Texas Commission on Environmental Quality
Responses to Public Comments

Date

Texas Commission on Environmental Quality
Toxicology Division
Office of the Executive Director

Prepared by:

Toxicology Division
Office of the Executive Director

Comments from the Public

The Toxicology Division (TD) of the TCEQ has prepared responses to comments from the public on the updated TCEQ Guidelines to Develop Toxicity Factors (hereafter referred to as the Guidelines), posted April 10, 2012, for a 60 day public comment period. Only one set of comments was received by the TD from the American Chemistry Council Center for Advancing Risk Assessment Science and Policy (ARASP). The TD would like to thank the public commenter for their careful review of the updated Guidelines and the TCEQ Responses to Peer Review Report, and for their constructive suggestions. The TD also chose to include suggestions for improvement of the methodology in Section 3.13 from the Science Panel of the Alliance for Risk Assessment fourth workshop on *Beyond Science and Decisions: From Problem Formulation to Dose-Response* (May 22-24, 2012, Austin, Texas) as public comments. After careful review and due consideration, the TD revised the Guidelines document to incorporate suggested revisions of the public comments which increased the scientific and technical merit of the Guidelines or clarity for agency staff. Public comments are addressed in the order they were received.

Public Comments
- Public Comment 1
  - ARASP Introduction Letter
  - ARASP Comments on TCEQ Guidance
    - 1 Introduction
      - 1.1 Documents Considered in this Review
    - 2 General Comments
      - 2.1 Mode of Action
      - 2.2 Determination of a Mutagenic MOA
      - 2.3 Choice of a Point of Departure (POD)
      - 2.4 Identification of Inhalation Effect Screening Levels
      - 2.5 Acute Inhalation Toxicity Factors
      - 2.6 Odor as a Basis for Acute Screening Levels
      - 2.7 Use of Epidemiological Studies for Toxicity Factor Derivation
        - 2.7.1 Classification Schemes for Uncertainty
        - 2.7.2 Should the Linear Slope be Based on the Upper Bound on the MLE Slope (q1*) or the BMDL?
    - 3 Specific Comments
      - 3.1 Chapter 1
      - 3.2 Chapter 2
      - 3.3 Chapter 3
      - 3.4 Chapter 4
      - 3.5 Chapter 5
      - 3.6 Chapter 6
      - 3.7 Chapter 7
    - 4 References
- Public Comment 2
  - Suggestions for improvement of the methodology in Section 3.13

Public Comments

Public Comment 1

Kimberly Wise Ph.D.

The American Chemistry Council Center for Advancing Risk Assessment Science and Policy (ARASP)

ARASP Introduction Letter

The Center for Advancing Risk Assessment Science and Policy (ARASP), which is managed by the American Chemistry Council (ACC), is pleased to provide comments in response to the Texas Commission on Environmental Quality's (TCEQ) request for public comment on the revision of the TCEQ Guidelines to Develop Toxicity Factors (Guidelines). ARASP[1] is a coalition of independent groups and associations that promotes the development and application of up-to-date, scientifically sound methods for conducting chemical assessments. ARASP also fosters activities to promote adoption of policies and practices, both within and outside government, that assure that the best available science underlies chemical safety determinations.

The TCEQ's 2012 Guidelines include the methodology for derivation of effects screening levels, reference values, unit risk factors, and slope factor (SFo) values. In the attached, ARASP offers general comments on the overarching risk assessment issues presented throughout the Guidelines and also provides more detailed comments on specific aspects of the Guidelines (i.e. mode of action, the use of linear low dose extrapolation, point of departures and the use of epidemiological studies).

In general, the Guidelines are well written and balanced. TCEQ presents a thoughtful approach to the consideration of mode of action (MOA) in risk assessment. Specifically, the discussion regarding the determination of mutagenic MOA and considerations for dose-response in determining threshold and non-threshold MOAs for cancer and non-cancer is scientifically balanced and appropriate. However, the guidelines would benefit from greater discussion and analysis regarding the following: (1) the use of odor as a basis for an effect screening level; (2) the use of the shape of the dose-response curve in the selection of point of departures, and (3) the uncertainty associated with the derivation of the effect levels.

Response: See responses below to specific comments.

We hope you will give full consideration to the general comments included above and the detailed comments provided in the attachment.

[1] ARASP is comprised of the following member organizations: Acrylonitrile Group, ACC Chlorine Chemistry Division, ACC Ethylene Oxide Panel, ACC Formaldehyde Panel, ACC Hexavalent Chromium Panel, ACC High Phthalates Panel, ACC Hydrocarbon Solvents Panel, ACC Oxo Process Panel, ACC Propylene Oxide/Propylene Glycol Panel, ACC Public Health and Science Policy Team, ACC Olefins Panel, American Cleaning Institute American Petroleum Institute, CropLife America, Halogenated Solvents Industry Alliance, Silicones Environmental, Health and Safety Council of North America, and the Styrene Information and Research Center.

ARASP Comments on TCEQ Guidance

1 Introduction

The 2012 guidance document from the Texas Commission on Environmental Quality, titled TCEQ Guidelines to Develop Toxicity Factors (RG-442) (Guidelines), includes methodology for derivation of 1) acute and chronic inhalation effects screening levels (ESLs); 2) acute and chronic inhalation reference values (ReVs); 3) chronic inhalation unit risk factor (URF) values; and 4) chronic oral reference dose (RfD) values and slope factor (SFo) values.[2]

A draft of the Guidelines was published in April 2011 and was the subject of a Letter Peer Review conducted by Toxicology Excellence for Risk Assessment (TERA). The final Peer Review Report was released on August 31, 2011. These peer review comments were addressed and the response to comments document was released in early April of 2012. TCEQ Guidelines to Develop Toxicity Factors, the subject of this review, was released at the same time for public comment with the comment period ending on June 8.

The following comments consider overarching risk assessment issues presented throughout the Guidelines such as mode of action (MOA), the use of linear low dose extrapolation, and the use of epidemiological studies. Additionally, specific comments on the narrative are provided.

1.1 Documents Considered in this Review

This review focused primarily on five documents:

- *TCEQ (2012) TCEQ Guidelines to Develop Toxicity Factors*
- *TCEQ (2012) Texas Commission on Environmental Quality Responses to Peer Review Report, April 5, 2012*
- *TERA (2011) Report of a Letter Peer Review of the Texas Commission on Environmental Quality's (TCEQ) updates to its Guidelines to Develop Inhalation and Oral Cancer and Non-Cancer Toxicity Factors, August 31, 2011*
- *TCEQ (2011) Chapter 7. Hazard Characterization and Exposure-Response Assessment Using Epidemiology Data. Draft June 7, 2011.*
- *TCEQ (2011) Guidelines to Develop Inhalation and Oral Cancer and Non-Cancer Toxicity Factors. Chapters 1-6. Draft June 7, 2011*

These comments also considered the TCEQ's responses to the peer review comments received on the 2011 draft document, the subsequent changes made to the draft 2011 document as well as the content of the 2012 Guidelines. The following earlier documents were also reviewed:

- *TCEQ (2006) Guidelines to Develop Effects Screening Levels, Reference Values, and Unit Risk Factors. RG-442*

[2] The guidance document has gone through several iterations of comment and review since its initial publication in 2006. For convenience, this document will be referred to in these comments as "the Guidelines."

