TCEQ's Response and Additional Questions

The purpose of this conference call is to provide TCEQ an opportunity to ask clarifying questions regarding the comments submitted by the peer review panel and to obtain 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. 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, databases that are lacking 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. As TD typically evaluates total nickel and does not have nickel compound-specific information, TD seeks appropriately conservative surrogate nickel compounds (noncarcinogenic assessment) or mixtures (carcinogenic assessment) for the derivation of values and the protection of public health. While TD believes that for the noncarcinogenic assessment there is not a better alternative than use of the most toxic forms for derivation of likely very conservative comparison values, there is no need to select one particular compound (or two) for the carcinogenic assessment, which may be greatly over- or under-conservative depending on the specific compound(s) selected, when data on mixtures (occupational cohorts) are available and likely conservative (but not unduly conservative without necessity) for the evaluation of nickel emissions in Texas. Due to difficulties in applying unit risk factors (URFs) for specific nickel compounds to Texas air data where only total nickel is known, TD believes nickel compound-specific URFs to be less useful than a URF derived based on a mixture which is still likely to be conservative for the evaluation of Texas nickel emissions.

TD appreciates the reviewers' thoughtful comments and sufficiently understands most of the comments. However, TD would like clarification and/or additional information concerning a few of the comments, as outlined below.

General Issues

On page 5 of the combined comments, Reviewer 3 writes…

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.

(1) What specific epidemiological study limitations are most important to concentrate on for the cancer assessment, given that the DSD is intended to be a relatively short documentation of how health-protective values are derived as opposed to an extensive discussion of all limitations associated with toxicological and epidemiological studies, which are numerous and would alone require a lengthy document for a full discussion? As TD has no epidemiologist on staff and limited resources, can a thorough review paper such as Goodman et al. (2009) be relied upon to provide a discussion of relevant limitations?

On page 5, Reviewer 3 writes…

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 (to the Grimsrud et al. 2003 cohort).

(2) Which Grimsrud study (2002 case-control versus 2003 cohort) is best for a carcinogenic assessment and why?

Acute Evaluation

On page 7 of the combined comments, Reviewer 2 writes…

Nickel sulfate is the correct form for deriving the ReV/ESL. However, I would recommend that the Cirla study be a co-principle study.

(3) 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?

On pages 12 of the combined comments, Reviewer 2 writes…

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.

(4) Why should the RDDR be calculated only for the extrarespiratory region when the reviewer indicates that the appropriate internal dose measure for a systemic effect is the amount of nickel absorbed from the respiratory tract and systemically available? Why would the absorbed dose not depend on the surface area dose to the lung and only depend on the extrarespiratory region? Are the reasons given on page 14 of the DSD for selection of the total respiratory tract not appropriate? If so, why? What is the most appropriate region for the RDDR and why based on available information?

On pages 13 - 15 of the combined comments, the reviewers expressed various opinions on the LOAEL-to-NOAEL uncertainty factor (UF).

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

Chronic Non-carcinogenic Assessment

On page 19, Reviewer 2 writes…

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.

Which human and rat minute volumes do the reviewers suggest for use in the MPPD model and RDDR calculation and why?

Carcinogenic Assessment

On page 25, Reviewer 1 writes…

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.

TD read the IARC sentence quoted above to mean that evidence is sufficient in two separate cases, for nickel sulfate, and then separately for combinations of nickel sulfides and oxides encountered in refining. TD did not see language in this sentence which linked nickel sulfate to nickel sulfides and oxides (i.e., language which made it clear IARC was referring to exposure to a combination of all three forms). An example of such a sentence would be…*there is sufficient evidence in humans for the carcinogenicity of nickel sulfate* ***in combination with*** *nickel sulfides and oxides encountered in the nickel refining industry,* or*…there is sufficient evidence in humans for the carcinogenicity of* ***exposure to the combination of*** *nickel sulfate, nickel sulfides, and nickel oxides encountered in the nickel refining industry.* As written, TD interprets the sentence to mean evidence is sufficient for two types of nickel exposure, one being nickel sulfate, the other being the combination of nickel sulfides and oxides. This may just be an unclear IARC sentence. What language or additional information supports the assertion that IARC is referring to only the combination of the three forms as having sufficient evidence? In a related matter, what should the WOE be for the different forms of nickel and would a WOE for mixtures of nickel be defensible?

