Grimsrud et al. studies are much more reliable because they are based on several exposure levels and based on fewer assumptions than potential fabrications based on the one-point studies published by other authors.

The panel discussed the issue of how the analytical methods used in the epidemiology studies and the use of a job matrix to estimate exposures could have underestimated the amount of sulfidic nickel to which people had been exposed. Several panel members expressed concern that TCEQ selected their critical studies on the assumption that these studies had lower sulfidic nickel exposures, but the analytical issues may mean that this assumption is not correct.

The panel then discussed some concerns regarding the individual epidemiology studies. For Grimsrud, the panel expressed concern regarding the shape of the dose response curve. One reviewer stated that the dose-response curve had different shapes in the Grimsrud et al. 2002 and 2003 studies. Dr. Sielken explained that in both Grimsrud studies the dose-response were similar, with the response to the lower doses increasing somewhat more rapidly than the response to the middle doses. This is apparent for water-soluble nickel (which is the only form of nickel quantified in both Grimsrud et al. 2002 and 2003). In Figure 1a from Grimsrud et al. (2002), the odds ratios were approximately 1, 1.25, 1.75, 1.85, 2.5, and 3.8 at 2- to 3-fold increasing doses; this is also seen in Tables 7 and 8 in Grimsrud et al. (2003). The data for total nickel in Grimsrud et al. 2003 were used by TCEQ to derive the URFs that were based on the Grimsrud et al. data. Therefore, TCEQ only used a linear model to evaluate the dose response. One panel member asked about the goodness of fit. Dr. Sielken responded that the TCEQ assessment did not conduct a formal assessment, but did not observe any gross areas of lack of fit.

For the Clydach study, Dr. Sielken indicated that he and TCEQ believed the study had no useful exposure estimates. One reviewer noted that s/he had originally commented that TCEQ should have used the data from Clydach after 1930. This reviewer is now satisfied about why the study was not used; however, the reviewer suggested that TCEQ provide a better explanation of why the study has no useable data.

The panel then noted that three different approaches had been proposed for how to proceed with the dose response analysis, given the lack of data on relevant mixtures:

- Use animal data for the quantitative estimate
- Use all of the epidemiology data in a meta-regression
- Pick the most relevant epidemiology studies as the basis of the quantitative estimate.

Therefore, prior to making recommendations, the panel decided to discuss the advantages and disadvantages of each approach and discuss how the results could be applied in a regulatory setting.

Use of Animal Data

The reviewer proposing this approach said that the idea was to derive separate URFs for each nickel species based on the NTP study results and then weight the URFs based on some estimate of each in Texas air. The advantage is that it will attribute risk to a specific form of nickel. The disadvantage is that it requires extrapolation from animals and the composition of nickel species in Texas air is unknown. Another reviewer asked if this approach would take into account the interactions of the various nickel species in actual mixtures of air. The first reviewer replied that the approach assumes response additivity, which may not be true. Another reviewer indicated that interactions between nickel species would be due to something other than DNA interaction, such as an inflammatory process or inhibition of DNA repair. So, it may be possible to estimate the exposure levels at which these processes become of concern, and compare them with exposure for that species in the range of the ESL, to test whether greater than additive interactions are of concern. Another reviewer suggested that using animal data seems less appropriate when the epidemiology data base is rich.

Meta-regression of Epidemiology Studies

The reviewer proposing this approach stated that all studies with adequate data should be used, with the resulting values weighted by the size of studies. In addition, it would be reasonable to do a regression, and then obtain a slope factor on the resulting regression curve. The advantage of this approach is that it allows more of the epidemiology data to be used. The disadvantage is that special knowledge is needed and the end result likely will not be numerically different from TCEQ's current draft URF. Another reviewer did not believe that a "simple" meta-analysis would be an appropriate approach for this data set because of the different speciation of nickel in different studies. However, this panel member agreed that a meta-regression could be appropriate. One reviewer asked why the outliers would be excluded. The first reviewer replied that the average of many things that aren't similar is not a good representation of the data because it does not convey the range of the data.

Another panel member asked which of the available studies could be included in the meta regression. The first reviewer indicated that most of the studies could be included because, for most of the studies, it would be possible to develop at least a point estimate of average exposure. Another reviewer asked if a disadvantage would be losing information in the low dose region of individual studies. For example, several of the studies had 4 or 5 exposure levels and some of that data will be lost if only a single point from each study is used. The first reviewer agreed that to some extent, data would be lost. However, by using the meta regression approach, one would create a dose response from the slope of the regression curve by adding including data from some studies with low doses and some with high doses.

Selected Studies Approach

The panel noted that this was the approach used by TCEQ and, in fact, this approach is a simplified version of the approach described above in that two studies were given weights of "1" and all of the other studies were given weights of "0". One reviewer concluded that this approach seemed reasonable because of TCEQ's choice to use studies with a lower exposure to sulfidic nickel. This reviewer suggested that it might be appropriate to give the Grimsrud et al. (2003) study less weight than the Enterline and Marsh study, if the other reviewers concerns regarding the level of exposure to sulfidic nickel are correct. The primary advantage of this approach is that it is already done. The disadvantage is that the exact weighting of the studies is an issue.

A reviewer suggested that for the Enterline and Marsh study, the non-refinery workers after 1946 were the group most similar to Texas air mixture. Although the data for this group is not as good as for all workers, they are sufficient for a dose-response estimate. In addition, the exposure for this group (see page 45 of the DSD for a description of subcohorts) has a lower sulfidic nickel content, which makes this group more relevant to Texas. Dr. Sielken responded that they did not use this subcohort because the dose response for this group (workers after 1946) is flat for the first three doses. TCEQ pointed to table 12, pg 48 of the Enterline and Marsh study, which shows the data for subcohorts and explained that using "all workers" provided a better estimate. One reviewer stated that the flatness of curve for the subcohorts could be interpreted as indicating that when the sulfidic nickel component of exposure decreases, the risk decreases. Another reviewer stated that the complex modeling and weighting used for deriving the URF are inappropriate given the inherent uncertainty in the epidemiology data and speciation. A complex quantitative approach adds a suggestion of sophistication and accuracy that is not borne out by the underlying data. A simpler range-finding approach would seem to be more appropriate.

One reviewer suggested that TCEQ treat the separate cohorts as separate contributors to the risk estimate and give each a weight based on similarity to Texas exposures. Another panel member agreed, noting that if TCEQ used a weighting approach based on similarity to Texas air, then they would be able to add more studies. Other panel members agreed that could be a viable approach. For example, it would allow TCEQ to use the data from the Copper Cliff study, but that study would be given a low weight because of less similarity with Texas air.

Conclusion to Critical Study Discussion

Overall, the panel reached consensus that the animal studies should not be the primary approach for quantifying risks from nickel. However, two panel members concluded that deriving quantitative estimates from animals would be useful as a test for reasonableness of the proposed URF in light of the uncertainties in the epidemiology. In contrast, three panel members concluded that the challenges to deriving a quantitative estimate from the animal studies limit the usefulness of using the animal data directly to inform the dose-

response assessment. The panel agreed that an approach based on the epidemiology studies would be appropriate. Individual panel members provided their own preferences for refinements to the degree to which various studies should be used and the relative weight each study should be given in the ultimate dose-response used by TCEQ. However, the panel members all agreed that TCEQ should improve the description of the selection criteria for choice of studies in the document. Also, the panel suggested that TCEQ conduct a sensitivity analysis to evaluate the impact of adding epidemiology studies to the URF estimate. In addition, the panel recommended that TCEQ expand the qualitative characterization of uncertainties, including the concept that risk to the nickel mixture in Texas air may be as low as 0.