- *TCEQ (2010) Response to Public Comments Received on the Proposed Interim Guidelines for Setting Odor-Based Effects Screening Levels*
- *TCEQ (2010) Interim Guidelines for Setting Odor-Based Effects Screening Levels. May 28, 2010*

2 General Comments

This section presents overarching comments on the following six areas:

- *Mode of Action (MOA)*
- *Determination of a Mutagenic MOA*
- *Choice of a Point of Departure*
- *Acute Inhalation Toxicity Factors*
- *Identification of Inhalation Effect Screening Levels*
- *Odor as a Basis for Acute Screening Levels*
- *Use of Epidemiological Studies for Toxicity Factor Derivation*

2.1 Mode of Action

In the Guidelines, TCEQ presents a balanced and thoughtful approach to the consideration of MOA in risk assessment. Section 3.4 provides a discussion of how knowledge of mode of action (MOA) can inform the choice of dose metric, the occurrence of effect thresholds, human relevance of the adverse effect, and the identification of sensitive subpopulations. In this way, the centrality of the concept of MOA to risk assessment is highlighted and introduced in a relatively non-technical fashion. The term "MOA" is used 260 times in the Guidelines and, appropriately, is used to inform all aspects of the development of toxicity factors. Indeed, MOA is the organizing principle of Chapters 3, 4 and 5. TCEQ uses this concept to provide links between the various aspects of risk evaluation.

Section 5.7 on Non-threshold and Threshold Carcinogens examines the use of MOA in areas of risk assessment that have the most potential for regulatory impact. This section is well written, has clear definitions and provides useful examples.

Response: No response required.

2.2 Determination of a Mutagenic MOA

The discussion of a mutagenic MOA in Section 5.7.4.1 is well written and balanced. The use of the linear no-threshold assumption for risk assessment of chemical carcinogens began in 1977 with its adoption by the Safe Drinking Water Committee of the National Academy of Sciences (NAS). The belief, at that time, was that the dose-response of radiation carcinogenesis was linear at low doses.[3] Adoption of the linear no-threshold assumption, however, occurred prior to the advent of research on DNA repair.

DNA damage such as DNA adduct formation is not, by itself a mutational event (Jarabek et al., 2009), but in the late 1970's, prior to the advent of research on DNA repair, it seemed reasonable that DNA damage

[3] Dr. Edward Calabrese has explored the history of the scientific basis for and adoption of the linear no-threshold assumption for regulation (Calabrese, 2009, 2011a, 2011b).

itself be considered equivalent to a mutagenic event. A number of studies provide statistical support for the conclusion that DNA-reactive carcinogens can display non-linear/threshold dose-responses (Williams et al., 1996, 1999; Wadell, 2003; Wadell et al., 2006; Fukushima et al., 2002; Gocke and Wall, 2009; Pottenger and Gollapudi, 2010; Pottenger et al., 2009). These thresholds are likely due to DNA repair mechanisms and other compensatory processes (Swenberg et al., 2011).

TCEQ should consider adding the in vivo hypoxanthine phosphoribosyl transferase (HPRT) gene mutation assay to the list in Table 5-4. Specific rodent genes can serve as reporters for in vivo mutation in splenocytes (Chen et al., 1998; Bol et al., 1999; van Zeeland et al., 2008). The Pig-a gene mutation assay is currently being validated and while it is likely too early to include this assay in the Guidelines, TCEQ may wish to consider its inclusion in a future update (Bhalli et al., 2011; Cammerer et al., 2011).

Response: The hypoxanthine phosphoribosyl transferase (HPRT) gene mutation assay has been added to Table 5-4 as suggested.

2.3 Choice of a Point of Departure (POD)

The discussion of the choice of a POD on pages 168-171 conveys flexibility in determining the POD based on data being considered. TCEQ should consider referencing this section in earlier discussions of POD determination in chapters 3 and 4.

Response: No additional discussion/text was added. Chapter 3 describes common procedures used to derive both acute and chronic toxicity factors, whereas Chapter 5 on pages 168-171 describes procedures used to derive chronic toxicity factors and discusses derivation of a POD based on observed data.

This reliance on the data is continued in the discussion on pages 217-218. However, as one of the peer reviewers of the 2011 draft pointed out, it seems unlikely that an epidemiologic study would ever be large enough to use a 1E-05 or 1E-06 level as the POD. That said, the POD should be set as low as possible but still within the observable range (EPA, 2005a) and should also consider the nature of the adverse effect and the shape of the dose-response curve (see below).

While not directly related to selecting the POD as a numerical value, Section 3.10 (which was Section 3.11 in the 2011 draft) is somewhat confusing because it puts selection of the critical effect after derivation of the POD. If multiple PODs were derived corresponding to the various effects observed, then this scheme would make sense. However, the chosen POD should reflect the dose-response of the critical effect, and it would make no sense to derive a single POD without first selecting the critical effect from the set of observed effects. In this regard, the question posed to the peer reviewers was unclear. This section could be clarified by indicating that multiple PODs would correspond to multiple effects from which the critical effect (and associated POD) would be selected.

Response: No additional discussion/text was added. As a POD_{HEC} is derived for each potential critical effect prior to selection of the critical effect (see page 90, line 21, through page 91, line 5), the comment acknowledges the procedure is appropriate.

The peer reviewers were asked in question 2.1.7: "Should the critical effect be selected before or after uncertainty factors are applied?" Uncertainty factors (UFs) should be applied after the selection of the

critical effect and corresponding POD. Because the chosen critical effect may show a "hanging" LOAEL only, it is important to choose the critical effect before application of uncertainty factors. A "hanging" LOAEL occurs when the lowest non-zero dose in the dose range in a bioassay or experiment produces an adverse effect and the NOAEL remains unknown.

Response: No response required.

In section, 3.1.3, the narrative about using values of the POD lower than 1% is not entirely clear. The value of a POD, as noted in the Guidelines, is within the observable range of the data and representative of a finite level of risk when the confidence limits are considered. The quote from Seiler and Alvarez (1994) on page 109 is appropriate.

Response: It appears that this comment is referring to Section 3.13.3. The sentence concerning PODs lower than 1% has been clarified.

However, TCEQ should also point out the choice of a POD should include consideration of the shape of the dose-response curve. In Figure 1 below, two hypothetical dose-response curves, one sublinear and the other supralinear, are considered.