On page 29, Reviewer 1 writes…

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.

> (8) Given the elevated respiratory cancer SMRs, historical significance of this study in nickel risk assessment (e.g., USEPA did use Enterline and Marsh 1982 in the derivation of their URF), and lower sulfidic nickel exposure deemed desirable by TD, is exclusion of this study for use in URF derivation justified and why? How about for a supporting study/URF? Can/should an epidemiological study that does not have statistical significance be used to develop a URF and why?

On page 30, Reviewer 2 writes…

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.

> (9) If taken in conjunction with the Reviewer 1 comment above on Enterline and Marsh (1982), these comments would leave TD with no epidemiological studies in the document upon which to base a URF, and human data are preferred although there are always uncertainties associated with use of human epidemiological (and laboratory animal) data.

URFs based on the animal data for individual nickel forms (e.g., the most carcinogenic form or a noncarcinogenic form) may be even less applicable to the forms of nickel in Texas air and air permitting. Given this, TD believes that risk should be based on an occupational mixture more closely approximating the forms TD believes are most likely representative for Texans. Considering the purpose or application of the URF in Texas and that the specific form(s) of nickel in air will be unknown, how could URFs based on individual forms of nickel in animals be considered more applicable for this purpose and how could those URFs be applied in a reasonable manner (e.g., evaluating an air concentration using the most carcinogenic form would be an even more conservative comparison and more greatly overestimate risk, and using a noncarcinogenic form would set risk at zero)? In regard to the Kristiansand cohort, animal studies indicate that soluble nickel is not carcinogenic, but Grimsrud's research challenges that. What are the peer-reviewers' opinions of Grimsrud's research based upon the known facts (as opposed to speculative considerations)?

On page 30-31, Reviewer 3 writes…

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. 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.

(10) The only "predetermined goal" TD has is the adequate protection of public health. In furtherance of this goal, TD attempted to determine what studies had occupational nickel exposures most similar to what might be expected in Texas air in order to calculate a URF which is conservative from a public health perspective while at the same time limiting the gross overestimation of risk which may occur using human or animal studies with high sulfidic nickel. This requires balancing several considerations as opposed to only considering factors which may affect study quality in a vacuum. Although a particular study may be of higher quality relative to another, it may be less useful for the intended purpose. TD certainly recognizes that all epidemiological studies have associated uncertainty (e.g., Goodman et al. 2009 discusses several areas of uncertainty common to epidemiological studies), which does not necessarily preclude their use in a quantitative manner. From a public health perspective, TD believes generalizability is indeed important. Given this, objective study quality criteria, and the preference of TD for human data, which studies are considered superior by the reviewers for the end purpose (i.e., a somewhat conservative generalization of risk to the population of Texas) and why?

On page 32, Reviewer 1 writes…

It is more appropriate to use animal data as the primary data for the ReV/ESL calculations and the epidemiological data as supportive.

(11) Given that TD may not know what form(s) of nickel may be in air in a particular case, how do the reviewers envision TD being able to apply nickel compound-specific URFs derived based on animal data as suggested by Reviewer 1 in such a way as to not grossly over- or under-estimate risk?

On page 36, Reviewer 1 indicates that only insoluble (sulfidic and oxidic) nickel should be used to calculate URFs (excludes the possibility of using Enterline and Marsh 1982), and Reviewer 2 comments regarding possible (speculative) uncertainty as to the level of sulfidic nickel in the Kristiansand cohort as compared to what is expected in Texas air (may imply that this cohort should not be used)…

(12) Given that there is no cohort likely to have been exposed to a nickel mixture remarkably similar to what may be expected in Texas air, with the same being true for animal studies, and that limiting the carcinogenic assessment to sulfidic and oxidic nickel (suggested by Reviewer 1) would only exacerbate the overestimation of risk for the Texas population, what specific URF analyses do the reviewers suggest as more applicable for the evaluation of total nickel ambient air data? What epidemiology studies would be appropriate for the TD to use to develop a URF?

(13) Given the information in the draft DSD, is the draft URF likely to over- or under-estimate risk for Texans, and given that answer, do the reviewers believe use of the draft URF would be adequately protective of public health?

(14) What regression diagnostics should be used?

(15) Is the central estimate or the 95% UCL estimate the best estimate and why?

(16) Is the URF weighting procedure used to calculate the final URF reasonable and justified?