Dose-Response Approaches

Given the outcome of above discussions, what dose-response approaches are appropriate?

- **Is the central estimate or the 95% UCL estimate the best estimate and why**?

Two reviewers stated that the most appropriate dose response approach is to use the upper bound on risk because it addresses uncertainty in the modeling as well as other areas that are not addressed in other ways. The reasons given by TCEQ for using a central estimate are not relevant to this analysis. One reviewer added that the conservatism introduced by using the upper bound on risk is a minor fraction of the total uncertainty in the URF. Another reviewer asked if using the upper bound makes the estimated URF too conservative, given the fact that the epidemiology data are already conservative. The first reviewer replied that using the upper bound is not overly conservative because there are many other uncertainties in the assessment and we have no information on the direction or the magnitude of the effect that those uncertainties have on the actual risk. These reviewers also suggested, however, that adding two upper bound values together would compound the conservatism and provided additional support for not combining the potency estimates derived from the two epidemiology studies selected by TCEQ.

In summary, two panel members suggested use of an upper bound estimate, two members agreed with use of the central estimate, and one had no opinion. However, all reviewers agreed that TCEQ should better describe the uncertainties in the URF, including the possible direction and magnitude of the different factors contributing to uncertainties.

- **Is the URF weighting procedure used to calculate the final URF reasonable and justified?**

This panel discussion centered on the appropriateness of combining the lung cancer incidence data from the Grimsrud study with the respiratory tract cancer mortality data from the Enterline and Marsh study. Two reviewers stated that TCEQ had made a

reasonable argument that both endpoints are sufficiently related, so it is appropriate to combine the data. In contrast, one reviewer stated that the TCEQ had to make too many assumptions to justify combining the estimates for these two outcome measures. Another panel member suggested that it is reasonable to combine mortality and incidence data from the two studies, given that respiratory cancer is generally a rapidly fatal disease. However, cancer incidence and mortality data should not be combined within the same study, because these are dependent outcomes.

In summary, two reviewers thought the approach used by TCEQ was reasonable but two reviewers suggested that TCEQ should consider the two data sets separately, and choose the one that gives the highest URF. However, the panel agreed that the document needs to expand the discussion of the uncertainties in the approach. In addition, the weighting techniques are overly precise given the overall uncertainties.

General recommendations for the Cancer Assessment:

The panel recommended that the document should also characterize uncertainties associated with exposure. In addition, the panel suggested that TCEQ should expand the discussion of the epidemiology studies and animal studies so that the document gives a fuller picture of the available data. Finally, the inclusion criteria for studies used for calculation of the URF should be better discussed.

Health-Based Acute ReV and ^{acute}ESL

Overall, the panel agreed with the choice of Cirla et al, 1985 as the critical study. However, Reviewer #2 suggested also using the Graham et al. (1978) study as a co-critical study. There was panel consensus on the choice of uncertainty factors, except for the values selected for the UF_L and UF_D. Panel members made several different suggestions for appropriate choice of value for these two factors.

> **Should the Graham mouse study be designated a "co-principle" study since data needed for the MPPD model were not available? How would it made a difference in the final ReV value that was chosen?**

One reviewer noted that Cirla study was not ideal because it only included a single exposure level. However, this reviewer was equally concerned about the Graham study, because values for MMAD and sigma were not presented. This lack of data reduces the credibility of the results from Graham. This reviewer suggested that a sensitivity analysis would be needed to evaluate the impact of choosing different MMAD and sigma values, and so would not favor considering Graham as a co-critical study.

The panel then discussed the sensitivity of the population in the Cirla study, and whether the observations in the studied population were the result of sensitization from prior

nickel exposure or an inherent underlying condition. The panel did not reach any conclusions on this issue, but recommended that TCEQ expand the discussion on the relative sensitivity of this population. Reviewers discussed the appropriate interpretation of the Cirla in light of the previous exposure of the study group to nickel and resulting sensitization. A reviewer noted that the Texas ESL guidelines indicate that the ESL values are not designed to protect people who were previously sensitized. Another reviewer asked if there were other, more appropriate acute studies in humans. The only other available study is a 12-day study in rats. Although this exposure duration is too long for what Texas needs, that study can provide some useful perspective on concentration-response, and should not be ignored.

Overall the panel agreed with the use of the Cirla study as the critical study. The panel also concluded that adding the Graham study as a co-critical study would not provide significant additional relevant information.

Upon review of all opinions and rationales, what is the reviewers' consensus on the most appropriate value for the LOAEL-to-NOAEL UF and the database UF ?

One panel member stated that the uncertainty factor to account for LOAEL-to-NOAEL extrapolation is used to address two different components – incidence and severity. To evaluate severity, one should evaluate the magnitude of decrease in FEV_1 from the Cirla study. A 15% decrease in FEV_1 is typically considered to be a mild effect. Another reviewer agreed that a 15% change is not a severe effect. Therefore, a 3 should be adequate. In addition, this reviewer was concerned about the significance of this effect in a sensitized population. Since this is a sensitized population, it is likely that they will respond at a lower concentration than a non-sensitized population. In addition, there is no information about the nickel concentration that resulted in sensitization. Therefore, this reviewer considered that a value of 3 or 1 would be appropriate for the LOAEL uncertainty factor.

Other members of the panel noted that there were database limitations and study limitations that warranted an increase the database uncertainty factor. For example, the fact that there is only 1 exposure level means that nothing is known about the shape of the dose response.

Although individual panel members made recommendations to increase or decrease individual factors based on the arguments noted above, the panel reached consensus that the composite uncertainty factor of 30 is adequate. The panel suggested that TCEQ add more discussion about the contributions of the available animal data to the limited database and add more description of considerations that might increase or decrease the selected values.

Health-Based Chronic ReV and $^{chronic}ESL_{noncancer}$

As discussed for the cancer assessment, TCEQ must derive a chronic, non-cancer ReV and ESL based on total nickel. For the non-cancer assessment, TCEQ selected the most toxic form of nickel as a conservative surrogate to derive a conservative ESL. Overall, the reviewers agreed with this approach, and agreed with the choices of critical study, critical effect, dosimetry, and uncertainty factors. There was some disagreement on the choice of nickel sulfate as the surrogate compound, and Reviewer #1 suggested deriving nickel species-specific ESLs.

> **Which form of nickel should be selected for deriving the ReV/ESL appropriate given the purpose of these values?**
>
> **If nickel species-specific ESLs were derived from the NTP study, how should they be applied to a nickel mixture?**

One reviewer noted that the same issues discussed for the cancer assessment with regard to discussion of the animal data also apply to the noncancer assessment. This reviewer stated that if TCEQ has decided to choose a single representative nickel species, then nickel sulfate is the most appropriate surrogate for noncancer effects. The panel agreed with this conclusion.