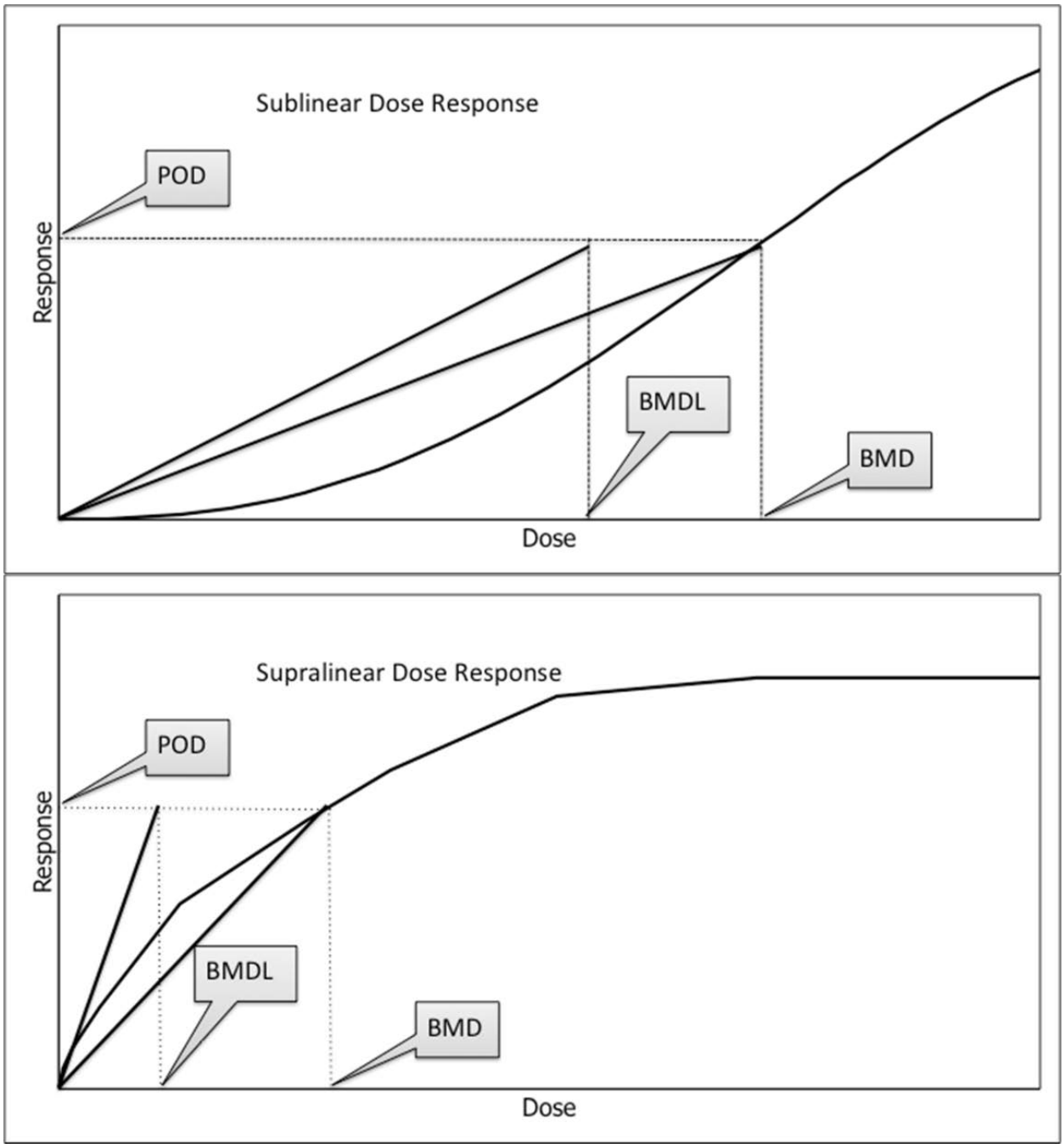

Figure 1. Comparison of the linear slope estimates from sublinear and supralinear DR curves. Upper panel, hypothetical sublinear DR curve illustrating that both the linear slope from the BMD and BMDL to the origin are much more conservative than the actiual low dose slope. Lower panel, hypothetical DR curve illustrating that the linear slope from the BMD may be less conservative than the actual low dose slope but the conservatism of the linear slope from the BMDL relative to the actual low dose slope is unknown. The majority of DR curves will be sublinear in the low dose region (Waddell, 2003; Waddell et al., 2006).

Response: This comment is relevant to dose-response assessment for nonthreshold MOA chemicals, most specifically, linear low-dose extrapolation for mutagenic or unknown MOA carcinogens. The comment indicates that the choice of a POD should include consideration of the shape of the modeled dose-response curve, and in some cases (sub/supralinear dose-response) use of a BMD/BMDL for linear low-dose extrapolation may lead to an over/underestimation of low-dose slope compared to the dose-response curve. As the dose-response curve is defined by the empirical data, this comment is also implicitly expressing concern regarding cases when a BMD/BMDL is used as a POD for linear low-dose extrapolation but is associated with a response level significantly higher than what would be expected based on the actual data. To address this comment, language was added to Section 5.7.3.1 (regarding carcinogen dose-response assessment) that indicates: (1) when using BMD modeling to derive a POD, the estimated values (e.g., EC10) are compared with the empirical dose-response data and values that significantly conflict with empirical data will generally not be used because of uncertainties associated with extrapolations beyond the experimental data (NRC 2001); (2) if a POD is to be used for linear low-dose extrapolation, the TCEQ may consider how the resulting slope compares to the low-dose slope on the modeled dose-response curve if sufficiently informed by empirical data in the low-dose region; and (3) the TCEQ evaluates potential PODs and determines the most appropriate POD for use on a case-by-case basis using best scientific judgment.

Figure 1 shows both sub-linear and supra-linear D-R curves. In the sub-linear curve, choosing a lower value for the POD will generally cause the maximum likelihood estimate of the slope to be lower (less risky) than for a higher POD value. The BMDL will depend on the choice of the POD and the confidence interval around the BMD. As can be seen in the upper panel of Figure 1, both the linear slopes at the BMD and BMDL tend to overestimate the slope of the DR curve in the low dose region.

In the case of a supralinear DR curve, the choice of a lower POD will have the opposite effect—the linear slope from the BMD will likely tend to underestimate the slope of the D-R curve in the low dose region. The linear slope from BMDL may overestimate or underestimate the low dose slope depending on the confidence of the study.

In general, consideration of basic biology suggests that most D-R curves are sub-linear (Waddell, 2003; Waddell et al., 2006). Nonetheless, TCEQ should include additional narrative about the influence of the shape of the dose-response curve on the risk estimate and its consideration in the choice of POD.

Response: Per section 5.7.3.1 of the Guidelines, slope of the dose-response curve at the POD is one factor to be discussed in a narrative concerning the choice of POD, which is one aspect of what the comment concerns (i.e., over/underestimation of the slope of the D-R curve in the low dose region). Discussing the slope of the dose-response curve at the POD per the Guidelines has implications for the influence of the shape of the dose-response curve (at the POD at least) on the risk estimate. That is, for example, if the slope of the dose-response curve at the POD is significantly lower than the slope factor derived using the POD, then the risk estimate will obviously be more conservative than that using the modeled dose-response curve. However, this does not fully address the concern expressed as the low-dose slope below the POD may significantly differ from the slope at the POD (e.g., very sub/supra linear dose-responses). Thus, to fully address the comment, language was added to Section 5.7.3.1 (regarding carcinogen dose-response assessment) that indicates if a POD is to be used for linear low-dose

extrapolation, the TCEQ may consider how the resulting slope compares to the low-dose slope on the modeled dose-response curve if sufficiently informed by empirical data in the low-dose region. Such issues will be discussed in DSDs on a case-by-case and as-needed basis.

Some readers interpreted the narrative on page 109 as suggesting that low POD values would result in overly conservative toxicity factors. The Guidelines did not make clear that the choice of a low risk level for the POD, such as 10-3, would not necessarily lead to a more stringent non-threshold slope factor. TCEQ should clarify this point.

Response: The comment did not identify the specific language being referred to in Section 3.13.3, and any such interpretation is unintended. This section does not address the derivation of nonthreshold (e.g., cancer) slope factors. The Guidelines generally indicate that choice of risk level for the POD should be supported by the data (Section 3.6.3.1) and make no representations about the implications of the choice of the specific excess risk level (e.g., high or low) for the stringency of the resulting cancer slope factor. Thus, no further clarification is needed.

2.4 Identification of Inhalation Effect Screening Levels

Section 3.13 of the Guidelines discusses the generation of inhalation effect levels where health effects would be expected to occur for threshold and non-threshold chemicals. For threshold carcinogenic or non-carcinogenic chemicals TCEQ notes that the value should be based on the lowest observable adverse effect level (LOAEL) from the study that identified the critical effect. For non-threshold carcinogenic or non-carcinogenic effects, TCEQ recommends setting the value based on no significant excess lifetime risk of cancer risk of 1 in 1,000 or 1 in 10,000. TCEQ also discusses the use of conventional cancer bioassays and epidemiology studies to model down to an increased incidence of 1 in 100 or 1 in 10 based on the study sample size and its statistical power.

Response: No response required.

For any value that is derived it must be noted that there is considerable uncertainty with extrapolation from high dose exposures down to the low dose region, sometimes quantitative extrapolations over several orders of magnitude may be required. When deriving a chronic inhalation effect level for chemicals that exhibit a non-threshold MOA, TCEQ should utilize the available animal and epidemiological data to derive a dose-response curve that would identify doses associated with statistically significant increases in tumor incidences. Utilizing this type of dose-response curve to derive the inhalation effect level would offer TCEQ a level that would likely be more in line with the observed study effects. This approach is more concordant with TCEQ's proposed approach for chemicals with a threshold MOA, in which the LOAEL is used to set the effect screening level. As well, it would also be helpful if TCEQ provided, in the Decision Support Documents (DSDs), a graph of both the PODs and inhalation effect levels to show how far effect levels are from ESLs. The Guidelines should provide a discussion of such graphs and provide at least one example.

Response: Text regarding using generic risk levels in an attempt to identify inhalation observed effects levels has been removed from Section 3.13.3. Considerable uncertainty can be associated with extrapolation from higher dose exposures where excess cancer risk higher than these generic risk levels has been observed down to a lower dose associated with a generic risk level. Instead, the Guidelines

emphasize use of the actual study-specific excess risk levels and doses where statistically elevated risk was observed, consistent with the goal of identifying "observed" effects levels (as opposed to uncertain, downwardly extrapolated theoretical effects levels at generic target risks). Text has been added to Section 3.13.3 indicating that the air concentration (or range of concentrations) associated with the lowest exposure group for which statistically elevated cancer risk was observed in an epidemiological or animal study can be used to derive the inhalation observed effect level. Consistent with the comment, the guidance provided in this section focuses on identifying the lowest doses associated with statistically significant increases for nonthreshold MOA chemicals (e.g., carcinogens). Graphical depictions of ESLs, PODs, and observed effects levels may be provided in DSDs, but this decision will be made on a case-by-case basis.

Additionally, TCEQ's approach for evaluating potential risks and communicating these findings should also consider the magnitude, frequency and duration of exposure to the chemical for which the inhalation effect level is derived. For example, a single excursion over an established value for a chemical carcinogen will NOT likely result in an adverse health impact, since the exposure periods used to establish the value were based on averaged exposures over a person's lifetime (70 years). Thus, if TCEQ is to use these values to communicate about potential expected health effects associated with shorter durations of exposure, exposure-duration adjustments would be warranted.

Response: Limited text was added to emphasize that TCEQ's approach does consider the magnitude, frequency, and duration of exposure to the chemical for which the inhalation effect level is derived. The Guidance already indicated that the derived observed effects levels are applicable to exposure to the same level and for a similar or longer duration (i.e., they apply to exposure scenarios that are similar or greater in magnitude, frequency, and duration). Thus, short-term excursions over a long-term average effects level should not be interpreted as being associated with a likelihood of adverse effects. These conditions and caveats will be part of a narrative in DSDs which puts observed effects levels into proper context, as now indicated in the Guidance. As discussed in the Guidance, exposure-duration adjustments will only be performed if such adjustments are predictive of toxicity (as opposed to simply being conservative) in an attempt to maintain the probability of response associated with observed effects levels.

2.5 Acute Inhalation Toxicity Factors

Chapter 4 on developing acute toxicity factors is generally well constructed. It received many peer review comments regarding the use of Haber's rule and the modification of it by Ten Berge (1986). Section 4.3.3 on page 134 is thoughtful regarding the consideration of toxicokinetics (TK) and toxicodynamics (TD) in developing acute toxicity factors. Additional specific comments can be found below.

Response: No response required.

2.6 Odor as a Basis for Acute Screening Levels

Organoleptic effects such as taste and smell vary greatly both from individual to individual and, at different times, in the same individual. This inter- and intra-individual variability occurs for both odor detection and odor perception. For example, schizophrenics appear more sensitive to odors than others and those suffering from Parkinson's Disease appear to be less sensitive to odors (Masaoka et al., 2008;

Moberg et al., 2003). Odor and memory are strongly connected in the brain; this is known as the "Proust phenomenon" (Matsunaga et al., 2011). Odor perception may be altered by early experience (Poncelet et al., 2010). Odor perception is influenced by reproductive hormones in both men and women (Doty and Cameron, 2009). The use of drugs may also change odor perception (Elsner, 2001).

In addition to these observed phenotypic effects, the olfactory receptor gene superfamily contains 390 putatively functional genes and 465 pseudogenes. These genes undergo allelic exclusion and each olfactory receptor neuron expresses only a single allele (Olender et al., 2008). In addition, the set of olfactory receptor genes has changed greatly over the course of evolution, possibly due to variation in gene copy number (Niimura, 2009; Waszak et al., 2010).

There is also a psychological aspect to odor detection and perception. Cacosmia is a syndrome involving headache, nausea and dizziness in response to common odors such as perfume, gasoline, or tobacco smoke and is associated cognitive and emotional effects (Bell et al., 1993, 1996; Simon G.E. et al., 1993). Cacosmia has been referred to as idiopathic environmental intolerance (IEI), also known as multiple chemical sensitivity. Because medicine or toxicology cannot provide the etiology of the symptoms in IEI, these symptoms are commonly associated with psychogenic illness since those with IEI have a significantly higher lifetime prevalence of mood and anxiety disorders (Tarlo et al., 2002; Staudenmayer et al., 2003a, 2003b). A number of illness outbreaks have been attributed to psychogenic causes (Jones et al., 2000; Page et al., 2010).

A number of studies have noted an "awareness bias" in which individuals who perceive an environmental threat such as proximity to a landfill or industrial facility, and who also worry about potential health effects associated with the perceived threat, tend to report more ill health in the absence of any measurable medical or biological effect (Moffat et al., 2000).

In light of both variability and the emotional connection with odor, there is a great deal of uncertainty regarding the use of odor panels to set welfare-based ESLs. TCEQ should acknowledge this uncertainty.

Response: It is correct that odor perception and the emotional response thereto is highly variable. Thus, there is an inherent level of uncertainty using odor panels to set welfare-based ESLs. TCEQ utilizes the geometric mean (Level 1 & 2) or the lowest value (Level 3) of the odor detection threshold from high quality odor studies to set the odor-based ESL. This approach is both conservative and representative. Furthermore, it is not necessarily the goal of an odor-based ESL to prevent 100% of odors, particularly not idiopathic responses to odor. The purpose of these values is that of scientifically defensible screening levels, which may be used in the evaluation of air permits or air monitoring data. If TCEQ finds that a facility is giving rise to odor nuisance, typically accomplished via complaints and subsequent investigations, then the agency may respond through enforcement.

2.7 Use of Epidemiological Studies for Toxicity Factor Derivation

TCEQ should add a general section on uncertainty analysis with an appropriate classification of the various types of uncertainty identified on page 228. Such a classification might help tie various portions of the document together. Please see section 2.6.1 below.

Response: It appears this comment is actually referring to section 2.7.1 below. A general section on uncertainty has been included in Section 5.7 *Nonthreshold Carcinogens and Threshold Carcinogens* (see

Section 5.7.4). Model uncertainty, parameter uncertainty, and human variation are included, with a reference to USEPA (2005a) and other documents that provide additional information about uncertainty analysis. Chapter 7 has an adequate section on uncertainty analysis and will not be revised. As stated in Chapter 7 and in the revised Section 5.7.4, "Uncertainty analyses are done on a case-by-case basis, and their components may be numerous and variable".