Another reviewer recommended that TCEQ provide better documentation on why it decided to not use a BMD approach for the chronic ESL. Other reviewers agreed and suggested that TCEQ expand the discussion of the modeling, moving the discussion from the appendices to the body of the document, and add a discussion of the diagnostics used.

Overall, the panel members agreed with the UF choices presented in the document based on the information cited. However, regarding the choice of uncertainty factors for animal to human differences, one reviewer agreed with the submitted public comment that human data (Berge and Skyberg, 2003) support the conclusion that humans are less sensitive than animals. Therefore, this reviewer concluded that it would be reasonable to decrease the uncertainty factor to account for extrapolating from an animal study. Other panel members agreed that TCEQ should reevaluate the factor on animal to human differences in toxicodynamics in light of these data.

This page intentionally left blank.

References

Anttila, A.; Pukkala, A.; Aitio, A. (1998). Update of cancer incidence among workers at a copper/nickel smelter and nickel refinery. Int Arch Occup Environ Health 71: 245-250.

Arena, V.C: Sussman, N.B: Redmond, C.K: Costantino, J.P: Trauth, J.M. (1998). Using alternative comparison populations to assess occupation-related mortality risk. J Occup Environ Med 40: 907-916.

Agency for Toxic Substances and Disease Registry (ATSDR). (2005). Toxicological profile for Nickel. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service. Available online at: http://www.atsdr.cdc.gov/toxprofiles/tp15.html.

Berge, S; Skyberg, K. (2003). Radiographic evidence of pulmonary fibrosis and possible etiologic factors at a nickel refinery in Norway. J Environ Monit 5:681-688.

Circla, A.M: Bernabeo, F: Ottoboni, F: Ratti, R. (1985). Nickel induced occupational asthma: Immunological and clinical aspects. IN: Progress in Nickel Toxicology. Eds. Brown S, Sunderman FW. International Union of Pure and Applied Chemistry, Blackwell Scientific Publications.

Doll, R: Andersen, A: Cooper, W.C: et al. (1990). Report of the International Committee on nickel carcinogenesis in man. Scand J Work Environ Health 16:1-82.

Enterline, P: Marsh, G. (1982). Mortality among workers in a nickel refinery and alloy manufacturing plant in West Virginia. J Nat Cancer Inst 68(6):925-33.

Fernandez-Nieto, M: Quirce, S: Carnes, J: Sastre, J. (2006). Occupational asthma due to chromium and nickel salts. Int Arch Occup Environ Health 79:483-486.

Galbreath, K.C.; Crocker, C.R.; Nyberg, C.M.; Huggins, F.E.; Huffman, G.P.; Larson, K.P. (2003). Nickel speciation measurements of urban particulate matter: method evaluation and relevance to risk assessment. J Environ Monit 5: 56N-61N.

Goodman, J.E: Prueitt, R.L: Dodge, D.G: Thakali, S. (2009). Carcinogenicity assessment of water-soluble nickel compounds. Gradient Corporation. 20 University Rd, Cambridge, MA.

Graham, J.A: Miller, F.J: Daniels, M.J: Payne, E.A: Gardner, D.E. (1978). Influence of cadmium, nickel, and chromium on primary immunity in mice. Environ Res 16:77-87.

Greim, H; Ziegler-Skylakakis, K. (2007). Risk assessment for biopersistent granular particles. Inhal Toxicol.;19 Suppl 1:199-204; ILSI Risk Science Institute. (2000). The relevance of the rat lung response to particle overload for human risk assessment: A workshop consensus report. Inhalation Toxicology 12:1-17.

Grimsrund, T.K: Berge, S.R: Resmann, F: et al. (2000). Assessment of historical exposures in a nickel refinery in Norway. Scand J Work Environ Health 26(4):338-345.

Grimsrud, T: Berge, S: Haldorsen, T: Andersen, A. (2002). Exposure to different forms of nickel and risk of lung cancer. Am J Epidemiol 156(12):1123-1132.

Grimsrud, T.K: Berge, S.R: Martinsen, J.I: Andersen, A. (2003). Lung cancer incidence among Norwegian nickel-refinery workers 1953-2000. JEM 5:190-197.

Karjalainen, S.; Kerttula, R.; Pukkala, E. (1992). Cancer risk among workers at a coppper/nickel smelter and nickel refinery in Finland. Int Arch Occup Health 63: 547-551.

Magnus, K: Andersen, A: Hogetveit, A.C. (1982). Cancer of respiratory organs among workers at a nickel refinery in Norway. International Journal of Cancer 30(6):681-685.

Mauderly, J.L. (1997). Relevance of Particle-Induced Rat Lung Tumors for Assessing Lung Carcinogenic Hazard and Human Lung Cancer Risk. Environmental Health Perspectives 105:1337-1346.

Nikula, K.J: Vallyathan, V: Green, F.H: Hahn, F.F. (2001). Influence of exposure concentration or dose on the distribution of particulate material in rat and human lungs. Environ Health Perspect 109(4):311-8.

National Toxicology Program (NTP). (1996). NTP technical report on the toxicology and carcinogenesis studies of nickel sulfate hexahydrate (CAS no. 10101-97-0) in F344/N rats and B6C3F1s mice (inhalation studies). In NIH publication, National Toxicology Program technical report series; no. 454.

Oberdörster, G. (1995). Lung particle overload: Implications for occupational exposures to particles. Regulatory Toxicology and Pharmacology 27:123-135.

Oller, A.R; Kirkpatrick, D.T; Radovsky, A; Bates, H.K. (2008). Inhalation carcinogegnicity study with nickel metal powder in Wistar rats. Toxicol Appl Pharmacol 233(2):262-75.

Redmond, C.K: LaGasse, A.A: Bass, G. (1983). Cancer mortality in workers in the high nickel alloys industry. Unpublished Report. University of Pittsburgh. Submitted to the U.S. Environmental Protection Agency, Central Docket Section, Washington, DC, Docket No. ECAO-HA-81-1 IIA.E.4.

Redmond, C.K: Sussman, N.B: Arena, V.C: Constantino, J.P. (1996). Supplemental Analysis of High Nickel Alloy Workers. Final Report to NiPERA. July 26, 1996. Available upon request.

Schaumloffel, D. (2005). Speciation of Nickel. In Handbook of Element Speciation II, Species in the Environment, Food, Medicine and Occupational Health, edited by R. Cornelis, J. Caruso, H. Crews and K. Heumann: John Wiley & Sons, Ltd.

Seilkop, S.K: Oller, A.R. (2003). Respiratory cancer risks associated with low-level nickel exposure: an integrated assessment based on animal, epidemiological, and mechanistic data. Regulatory Toxicology and Pharmacology 37:173-190.

Serita, F. (1999). An animal exposure system using ultrasonic nebulizer that generates well controlled aerosols from liquids. Industrial Health 37:82-87.

Shihara, Y: Kyono, H: Serita, F: Toya, T: Kawashima, H: Miyasaka, M. (2002). Inflammatory responses and mucus secretion in rats with acute bronchiolitis induced by nickel chloride. Inhalation Toxicology 14:417-430.

Sivulka, D.J.; Seilkop, S.K. (2009). Reconstruction of historical exposures in the US nickel alloy industry and the implications for carcinogenicity hazard and risk assessment. Reg Tox and Pharmacol 53:174-185.