TCEQ should also add narrative to place the choice of the BMD or the BMDL as the POD in a historical context. Please see section 2.6.2 below on the use of the upper bound slope for non-threshold dose-response evaluation.

Response: It appears this comment is actually referring to section 2.7.2 below. No changes were made. It is beyond the scope of the Guidelines to compare models used to calculate the BMD or the BMDL versus the historical use of other models to calculate the upper Bound on the MLE Slope (q1*) as the POD.

2.7.1 Classification Schemes for Uncertainty

It may also be appropriate to have more than one classification scheme—possibly one for evaluation of animal data and another for epidemiologic data. Some chemicals will not have as extensive epidemiological analyses as other chemicals. EPA (2001) divides uncertainty into parameter uncertainty, model uncertainty and scenario uncertainty, but this classification is somewhat specific to Monte Carlo analysis and probabilistic risk assessment and may not be generally applicable. In this regard, Spiegelhalter et al. (2011) provide a discussion of the graphic communication of uncertainty that could be a starting point for thinking about various types of uncertainty in a chemical risk assessment.

Response: See response to2.7 above. Uncertainty analyses are done on a case-by-case basis, and their components may be numerous and variable. Section 5.7.4 acknowledges that uncertainties, when using animal data to determine cancer risk estimates, are different from uncertainties using human epidemiological data. Spiegelhalter et al. (2011) provides useful information on risk communication, which may be considered by the TD staff, but is not included in the Guidelines.

2.7.2 Should the Linear Slope be Based on the Upper Bound on the MLE Slope (q1) or the BMDL?*

On page 219, the Guidelines indicate the lower bounds on the concentration are not very responsive to the observed dose-response data and that the central maximum likelihood estimate is a better comparator of risk between chemicals. This raises the question of the purpose of using the benchmark dose lower confidence limit (BMDL), lowest effective concentration (LEC) or, as in earlier practices, the upper bound on the slope, q1, as the risk estimator rather than the value of the maximum likelihood estimate (MLE) (Crump, 1984).*

Response: No changes were made. Section 5.7.3.1 discusses the purpose of using the BMDL or LEC as the risk estimator POD as opposed to the BMD or EC: "The 2005 Cancer Guidelines refers to the Science Advisory Board who stated 'it may be appropriate to emphasize lower statistical bounds in screening analyses and in activities designed to develop an appropriate human exposure value, since such activities require accounting for various types of uncertainties and a lower bound on the central estimate is a scientifically-based approach accounting for the uncertainty in the true value of the ED10 [or central estimate].'" This section then provides a discussion of when it may be appropriate to use the

central estimate (BMC or EC) as opposed to the 95% UCL of the central estimate (BMCL or LEC) when human epidemiology exposure response studies are used for dose-response modeling. The TCEQ provides both the central estimate (BMD, EC or MLE) and the 95% UCL and 95% LCL on the MLE (BMDL, LEC) for any cancer evaluation in the chemical-specific DSDs. This information is necessary to understand and discuss the uncertainties in cancer estimate.

A number of scientists in the 1960's and 1970's developed statistical models for cancer dose-response and then suggested using the upper confidence limit (UCL) of the low dose slope for regulatory purposes (Mantel et al., 1961, 1975; Mantel and Schneiderman, 1975; Crump et al., 1976, 1977; Guess et al., 1977; Hartley and Sielken, 1977; Crump, 1984). The reasons for this choice are not provided in these papers.

Additionally, in the 1980's, EPA chose to use the multistage model and a maximum likelihood method for estimation of the UCL on the low dose linear slope or q1 (Anderson et al., 1983). The multistage model uses a polynomial to model the dose response as follows:*

P(d) = 1 – exp[-(q1d + q2d2 + … + qKdK)]

Where P(d) = probability of tumors as a function of dose

d = dose

qx = polynomial coefficients of the multistage model

q1 was designated as the UCL of the MLE of q1, the linear term in the polynomial (Crump, 1984). Please note that this procedure is not the same as using the BMDL to obtain a linear slope. This should be clearly stated in the Guidance.*

Response: It is beyond the scope of the Guidelines to compare the procedures used to calculate the BMDL versus the historical use of other procedures to calculate the upper Bound on the MLE Slope (q1*).

The reasons for choosing the UCL of the slope rather than the MLE are not made explicit in Anderson et al. (1983), but the choice of the UCL may be to account for the general uncertainty associated with low dose extrapolation.

Response: See above response.

EPA began the implementation of benchmark dose modeling in 1999 and began using the linear slope from the BMDL rather than the q1. Hence, the practice of using a conservative confidence limit for estimating the linear slope factor used for regulation persisted even though the modeling procedures were different. The lack of a specific rationale for the use of a conservative confidence limit as the POD rather than the central estimate provides TCEQ the opportunity to explore this rationale.*

Grant et al. (2007, 2009) are good examples of the type of effort needed to understand (1) why no explicit reason for choosing the upper bound has been given; (2) the ramifications of choosing the MLE rather than the upper bound estimate on risk; and (3) the most appropriate policy choices for different types of dose-response data. TCEQ should consider conducting this analysis and including the results in a future update to the guidance. TCEQ can greatly enhance its leadership role among regulatory agencies

by careful, explicit, and quantitative consideration of this issue and publication of the outcome and any recommendations.

Response: No changes were made. Section 5.7.3.1 provides a discussion of when it may be appropriate to use the central estimate (BMC or EC) as opposed to the 95% UCL of the central estimate (BMCL or LEC). The TCEQ provides both the central estimate (BMD, EC or MLE) and the 95% UCL and 95% LCL (BMDL, LEC) for any cancer evaluation in the chemical-specific DSD. This information is necessary to understand and discuss the uncertainties in cancer estimates.

3 Specific Comments

This section provides specific comments on the narrative, figures and tables. Both the Guidelines and the TCEQ responses to the peer review comments on the 2011 draft are considered. These comments are divided by chapters and include pin-point cites, with page and line numbers.

3.1 Chapter 1

Page 3, line 38:

Sub-linear and hockey-stick curves are interpreted as nonlinear. In fact, a method of fitting the hockey stick function to dose-response data has been developed for the purpose of identifying thresholds quantitatively for both continuous and dichotomous data (Lutz and Lutz, 2009; Bogen, 2011).

Response: The following text was deleted from Chapter 1, page 3: "(e.g., as opposed to sub/supralinear, hockey-stick shaped, etc.)."

Page 4, lines 12-18:

It is appropriate to distinguish the difference between threshold and non-threshold effects from the difference between cancer and non-cancer effects. Dr. Lorenz Rhomberg has opined further on this issue (Rhomberg et al., 2011a, 2011b)

Response: No additional discussion/text was added. Chapter 1 is an introduction to the guidelines. The distinction between threshold and non-threshold effects and cancer and non-cancer effects is thoroughly discussed in subsequent sections of the guidelines.

Page 6, lines 17-24:

Margin of Exposure (MOE) is not typically used as a risk assessment method; however, TCEQ should provide an MOE value, or range of values, to be considered in certain cases. In other words, if the MOE is very high (e.g., 100,000) there is very low concern whereas an MOE of 1 may be a reason for higher prioritization and/or regulatory action. Since the Guidelines are not regulation, it may be appropriate for TCEQ to suggest a default range for the MOE at which there would be some regulatory concern.