Texas Commission on Environmental Quality (TCEQ). (2006). Guidelines to Develop Effects Screening Levels, Reference Values, and Unit Risk Factors. RG-442, Toxicology Section, Chief Engineer's Office. Available at: www.tceq.state.tx.us/publications.

United States Environmental Protection Agency (U.S. EPA). (1986). Health assessment document for nickel and nickel compounds: final report. Research Triangle Park, NC: U.S. Environmental Protection Agency, Office of Research and Development, Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office.

This page intentionally left blank.

Appendix A - Peer Review Charge

Scientific Peer-Review of the
Noncarcinogenic and Carcinogenic Sections of the
Nickel Development Support Document
Charge to Peer Reviewers

The Toxicology Division of the Chief Engineer's Office, Texas Commission on Environmental Quality (TCEQ) has prepared a draft Development Support Document (DSD) that outlines the hazard assessment and dose-response processes used to derive Effects Screening Levels (ESLs), Reference Values (ReV), and a Unit Risk Factor (URF) for nickel. The toxicity values were developed using RG-442 *Guidelines to Develop Effects Screening Levels, Reference Values, and Unit Risk Factors* (TCEQ 2006). ESLs are chemical-specific air concentrations set to protect human health and welfare. Short-term ESLs are based on data concerning acute health effects, odor potential, and vegetative effects, while long-term ESLs are generally based on data concerning chronic noncarcinogenic and/or carcinogenic health effects. ESLs are used in the evaluation of air permit applications as well as proposed rules and regulation (e.g. Permits by Rule). ReVs and URFs, used as the basis of ESLs, are used in the evaluation of air monitoring data and in the development of Protective Concentration Levels for remediation sites.

We are asking you to provide a review of the scientific approaches used by TCEQ in developing the toxicity values that are described in this draft document. The DSD is a summary document and does not provide a detailed description of every aspect of the toxicity assessment for a chemical. References to appropriate papers or documents are provided if more detailed information is needed.

There are a number of policy decisions the TCEQ has made and included in this assessment that the agency does not seek comment on. These risk management goals were approved by the Commissioners and Executive Director of the TCEQ and are consistent with other TCEQ programs. ESLs developed in accordance with these no significant risk levels are intended to prevent adverse effects potentially associated with cumulative and aggregate exposures as defined in Section 1.2 of RG-442 ESL Guidelines (TCEQ 2006). Therefore, please do not spend your time commenting on the following policy decisions:

- The use of a hazard quotient (HQ) of 1 for assessing an individual chemical with a nonlinear dose-response assessment for the ReV (e.g., noncarcinogenic ReVs for nickel).
- The use of a target excess cancer risk level of 1 in 100,000 (1E-05) for the nickel carcinogenic-based ESL.
- In consideration of cumulative and aggregate exposure, the use of an HQ of 0.3 to calculate short-term and long-term health-based ESLs for chemicals with a nonlinear (threshold) dose-response assessment (i.e., health-based ESLs = 0.3 x ReV).
- Assumption of a lifetime exposure of 70 years.

General Issues

Please consider all aspects of the nickel DSD and evaluate strengths and weaknesses of the procedures used to develop acute and chronic toxicity values based on the specific questions described below. Where possible, try to put the strengths and weaknesses in perspective by indicating their relative magnitude. Please try to avoid emphasizing minor technical details or making tutorial comments. Reviewers should identify scientific uncertainties and suggest ways to reduce or eliminate those uncertainties.

- Were procedures outlined in the ESL Guidelines followed by the TCEQ to perform nickel's toxicity assessment? If references to accepted procedures in federal, state, or other appropriate guidance documents were made in the ESL Guidelines, were those accepted procedures followed?
- Does the nickel DSD clearly describe the approaches used by TCEQ to perform the toxicity assessment (i.e., hazard identification and dose-response assessment)?
- Please identify any relevant studies or data that have not been cited. Explain how they may impact the assessment.

Health-Based Acute ReV and ^{acute}ESL

Chapter 3 of the nickel DSD describes the approaches used to derive the health-based acute ReV and acute ESL. The key decision points are listed below. For each decision, please consider the scientific defensibility of the decision and any additional approaches or analyses, or additional information that could improve that decision. Please indicate if there are other issues specific to developing acute toxicity factors that have not been adequately addressed in the document.

- The choice of the critical study (Cirla et al. 1985):
  - Was the human study selected for the non-cancer estimates the most appropriate study? Was the form of nickel selected (nickel sulfate) for deriving the ReV/ESL appropriate given the purpose of these values?
- The choice of critical effect (greater than a 15% decrease in FEV_1 in six of seven asthmatics):
  - Was the most appropriate critical effect selected? If not, what would be a more appropriate critical effect?
  - Is the endpoint relevant for human risk assessment?
- The choice of point of departure.
- The choice of dosimetric adjustments:
  - Was the most relevant, appropriate, and defensible dose metric selected?
  - Were the appropriate default exposure duration adjustments conducted?
  - For the supporting animal study (Graham et al. 1978), were the appropriate default dosimetry adjustments from animal-to-human exposure conducted? Specifically, were appropriate estimates (i.e. mass median aerodynamic diameter (MMAD) and geometric particle size distribution (σ_g)) for conducting the regional deposition dose ratio

(RDDR) chosen when the supporting study did not report the required parameters?
    - If the dosimetry adjustments were not acceptable, what would be a more appropriate dosimetric adjustment?
- The choice of uncertainty factors:
    - Have all of the appropriate uncertainty factors been considered and are the values assigned to the uncertainty factors clearly justified and defensible?
    - Would you make recommendations for a different approach to select uncertainty factors to calculate the acute ReV?

Health-Based Chronic ReV and $^{chronic}ESL_{noncancer}$

Chapter 4 of the nickel DSD describes the approaches used to derive the health-based chronic ReV and chronic ESL for noncancer endpoints. The key decision points are listed below. For each decision, please consider the scientific defensibility of the decision and any additional approaches or analyses, or additional information that could improve that decision. Please indicate if there are other issues specific to developing chronic toxicity factors that have not been adequately addressed in the document.

- The choice of the critical study (NTP 1996):
    - Was the animal study selected for the non-cancer estimates the most appropriate study? Was the form of nickel selected (nickel sulfate) for deriving the ReV/ESL appropriate given the purpose of these values?
- The choice of critical effect:
    - Was the most appropriate critical effect selected (lung fibrosis and chronic inflammation)? If not, what would be a more appropriate critical effect?
    - Is the endpoint relevant for human risk assessment?
- The choice of dosimetric adjustments:
    - Was the most relevant, appropriate, and defensible dose metric selected?
    - Were the appropriate default dosimetry adjustments from animal-to-human exposure conducted? Specifically, was the Multiple Pass Particle Dosimetry (MPPD) Model used appropriately and is the (RDDR) appropriate? Were the parameters used scientifically defensible?
    - If the dosimetry adjustments were not acceptable, what would be a more appropriate dosimetric adjustment?
- The choice of point of departure.
- The choice of uncertainty factors.
    - Have all of the appropriate uncertainty factors been considered and are the values assigned to the uncertainty factors clearly justified and defensible?
    - Would you make recommendations for a different approach to select uncertainty factors to calculate the chronic ReV?

Welfare-Based Acute and Chronic ESLs

The TD did not find any data to allow the derivation of welfare-based acute or chronic ESLs. Please indicate if there are other issues specific to developing welfare-based ESLs that have not been adequately addressed in the document.

Cancer Weight of Evidence and Unit Risk Factor (URF)

The nickel DSD describes the approaches used to evaluate carcinogenicity and derive the URF and chronic ESL for cancer. Please review the key decisions made by TCEQ in deriving these values. For each decision, please comment on the consistency of the decision with TCEQ's ESL guidelines, the scientific appropriateness of the decision, and any additional approaches or additional information that would improve that decision. The key decisions and some specific issues to consider are listed below. Please discuss other issues specific to developing URFs for carcinogenic effects that have not been adequately addressed in the document.

- Was the proper weight of evidence (WOE) classification using the new USEPA carcinogenic guidelines given to nickel compounds? If not, what WOE classification should be given to nickel compounds, specifically metallic nickel?
- The cancer assessment relied upon human epidemiological studies. There are also animal studies; were the animal data used appropriately to support the weight of evidence conclusions?
- Is the epidemiological evidence in Grimsrud et al. (2003) and Enterline and Marsh (1982) properly used in the characterization of chronic cancer risks? Is use of these two studies for calculating URFs justified?
- Were the reasons for not using the following epidemiological studies to develop URFs clearly described and justified: Copper Cliff, Ontario (Chovil et al. 1981) and Clydach, Wales (Peto et al. 1984)?
- Were the statistical and modeling approaches used for calculating URFs appropriate?
- Was the dose metric selected (mg/m^3-months or -years) the most relevant and appropriate choice?
- Is use of total nickel for both studies, and all workers for Enterline and Marsh (1982), justified given the purpose of the URF and carcinogenic ESL and in light of the recent work by Goodman et al. (2009)?
- Are the most appropriate URFs from each study used to calculate the final URF?
- Is use of the central estimate URFs justified for reasons discussed in the DSD?
- Was cancer endpoint selected as the basis of the potency estimates (lung cancer) the most appropriate and relevant choice?
- Are respiratory cancer data from Enterline and Marsh (1982) a reasonable surrogate for lung cancer for reasons discussed in the DSD?
- Are lung cancer incidence and mortality sufficiently similar as to be comparable for purposes of this assessment for reasons discussed in the DSD?
- Is the URF weighting procedure used to calculate the final URF reasonable and justified?
- Is the decision not to apply age-dependent adjustment factors (ADAFs) to the

URF to account for potential increased sensitivity of children justified and properly considered?

Appendix B – Conference Call Charge

Conference Call Charge on the Peer Review of the Development Support Document for Nickel

The purpose of this conference call is to allow the panel members the opportunity for discussion and consensus on issues that remained unresolved in the written comments and to provide TCEQ an opportunity to ask clarifying questions regarding the reviewers written comments. Ideally, the outcome of the conference call will give TCEQ clear feedback on additional options for developing the ESL values that TCEQ is considering in response to the individual panel members' written comments.

As an aid to the panel, clarifications on methodology have been identified by the Toxicology Division (TD) of the TCEQ that respond to some of the panel comments.

- First, TD relies on toxicity assessments conducted by other federal, state, and international agencies that have undergone a peer-review process as a starting point in their toxicity assessments, because of time and resource constraints. However, the TD obtains copies of key studies and supporting studies and critically reviews these studies. TD does not routinely adopt toxicity values developed by other organizations. The toxicity assessments conducted by others are critically reviewed. These toxicity values may be adopted if procedures outlined in the TCEQ ESL Guidelines document are followed. Because the DSD is a summary document, the procedures discussed above are not included in the DSD.

- Second, the legislature dictates that the TD develop acute (usually 1-hour averaging time) and chronic values to evaluate all chemicals. These values are used to evaluate emissions from facilities during the air permit review process and to evaluate ambient air monitoring data, which are reported as total nickel. No nickel speciation data are routinely available. Other federal or state agencies may decide not to develop values, but the TD must have procedures and comparison values in place to evaluate air emissions of all chemicals. Since the statutes require that ESL values be developed, incomplete toxicity databases present a complex challenge for the TD.

- Third, the toxicity values will be used to evaluate emissions from facilities during the air permit review process, for evaluation of ambient air monitoring data, and as toxicity factors in the Texas Risk Reduction Program remediation program. Because the nickel concentrations in air from these processes are measured as total nickel, not as individual nickel species, both the non-cancer and cancer ESLs derived for nickel must be health protective for total nickel without being overly-conservative.

Cancer Weight of Evidence and Unit Risk Factor (URF)

Hazard Characterization of Nickel Mixtures

The toxicity values generated in the Nickel DSD will be used to evaluate emissions from facilities during the air permit review process, for evaluation of ambient air monitoring data, and as toxicity factors in the Texas Risk Reduction Program remediation program. Because TCEQ typically evaluates total nickel and does not have nickel compound-specific air data, TCEQ used a mixtures approach for the derivation of URFs. The challenge is to select an appropriate mixture, which will be health protective without being unduly conservative, on which to base the cancer assessment.

1. Given the wide spectrum of toxicity of the different nickel species, the needs that the TCEQ faces in dealing with mixtures, and the fact that the TCEQ only receives total nickel data, is using a mixtures approach appropriate for a nickel cancer assessment?

2. Does the DSD adequately bound the hazard posed by nickel mixtures in air?

3. Does the DSD accurately and adequately characterize the nature and/or composition of the nickel mixtures likely to be found in Texas air so that a similar "reference mixture" can be identified?

4. Generally, considering both animal and human studies, has the nature and/or composition of the nickel mixture to which exposure occurred been identified sufficiently to allow TCEQ to identify those studies that had a nickel mixture similar to Texas air?

Weight of Evidence

There was no consensus among the panel members on the appropriateness of the weight of evidence statement. In particular, there was disagreement on whether it was appropriate to assess WOE for nickel compounds as a group compared to WOE for the individual nickel species, or whether it is feasible to assess WOE for the mixture of nickel species that is typical of Texas air. In addition, Reviewer #1 indicated that there is not sufficient evidence for nickel sulfate alone, but only in the presence of other nickel compounds.

5. If a mixtures approach to nickel assessment is appropriate, would a WOE for mixtures of nickel be defensible?

6. If so, how should data for the different nickel species be taken into account and how should the WOE narrative address the different forms of nickel likely to be present in the mixture?

Selection of Critical Study

The reviewer comments expressed a broad range of opinions on which studies would be appropriate as the critical study. Reviewer recommendations included using the animal studies to derive the URF, adding some of the other human studies in addition to Grimsrud and Enterline and Marsh to the analysis, and using all of the human studies to derive the URF via a meta-analysis. Specifically, Reviewer #1 suggested not using the Enterline and Marsh study because there were no statistically significantly increased risks. Reviewer #2 suggested not using the Kristiansand study due uncertainties in the exact nature of the exposure. Reviewer #3 suggested conducting a meta-analysis using data from all human studies with suitable exposure estimates. For the URF based on the Enterline and Marsh study, the reviewers' comments were divided on the appropriateness of using total nickel and all workers. Two reviewers commented that only insoluable forms of nickel are carcinogenic and so only exposure estimates from these species should be used. The remaining reviewers suggested that the approach was appropriate and conservative given the uncertainties in the literature regarding the carcinogenic potential the various species that make up soluble and insoluble nickel compounds.