Response: The MOE is a valuable tool for the management of risk. While the TD may choose to use these values to make an informed decision, we would not aim to set an acceptable MOE to a finite range of numbers. Thus, we will not identify an acceptable range for the MOE in the guidelines.

Page 7, lines 3-8:

Is there a reference for cumulative and aggregate exposures? Paloma Beamer uses these terms in her recent work on children's pesticide exposure (Beamer et al., 2012), but it is unclear if she is the originator of the terms. Ideally, a regulatory guidance document should be referenced as the source of these terms. Additionally, the most recent language noted by WHO, the International Programme on Chemical Safety (IPCS) utilizes "combined exposures" rather than "cumulative or aggregate exposure". TCEQ should review current nomenclature and provide a reference for the term it utilizes in the Guidance to describe these types of exposures.

Response: No additional discussion/text was added. As described on Page 6, lines 26-31, the terms "cumulative" and "aggregate" have been clearly defined in the USEPA Framework for Cumulative Risk Assessment (USEPA 2003a). The TCEQ does not believe that it is necessary to utilize combined exposure. The TCEQ air permitting process and ambient air monitoring program minimize potential cumulative or aggregate exposure in an area where multiple sources for a chemical and/or multiple chemicals exist.

Page 8, lines 9-12:

APWL (air pollutant watch list) should be included in the list of acronyms.

Response: APWL was added to the list of acronyms.

Page 9, lines 13-29:

It is somewhat arbitrary to use a hazard quotient (HQ) of 0.3 for consideration of cumulative effects of multiple chemicals. This choice assumes that three chemicals present at 30% of their effect screening levels is acceptable, but not four chemicals. Are there typically three or fewer chemicals considered? Is this choice reasonable based on historical monitoring data? Dr. Laurie Haws of ToxStrategies, Inc. noted this issue in public comments on the 2011 draft. TCEQ responded that an HQ of 0.3 is used because there may be multiple sources of a chemical. Narrative explaining this reasoning should be included in the Guidelines.

Response: The HQ of 0.3 used for the calculation of health-based ESLs from derived referenced values (ReVs) is a policy decision made by the TCEQ. The application of an HQ is to consider potential cumulative and/or aggregate exposure. During the air permit review process, the predicted maximum ground level concentrations (GLC_{max}s) from the potential emissions are evaluated. The GLC_{max}s are predicted using the maximum allowable emission rates and worst-case meteorological conditions, which may or may not actually occur. Typically, when evaluating the GLC_{max} predicted to occur at a sensitive receptor, the concentration must be at or below the ESL. There is a lot of conservatism incorporated into the ESL (e.g., the ESL is 70% lower than the ReV) and layers of conservative assumptions are made in the worst-case modeling analysis itself. Thus, in the event that multiple facilities in an area emit the same chemicals, it is very unlikely that the maximum concentrations of emissions from other facilities emitting the same chemicals would occur at the same place. It is also very unlikely that the maximum concentrations of emissions from multiple chemicals from a facility and other facilities (if any) would occur at the same time and place. Furthermore, air concentrations of chemicals collected in air monitoring samples represent emissions from multiple chemicals and from different facilities and sources (i.e., can be both cumulative across chemicals and aggregate across sources and time). As

described in Section 1.3.1, exposure to chemicals in ambient air can be both cumulative across chemicals and aggregate across sources and time. Therefore, air monitoring data would indicate whether or not cumulative and/or aggregate exposure problems exist in an area of concern. Clarifying text was added to Section 1.4.2.

Pages 19-20. Table 1-3.
For Air Permitting, using the modeled maximum ground-level concentration (GLCmax) and the ESL to evaluate modeling data may be overly conservative. A statement in the narrative preceding the table about the inherent conservatism (or lack of it) in the air models would be appropriate.

Response: TCEQ agrees with the above comment. A clarifying statement has been added to page 20.

3.2 Chapter 2

Pages 30-31, Figures 2-1a, 2-1b:
The "reality check" of TCEQ personnel confirming the presence of the odor is a good idea. This is true even with the uncertainty in odor detection/perception discussed in the general comments.

Response: No response required.

Pages 33-38:
The exploration of curve fitting and prediction models for estimating odor thresholds is an example of quantitative structure-activity relationships (QSAR). OECD (2007) provides guidance on validation of QSAR methods. Information on the OECD QSAR project can be found at http://www.oecd.org/document/14/0,3746,en_2649_34365_33957015_1_1_1_1,00.html. TCEQ may wish to use this information to support its use of QSAR for estimation of odor thresholds.

Response: A reference to the OECD QSAR website was added to the QSAR section in Chapter 3.

3.3 Chapter 3

Page 46, lines 1-10:
This list is necessarily incomplete because for a data-rich chemical, there may be other steps involved in a credible dose-response assessment. However, this list provides an accurate summary of the general methodology.

Response: No response required.

Page 47, lines 16-37:
Log Kow is also known as Log P, where P stands for partition. Recently, both QSAR models and an assay for intestinal permeability using Caco-2 cells have been developed to estimate gastrointestinal absorption (Volpe, 2011; Press and Grandi, 2008). These are both used in physiologically-based pharmacokinetic models and TCEQ should mention these in this discussion.

Response: No revisions were made. The suggestion to mention methods to estimate gastrointestinal absorption is beyond the scope of the referenced section, which focuses on rule of thumb implications of log K_{OW} values for solubility and absorption as opposed to estimation methods.

Pages 46-48:
TCEQ should mention Lyman's Handbook of Chemical Property Estimation Methods (Lyman et al., 1990). This compendium of methods is often useful when chemical property data cannot be found.

Response: Appropriate text and the reference were added to Section 3.3.1.

Page 50, lines 26-27:
The text reads "… factors area dose-response relationship …" It should read "… factors are a dose-response relationship …" A space needs to be added.

Response: The suggested change was made.

Pages 50-58:
This discussion of adult-child differences, their incorporation into the risk assessment methodology and the inclusion of the white paper by Haber et al. as Appendix C add greatly to quality of this guidance document.

Response: No response required.

Page 59, line 20:
The study is by Adami not Adam.

Response: The suggested change was made.

Pages 65-69:
The discussion of adverse vs. non-adverse is well-reasoned and thoughtful but incomplete. Toxicity Testing in the 21st Century: A Vision and A Strategy (NRC, 2007) is not included in the references. This document calls for increasing substitution of in vitro assays and accompanying prediction models for traditional animal toxicity testing. This substitution will require enhanced knowledge of MOA and initially, the determination of adverse versus non-adverse may become more difficult as partial knowledge of MOA or mechanism is obtained. TCEQ may wish to mention this document, if only to say why it was not considered. One of the peer reviewers also mentioned this point.

One additional reference to consider in the discussion on adverse vs. non-adverse would be Boekelheide and Andersen (2010). These authors considered the ability to distinguish acceptable, homeostatic, or adaptive perturbations of a pathway from excessive or adverse perturbations to be the key challenge in the use of abundant high through-put data and increasing knowledge of toxicity pathways. A new approach they discussed was to evaluate dose-response relationships as functions of the probabilities of biological system failure, determined in a stepwise manner through assays that measure progressive perturbation along toxicity pathways.

Response: Appropriate text and references were added to Section 3.6.1.

Page 71, Figure 3-3:
Generally, UFs are applied to extrapolate from the NOAEL, not the NOEL, as the POD. TCEQ should change this figure.

Response: This figure has been modified as suggested.