7. Given the issues discussed above for hazard characterization, and the preference of TCEQ for human data, which studies are considered superior by the reviewers for deriving a somewhat conservative generalization of risk to the population of Texas and why?

8. What specific URF analyses do the reviewers suggest as more applicable for the evaluation of total nickel in ambient air data? What epidemiology studies would be appropriate for TCEQ to use to develop a URF? The following studies are suggestions from either reviewers or the public for alternate studies other than Enterline and Marsh and Grimsrud:

- The Clydach cohort after 1930
- Harjavalta, Finland (Karjalainen et al., 1992 and Anttila et al. 1998)
- Public Comments (Nickel Producers Environmental Research Association) : Arena et al. (1998) and Sivulka and Seilkop (2009)

9. If nickel compound-specific URFs were derived based on animal data, how should they be applied to a nickel mixture so as to not grossly over- or under-estimate risk?

10. Given the cohort in Enterline and Marsh 1982 appears to have been exposed to a nickel mixture appropriate for Texas air, is it reasonable to exclude this study from use in URF derivation because it does not have statistical significance? Could the Enterline and Marsh study be used as a supporting study and how?

11. What is the panel's opinion of the Grimsrud study given that is appears to contradict findings in animals regarding carcinogenicity of soluble nickel compounds?

Dose-Response Approaches

12. Given the outcome of above discussions, what dose-response approaches are appropriate?

- Is the central estimate or the 95% UCL estimate the best estimate and why?

- Is the URF weighting procedure used to calculate the final URF reasonable and justified?

Health-Based Acute ReV and ^{acute}ESL

Overall, the panel agreed with the choice of Cirla et al, 1985 as the critical study. However, Reviewer #2 suggested also using the Graham study as a co-critical study. There was panel consensus on the choice of uncertainty factors, except for the values selected for the UF_L and UF_D. Panel members made several different suggestions for appropriate choice of value for these two factors.

13. Should the Graham mouse study be designated a "co-principle" study since data needed for the MPPD model were not available? How would it made a difference in the final ReV value that was chosen?

14. Upon review of all opinions and rationales, what is the reviewers' consensus on the most appropriate value for the LOAEL-to-NOAEL UF and the database UF ?

Health-Based Chronic ReV and $^{chronic}ESL_{noncancer}$

As discussed for the cancer assessment, TCEQ must derive a chronic, non-cancer ReV and ESL based on total nickel. For the non-cancer assessment, TCEQ selected the most toxic form of nickel as a conservative surrogate to derive a conservative ESL. Overall, the reviewers agreed with this approach, and agreed with the choices of critical study, critical effect, dosimetry, and uncertainty factors. There was some disagreement on the choice of nickel sulfate as the surrogate compound, and Reviewer #1 suggested deriving nickel species-specific ESLs.

15. Which form of nickel should be selected for deriving the ReV/ESL appropriate given the purpose of these values?

16. If nickel species-specific ESLs were derived from the NTP study, how should they be applied to a nickel mixture?

Appendix C – Public Comments

Comments of the
Nickel Producers Environmental Research Association
on the
Nickel and Inorganic Nickel Compounds Development
Support Document (DSD) prepared by the TCEQ

These comments focus on a few issues not already brought up by the peer-reviewers of this document.

Speciation of air nickel exposures

The TCQS nickel Development Support Document (DSD) does not provide any measured data on the speciation of nickel in ambient air in Texas. Rather, it relies on personal communications regarding the kind of processes that are present in Texas and the expected emissions to ambient air based on the nature of those processes. Although this approach can be acceptable as a first approximation, it is only a first approximation. For example, some activities like grinding nickel metal or alloys in massive forms can produce very large particles (visible dusts). These particles (containing metallic nickel) will be fairly course for the most part and not contribute to PM10 or PM2.5. The small particles that contribute to Texas ambient air could have a completely different composition. The enclosed papers by Galbreath et al. report on nickel in ambient air measurements in Florida and provide concrete speciation data on these exposures. Galbreath et al. report that complex nickel oxides and nickel sulfate are the predominant forms of nickel in ambient air, with very small amounts of nickel sulfide (not nickel subsulfide) present and no metallic nickel. It may be prudent to take these published data into consideration in the DSD report.

Mode of action and carcinogenicity tumor sites

The DSD does not cite the Heim et al (2007) study (listed *as Reference not cited*). Yet, this study is important for two reasons: 1) it confirms that nickel cannot cause tumors at sites other than the respiratory tract, and 2) it adds to the WOE evaluation of the carcinogenicity of soluble nickel compounds. In the Heim et al. study, rats exposed by gavage to nickel sulfate hexahydrate did not demonstrate increased incidence of tumors even though blood nickel levels were several hundred-fold higher than in control rats. This also indicates that it is unlikely that cell membrane-mediated effects of nickel ions (e.g., HIF-mediated effects) can result in tumor induction or promotion of naturally occurring tumors. Rather, this supports the premise that Ni ions have to be present in the nucleus of target respiratory cells to see any tumors and that Ni ions (from water soluble nickel compounds) do not have an efficient way to get to the nucleus of the cells.

Health Based Chronic ReV

The use of an Uncertainty factor of 3 to account for toxicodynamic differences in response between rats (assumed to be less sensitive) and humans (assumed to be more sensitive) for respiratory toxicity/inflammation effects is likely to be overly conservative.

Rats are known to be more (not less) sensitive to the toxicity effects of particulates than mice and primates (Oberdorster, 1995; Mauderly, 1997; ILSI, 2000; Nikula, 2001; Greim and Ziegler-Skylakakis, 2007).

Selection of Studies Used in Determination of URF

As mentioned by some reviewers, the most thorough approach to URF determination would be to use a combination of derivations based on animal studies with single pure exposures to various nickel compounds expected to be present in Texas air and key epidemiological studies of cohorts with exposures that closely match the composition of Texas air. Since increased cancer risks have not been observed outside workers refining sulfidic nickel mattes (who had mixed and complex exposures), the use of any of these cohorts to represent ambient air composition will overestimate risks. We concur that it is appropriate to also include non refinery cohorts. In this regards, the Arena et al. (1998) study together with the recent Sivulka and Seilkop (2009) study should be considered for incorporation into the derivation of an URF for the following reasons: 1) the Enterline and Marsh cohort (West Virginia) is just one of the 13 cohorts included in the Redmond et al. (1983; 1996) and Arena et al. (1998) studies, 2) improved information on exposures for these cohorts is now available through the work of Sivulka and Seilkop (2009), indicating that exposures are mostly to oxidic nickel with some metallic nickel exposures, and 3) the cohort is very large (31,000 workers). The Arena study and the earlier Redmond study are cited in the text but are not included in the Reference list. The references cited here are provided below.

Arena, V. C.; Sussman, N.B.; Redmond, C.K.; Costantino, J.P.; Trauth, J. M. (1998). Using alternative comparison populations to assess occupation-related mortality risk. J. Occup. Environ. Med. 40:907-916.

Greim, H.; Ziegler-Skylakakis, K. (2007). Risk assessment for biopersistent granular particles. Inhal Toxicol 19 Suppl 1:199-204; ILSI Risk Science Institute. (2000). The relevance of the rat lung response to particle overload for human risk assessment: A workshop consensus report. Inhalation Toxicology 12, 1-17.

Mauderly, J.L. (1997). Relevance of Particle-Induced Rat Lung Tumors for Assessing Lung Carcinogenic Hazard and Human Lung Cancer Risk. Environmental Health Perspectives 105:1337-1346.

Nikula, K.J.; Vallyathan, V.; Green, F.H.; Hahn, F.F. (2001). Influence of exposure concentration or dose on the distribution of particulate material in rat and human lungs. Environ Health Perspect 109(4):311-8.

Oberdörster, G. (1995). Lung particle overload: Implications for occupational exposures to particles. Regulatory Toxicology and Pharmacology 27:123-135.

Redmond, C.K.; LaGasse, A.A.; Bass, G. (1983). Cancer mortality in workers in the high nickel alloys industry. Unpublished Report. University of Pittsburgh. Submitted to the U.S. Environmental Protection Agency, Central Docket Section, Washington, D.C., Docket No. ECAO-HA-81-1 IIA.E.4.

Redmond, C.K.; Sussman, N.B.; Arena, V.C.; Constantino, J.P. (1996). Supplemental Analysis of High Nickel Alloy Workers. Final Report to NiPERA. July 26, 1996. Available upon request.

Sivulka, D.J.; Seilkop, S.K. (2009). Reconstruction of historical exposures in the US nickel alloy industry and the implications for carcinogenicity hazard and risk assessment. Reg Tox and Pharmacol 53:174-185.

Appendix D – Panel Information

Conflict of Interest

An essential part of an independent expert review is the identification of conflicts of interest and biases that might interfere with a candidate's objectivity and be reason to disqualify a candidate, as well as the identification of situations which may appear to be a conflict or bias. *TERA* was selected by TCEQ to independently organize and conduct this expert panel review and is solely responsible for the selection of the panel. TCEQ has had no influence on the selection of the panel or implementation of the process. Prior to being selected to conduct this expert review, *TERA* provided information to TCEQ regarding its past and current relevant work, in order to assure *TERA*'s corporate independence to organize and conduct this review for TCEQ. *TERA* has experience in risk assessment and toxicity of nickel from project work that has been done for a variety of public and private sponsors in the past. One ongoing project is described below under the disclosure for Dr. Haber. *TERA* has not participated in the development or preparation of the document that is the subject of this meeting. *TERA* has an ongoing contract with TCEQ to organize peer reviews and is being paid for its level of effort from funds in this contract. After the initial COI check and report to TCEQ, *TERA* was approached by NiPERA (the Nickel Producers Environmental Research Association) to organize a workshop related to issues of the carcinogenicity and bioavailability of nickel metal. *TERA* did not think assisting NiPERA with this workshop would be a conflict of interest because there are no nickel producers in the State of Texas and there is no financial relationship between NiPERA and any of the parties that could be affected by the nickel ESL. In addition, the purpose of the workshop is to provide a forum for sharing of information and opinions regarding the toxicity of nickel, which will aid in future assessment efforts by both governmental agencies and private parties. Prior to agreeing to conduct the work for NiPERA, *TERA* discussed the potential NiPERA works with TCEQ and TCEQ did not see any COI issues.

The evaluation of real and perceived bias or conflict of interest is an important consideration in panel selection to ensure that the public and others can have confidence that the peer reviewers do not have financial or other interests that would interfere with their ability to carry out their duties objectively. *TERA* follows the U.S. National Academy of Sciences (NAS) guidance on selection of panel members to create panels that have a balance of scientific viewpoints on the issues to be discussed. As a result, the expert panels have a broad and diverse range of knowledge, experience, and perspective, including diversity of scientific expertise and affiliation. Panel members serve as *individuals*, representing their own personal scientific opinions. They do not serve as representatives of their companies, agencies, funding organizations, or other entities with which they are associated. Their opinions should not be construed to represent the opinions of their employers or those with whom they are affiliated.

Prior to selection, the candidates completed a questionnaire, which *TERA* used to determine whether their activities, financial holdings, or affiliations could pose a real or perceived conflict of interest or bias. *TERA* asked each promising candidate to report on his or her financial and other relationships with TCEQ, and with Texas companies that have reported releases of nickel on the Toxic Release Inventory. The completed

questionnaires were reviewed by *TERA* staff and discussed further with panel candidates as needed. (See www.tera.org/peer/COI.html for *TERA* conflict of interest and bias policy and procedures for panelist selection.)

TERA has determined that the selected panel members have no conflicts of interest and are able to objectively participate in this peer consultation. None of the panel members has a financial or other interest that would interfere with his or her abilities to objectively participate on the panel. None of the panel members is employed by TCEQ, or Texas companies releasing nickel. Nor do the panel members have any financial interests in these organizations or in the outcome of the review. None of the panel members was involved in the preparation of the document.

A brief biographical sketch of each panel member is provided below. To promote transparency, a short statement describing situations which might appear to present a conflict of interest or bias are included, as appropriate.

Biographical Sketches of Panel Members

Mr. Bruce Allen. Mr. Allen has 27 years of experience related to human and environmental health and safety. He has expertise as a biomathematician involved in risk assessment, modeling, statistical analysis, and clinical trials, having worked for a variety of government and private clients. Mr. Allen's primary interest is in the quantitative aspects of risk analysis, reflecting his experience with dose-response analysis; with the statistical appraisal of data, models, and modeling results; and with developing rigorous approaches to decision making in risk assessment contexts. His expertise in dose-response analysis extends to the modeling, including biologically motivated modeling, of cancer, noncancer, and genotoxic endpoints as well as genomics data. Mr. Allen's statistical expertise includes computer-intensive approaches – such as Monte Carlo simulation, bootstrap analysis, and Markov chain Monte Carlo approaches for Bayesian analyses – as well as other techniques for uncertainty analyses, data quality objectives, quality control/assurance, statistical analyses for site risk assessments, and analysis of clinical trials data. In particular, Mr. Allen has conducted research to study dose-response modeling approaches for developmental toxicants and analyzed cancer dose-response relationships and the issues associated with cancer risk assessment. In addition, Mr. Allen has participated in the development of methods that allow the estimation of risks from epidemiological data.

Mr. Allen received his B.S. in Mathematics and Philosophy from the University of Washington and his M.S. in Biomathematics with a Statistics minor from North Carolina State University. Mr. Allen has numerous publications in the areas of dose response modeling. Mr. Allen has provided expert testimony and has served as manager for numerous projects including multi-disciplinary, multi-year efforts. Mr. Allen has also served on numerous scientific peer review panels for *TERA*, U.S. EPA, and other organizations.

Disclosure: None.

Dr. John Bukowski. Dr. Bukowski is an epidemiologist and Senior Associate at WordsWorld Consulting, a biomedical and medical-writing consultancy located in Dayton, Ohio. His epidemiology and public health career has spanned 20 years, including a broad base of experience within government, academia, and private industry. His clinical research experience includes a post as Director of the Clinical Research Centre at the University of Prince Edward Island, Canada. He has most recently worked as a senior scientist and epidemiologist for ExxonMobil Biomedical Sciences, focusing on such varied topics as air pollution, health effects of solvents, children's health, reproductive health, neurological health, risk assessment, toxicity of metals, and emerging health issues.