Page 73, Figure 3-4.
The lower confidence limit (LCL) at both zero dose and the lowest positive dose appear to be less than zero. These presumably would be binomial confidence limits due to the quantal nature of the response and cannot go below zero. This may be due to the apparent lack of correspondence between y-axis ticks and the numbers associated with them. In any case, TCEQ should revise this graph.

Response: This figure has been modified to make the x- and y-axis tick labels more clear.

Pages 74-75:
Please also see Murrell et al. (1998) for another way to define the critical effect size (CES) as a 10% (or other value) response for continuous data.

Response: No change was made.

Page 75, lines 30-36:
The narrative notes that the vast majority of chemicals do not have sufficient data for the development of a biologically-based dose-response (BBDR) model. This situation is changing, and TCEQ should begin the sentence "At present" or "As of the publication of this guidance."

Response: The suggested change was made.

Page 76, Figure 3-5:
This was figure 3-7 in the 2011 draft. One of the peer reviewers indicated some confusion about the terms "protective" vs. "predictive." In the Response to Comments document, TCEQ indicated the figure came from EPA (1994), also known as the Inhalation Dosimetry Methodology. EPA (1994) attributes the figure to Conolly et al. (1990) and Andersen et al. (1992), both of which were reports of Chemical Industry Institute of Toxicology (CIIT) activities. A couple of sentences could be added to the narrative or the figure legend to clarify the meaning of these two words in the figure.

Response: No change was made. The relevance of this figure to the Guidelines does not focus on these two terms; rather, the associated text discusses the relevant and important points for purposes of the guidance.

Page 82, Figure 3-10:
The note under the figure indicates that TCEQ does not use a modifying factor (MF). Reviewer #12 asked for revision of this figure (3-12 in the draft), but all that was done was to add the one-sentence note. A new figure would be better.

Response: Figure 3-10 is provided as an example of a dose-response array from a USEPA publication and is not meant to accurately depict TCEQ's process (e.g., note that it refers to derivation of an RfC). No change was made.

Page 87, Equation 3-5:
It would be better to define, perhaps in the narrative, the normalizing factor (NF). While this is defined in EPA (1994), the reader must go to both Chapter 3 and Appendix G in EPA (1994) to figure out what is meant. This will require perhaps three or four sentences because NF can be defined differently for different chemicals.

Response: The suggested change was made in Section 3.9.2.

Page 88, lines 12-36:

The section on animal-to-human inhalation dosimetry is not entirely clear. First, minute ventilation (VE) and surface area (SA) should be defined. We assume these are minute ventilation and surface areas of the extra-thoracic (ET), tracheobronchial (TB) and pulmonary (PU) regions. However, the acronym "VE" without a subscript is used on the previous page to mean minute ventilation. Essentially, two different abbreviations are used for the same quantitative factor. This is confusing and should be corrected.

The papers from the workshop mentioned are published in the Journal of Toxicology and Environmental Health Part A, Volume 71, No. 3. Dr. John Stanek was not an author of any of the papers in that volume. Hence, it is difficult to obtain information on this section. Hence, for clarity, TCEQ should consider providing additional detail and sample calculations.

Response: Changes were made to Section 3.9.4.1 for clarification. A new reference, USEPA's 2011 *Status Report: Advances in Inhalation Dosimetry for Gases with Lower Respiratory Tract and Systemic Effects*, was added in place of the 2006 *Inhalation Risk Assessment in Children* presentation. VE and SA were removed from this section. However, additional details and sample calculations are beyond the scope of the guidance and are not provided.

Page 93 and elsewhere:

Increasingly, data are used to develop interspecies, intraspecies, and other types of extrapolation factors. On page 98 in section 3.11.1.2, the Guidelines indicate that chemical-specific adjustment factors (CSAFs) (WHO-IPCS, 2005) and data-derived extrapolation factors (DDEFs) (EPA, 2011) are different than a data-derived UF. TCEQ indicates that a data-derived UF may be based on information that shows variability in different species is small and when this is the case a smaller value could be used for interspecies uncertainty factors (UFA).

However, if quantitative information on the variability in both species is available, then a CSAF or DDEF can be developed. It is not clear what TCEQ means by "data-derived UF" in contrast to a CSAF or DDEF and a specific example is warranted.

This discussion also raises the question as to how much data is needed to call this number an extrapolation or adjustment factor rather than an uncertainty factor. TCEQ should clarify these concepts.

Response: A detailed comparison/contrast of the scientific nuances and data requirements of data-derived UFs, CSAFs, and DDEFs is beyond the scope of the guidance. Such a comparison is also unnecessary as users are referred to appropriate reference documents for detailed information and examples.

Pages 101-104:

The discussion of factors used to select the database uncertainty factor is excellent. Perhaps in a future guidance document, this could be expanded into a weight-of-evidence approach.

Response: No response required.

Page 108, lines 18-24:
These lines are comprised of two lengthy sentences that are unclear. This passage was also unclear in the 2011 draft, as noted by one of the reviewers. The only change was to add the phrase "assuming no available data on the sensitivity of animals versus humans." TCEQ should rewrite this section with shorter sentences and clarity in mind.

Response: Changes were made as suggested.

Page 113, lines 41-43 to Page 114, lines 1-2:
This sentence does not capture what Honma and Suda (1998) did in their publication. One of the peer reviewers also commented on the 2011 draft that the description of the correlations in Honma and Suda (1998) was unclear. The sentence was rewritten by adding a phrase. The 2012 Guidelines still do not accurately describe what was done by Honma and Suda (1998).

What Honma and Suda (1998) actually did was explore two different correlations—one correlation between the LD50 values for oral administration and LC50 inhalation values and the other correlation between the LD50 values for intraperitoneal administration and LC50 inhalation values. 146 chemicals were used to develop the correlations and separate evaluations were done for rats and mice. For both correlations, the LD50 and LC50 values were expressed as base ten logarithms. Both correlations improved when the LD50 values were expressed as mmol/kg rather than mg/kg and the LC50 values were expressed as ppm-hour or cumulative dose rather than ppm or concentration.

In rats, the correlation coefficient between the LD50 for oral administration and the LC50 was 0.742, and the correlation coefficient between the LD50 for intraperitoneal administration and the LC50 was 0.895. In mice the correlation coefficient between the LD50 for oral administration and the LC50 was 0.667, and the correlation coefficient between the LD50 for intraperitoneal administration and the LC50 was 0.726.

Response: Changes were made for clarification.

Pages 114-115:
TCEQ should mention that EPA's Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual (Part F, Supplemental Guidance for Inhalation Risk Assessment) (EPA, 2009) recommends against route-to-route extrapolation as presented on page 115. However, because of the requirement to develop ESLs for all chemicals, TCEQ may have to use these methods. In this regard, the Department of Defense Vapor Intrusion Handbook (TriServices Environmental Risk Assessment Workgroup , 2009) also calls for route-to-route extrapolation if no other options for inhalation toxicity factors are available.

Response: Relevant considerations are already included in this section for determining the appropriateness of route-to-route extrapolation. In actuality, rather than recommending against route-to-route extrapolation, Part F is essentially cautioning Superfund risk assessors not trained in evaluating these considerations against performing route-to-route extrapolations themselves without involvement of appropriate USEPA staff. In fact, some recent USEPA assessments have used it to derive toxicity factors (e.g., perchloroethylene). An additional sentence emphasizing the importance of evaluating the appropriateness of route-to-route extrapolation on a case-by-case basis was added.

Pages 116-119:
As well as QSAR and analog ID, read-across approaches might be useful here. These are detailed in ECETOC Technical Report No. 109 High information content technologies in support of read-across in chemical risk assessment (2010), Vink et al. (2010) and Wu et al. (2010). EPA also provides tools for Analog ID at http://www.epa.gov/oppt/sf/tools/aim.htm. TCEQ should consider citing these references and/or incorporating some of this information in a future update of the Guidelines.