Dr. Bukowski received his B.S. in Biology from Wayne State University, his M.P.H. from the University of Michigan and his Ph.D. in Epidemiology from the University of

Medicine and Dentistry of New Jersey. He also holds a D.V.M. from Michigan State University. During his career, he has authored numerous peer-reviewed articles as well as a multitude of reports, critiques, reviews, and white papers. John sits on the Editorial Board for the journal *Dose Response* (formerly *Nonlinearity in Biology, Toxicology, and Medicine*) and is an Adjunct Associate Professor in the School of Medicine at Wright State University.

Disclosure: None

Dr. Julie Goodman. Dr. Goodman is the Director of Epidemiology at Gradient and an adjunct faculty member in the Department of Epidemiology at the Harvard School of Public Health. Dr. Goodman's focus is on human health risks from chemicals in the environment and consumer products. Her primary responsibilities include the design, oversight, analysis, and interpretation of epidemiology studies, and the evaluation of chemical toxicology data, apparent disease clusters, and environmental chemical exposures. Before joining Gradient, Dr. Goodman was a Cancer Prevention Fellow at the National Cancer Institute, where she conducted molecular epidemiology studies on colon cancer risk. She was also instrumental in the development of Polymorphism Interaction Analysis, a powerful statistical tool for cancer risk assessment. Dr. Goodman has conducted a critical review and a weight-of-evidence assessment of soluble nickel compounds. This assessment assessed the causality of soluble nickel compounds and respiratory cancer risk based on animal carcinogenicity studies, mode-of-action studies, and occupational epidemiological studies. Based on this work, Dr. Goodman was an invited observer at the IARC Monograph 100 Meeting C: Metals, Particles and Fibres (March 2009).

Dr. Goodman received her S.B. in Environmental Engineering from the Massachusetts Institute of Technology and her Sc.M. (in Epidemiology) and Ph.D. (in Environmental Health Sciences/Toxicology) from Johns Hopkins University's Bloomberg School of Public Health. She is a Diplomate of the American Board of Toxicology. Dr. Goodman has published the results of her analyses in peer-reviewed toxicology and epidemiology journals, and has presented them to community groups and several regulatory and legislative bodies.

Disclosure: Dr. Goodman's research on nickel is funded by NiPERA. *TERA* has determined that this is not a conflict of interest because there are no nickel producers in the State of Texas and there is no financial relationship between NiPERA and any of the parties that could be affected by the nickel ESL. *TERA* has concluded that this relationship would not result in a bias that would interfere with Dr. Goodman's ability to carry out her duties objectively, but is disclosing the relationship for the purpose of transparency. Dr. Goodman's employer, Gradient Corporation, has ongoing financial relationships with some of the Texas companies that have been identified as releasing nickel to the environment. However, none of these relationships involve work on nickel and Dr. Goodman does not work on any of the projects for these companies.

Dr. Lynne Haber. Dr. Haber is Associate Director of Science at *TERA*. She has 17 years of experience in development of assessment documents, evaluating toxicity, toxicokinetics, and epidemiology studies, and in risk assessment methods development. She has conducted a variety of toxicological assessments, including evaluating and synthesizing data from acute, subchronic, and chronic animal and human toxicity studies, as well as toxicokinetics data, for more than 30 major documents for the U.S. EPA's Office of Water, other EPA offices, and other government agencies. In particular, Dr. Haber was a principal author of U.S. EPA's assessment of soluble nickel compounds. Dr. Haber's research interests include the improved use of mechanistic data in risk assessment, including incorporation of mode of action data in cancer risk assessment, and use of data to replace default uncertainty factors. Her quantitative experience includes serving on a peer review committee for EPA's Benchmark dose guidelines, participating in a workshop on benchmark dose (BMD) and categorical regression, and several publications on issues related to BMD and categorical regression modeling. She has also used both BMD modeling and categorical regression modeling in the development of acute and chronic risk values. Dr. Haber's experience covers all aspects of human health risk assessment including inhalation, oral and dermal toxicology, acute and chronic hazard identification and dose-response for cancer and non-cancer risk assessment and regulatory toxicology.

Dr Haber received her B.S. in Chemistry from University of California, Los Angeles and her Ph.D. in Biology from Massachusetts Institute of Technology. She is a Diplomate of the American Board of Toxicology. Dr. Haber has numerous peer reviewed publication on risk assessment methods and chemical-specific risk assessment. She has served as a panel chairperson or panel member for scientific peer reviews organized by *TERA*, EPA, and other U.S. and foreign government agencies. She has also served on two panels for the NAS/NRC.

Disclosure: Dr. Haber has worked on nickel assessments in the past for several different organizations including U.S. EPA, NiPERA, and a Japanese Government Ministry. In addition, Dr. Haber is currently participating in a peer review panel related to the Port Colbourne cleanup (where nickel was a key driver) for the Ontario Ministry of the Environment. *TERA* has determined that this is not a conflict because it does not involve any financial relationships with organizations in Texas that could be affected by this review.

Dr. Ralph Kodell. Dr. Kodell is a Professor of Biostatistics at the University of Arkansas for Medical Sciences in Little Rock. Previously, he was Director of the Division of Biometry and Risk Assessment at FDA's National Center for Toxicological Research. His research interests include using statistical models for probabilistic risk assessment, statistical methods for low-dose extrapolation in risk assessment, methods for quantifying uncertainty, and classification algorithms for biomedical decision making.

Dr. Kodell received his B.S. in Mathematics from the University of the Ozarks, his M.S. in Mathematics from Stephen F. Austin State University, and his Ph.D. in Statistics from Texas A&M University in 1974. He is a fellow of the American Statistical Association and the Academy of Toxicological Sciences, and a member of the Academy of Distinguished Former Students of the College of Science at Texas A&M University. Dr. Kodell has served on several NAS subcommittees and has served on several scientific peer review panels for U.S. EPA, the Executive Office of the President, and other national and international organizations.

Disclosures: None

This page intentionally left blank.

Appendix E - List of Conference Call Registrants

Panel

Dr. Andy Maier (Facilitator)
Toxicology Excellence for Risk Assessment (*TERA*)

Dr. Bruce Allen
Bruce Allen Consulting

Dr. John Bukowski
WordsWorld Consulting

Dr. Julie Goodman
Gradient

Dr. Lynne Haber
Toxicology Excellence for Risk Assessment (*TERA*)

Dr. Ralph Kodell
University of Arkansas

Observers

Roberta L. Grant
Texas Commission on Environmental Quality (TCEQ)

Robert L. Sielken
Sielken & Associates

Pam Rosett
Lockheed Martin Aeronautics

Akos Szakolcai
Ontario Ministry of the Environment

Brendan Birmingham
Environment Ontario

James Collins
The Dow Chemical Company

Glinda Cooper
U.S. Environmental Protection Agency (U.S. EPA)

Debra Kaden
DakTox, LLC

Ms. Joan Strawson
Toxicology Excellence for Risk Assessment (*TERA*)

Darrell McCant
Texas Commission on Environmental Quality (TCEQ)

Kip Haney
Texas Commission on Environmental Quality (TCEQ)