Response: This information will be considered in a future update of the guidelines.

Pages 121-122:
The rounding procedure for significant figures (SFs) is incorrect. If the number next to the SF is less than 5 (not 5 or less), then round down. If greater than or equal to 5, then round up. So 13563 with 2 SFs would become 14000 and 13463 with 2 SFs would be 13000. Please see the following websites:

http://ostermiller.org/calc/significant_figures.html
http://www.chem.sc.edu/faculty/morgan/resources/sigfigs/index.html

More than one of the peer reviewers also commented on this error.

Response: This is technically not an error; it is an arithmetically accepted method of rounding. Final values derived in the DSD are conservative values, and to ensure a uniform and transparent approach TCEQ has chosen to round these values to two significant figures. The number 5 is in the middle of the range of numbers and may be rounded up or down. There are no rules requiring the rounding of 5 up, rather as long as the rounding convention used is stated it is up to the rounder which direction to go. The choice to round a number 5 down is called half round down. When developing the rounding convention for DSDs, TCEQ chose to use the half round down method. Clarifying text on this rounding convention was added.

3.4 Chapter 4

Page 130, lines 31-34:
This narrative is unclear and needs rewriting. Certainly, direct maternal toxicity during pregnancy can affect offspring in utero. Exposure during pregnancy and lactation can also affect the fetus directly, but these are not "effects of toxic agents on the maternal system."

Response: This section was clarified in response to the comment.

Page 134, lines 5-31:
The section on using TK and TD to inform about the type of study to use for a 24hr reference value (ReV) is very good.

Response: No response required. The section on 24-h ReV and AMCV values was removed from the Guidelines. See explanation below.

Page 138, lines 10-12:
With a new exposure event every day, each day represents a potential toxic effect. Hence, TCEQ should add the word "potential" so that lines 10-11 read: "… each new day represents a potential toxic effect …"

Response: The section on 24-h ReV and AMCV values was removed from the Guidelines. Since this section introduces new concepts, TCEQ has decided to submit this section as a case study at a future Alliance for Risk Assessment workshop on *Beyond Science and Decisions: From Problem Formulation to Dose-Response* prior to adopting it in the Guidelines. This will allow for review by the *Beyond Science and Decisions* Science Panel. The ultimate goal is for TCEQ to published guidance on 24-h ReV and AMCV values through a white paper or in a peer-reviewed journal article.

Pages 146-147:

TCEQ should consider adding a comparison of the threshold of concern (TOC) approach and the NOAEL-to-LC50 (N-L) ratio approach. An example might make these more clear to the reader. TCEQ should consider n-hexane because it is a well studied chemical and not in the database in Grant et al. (2007), and thus would be new.

The 4-hour LC50 in rats for n-hexane is 48,000 ppm or about 170,000 mg/m3. Hence, according to the GHS classification in Table 1 of Grant et al. (2007), n-hexane would be a Category 5 gas and the 10th percentile TOC concentration would be 1 mg/m3 or 0.28 ppm. Multiplying the LC50 by the 10th percentile N-L factor of 8.3E-05 would give an estimated NOAEL of 3.98 ppm.

These values are respectively about 100 and 10 fold lower than the lowest benchmark concentration lower confidence limit (BMCL) value of 28 ppm from Table 5-1 in EPA's Toxicology Review of n-hexane (EPA, 2005b). The endpoint considered was a decrease in nerve conduction velocity from a rat study. This same effect is also observed in humans and the MOA is well established.

Whether n-hexane is used or not, an example in the Guidelines could support the choice to use the N-L method.

Response: No changes were made. Phillips et al. 2011, which is referenced and summarized in this section, does a comparison such as is suggested here. The end result was that, in general, the TOC method was more conservative than the N-L ratio approach. While n-hexane was not in the Grant et al. 2007 database, it was included in the Phillips et al. 2011 comparison as a n-hexane DSD was completed in October of 2007 (http://www.tceq.state.tx.us/assets/public/implementation/tox/dsd/final/hexane,n-_100-54-3_final_10_15_2007.pdf). In that DSD, a comparison of the TOC and N-L ratio approach was described. The TOC and N-L ratio approaches resulted in a Tier II acute generic ESL for n-hexane of 1,000 ug/m3 (280 ppb) and 6,200 ug/m3 (1,800 ppb), respectively. A weight-of-evidence approach is then used to determine which Tier II acute generic ESL is the most appropriate. Use of the EPA RfC for comparison here is inappropriate. The TOC and N-L ratio approaches are for acute values, not chronic, and the EPA RfC value is a chronic value. For examples of the TOC and N-L ratio approaches, readers should review chemicals that have gone through the DSD process.

3.5 Chapter 5

Page 152, Equation 5-3:

The stated units of daily dose (DD) in Eq. 5-3 should be mg/day for the units of point of departure adjusted for human equivalent concentration (PODHEC) to be correct. In addition, for the units to be

correct, DD is really an intake not a dose. It would be most correct to rename DD as ADI meaning "average daily intake." Then the units of mg/day and the name of the quantity would be correct.

Response: The units were corrected to mg/day. Since ADI is defined as acceptable daily intake in other parts of the Guidelines, and to avoid confusion, no change to the numerator was made in the equation.

Page 153, Equation 5-4:
The units are correct in Eq. 5-4. However, it would be most correct to rename DD to ADI for the reasons stated in the previous comment.

Response: See previous response.

It should also be noted that body weight (BW) will change over time and that the value for BW is assumed to be the average over the chronic time period. TCEQ should provide additional discussion on this point.

Response: No additional discussion/text was added. The TCEQ makes reference to USEPA 1988, which provides additional discussion regarding time-weighted average body weights for laboratory rodents and other laboratory mammals.

Page 160, lines 20-21:
The narrative reads: "… a compelling toxicity profile without great uncertainty …." This seems redundant and overstated. It would be better to characterize an evaluation based on limited data as "… a toxicity profile with the least possible uncertainty…"

Response: The sentence was adjusted to the following: "…or if there is great uncertainty in the toxicity assessment based on scientific judgment..."

Pages 163-165, Section 5.6:
When an RfD is needed for a limited toxicity data (LTD) chemical, in silico approaches mentioned earlier, such as SAR/QSAR, read-across or analog ID, should also be considered (e.g., Vink et al., 2010; Eriksson et al., 2003). In addition, while the use of in vitro data on LTD chemicals for hazard prediction (e.g., Martin et al., 2011; Judson et al., 2011; EPA's ToxCast™ Program) is not yet feasible, much effort is being expended toward this goal and TCEQ may want to mention this effort.

Response: No changes were made. The TCEQ Guidelines are updated periodically to incorporate appropriate procedures that have been developed and well established within the scientific regulatory community for chemical risk assessment, if it is applicable.

Page 168, lines 2-22:
TCEQ should also mention the route of exposure (e.g., oral, inhalation, dermal) as a consideration when evaluating the human relevance of tumors observed in animals.

Response: The 2005 Cancer Guidelines (Chapter 2 of USEPA 2005a) contain a detailed discussion of hazard identification based on a chemical's MOA using a WOE approach. Tumor data from different routes of exposure are discussed as well as relevance of tumors observed in animals to humans, so a detailed discussion is not included in the TCEQ Guidelines. However, the TCEQ does provide a brief

discussion and some examples of considerations on different routes of exposure when assessing the carcinogenic potential of a toxicant based on a WOE approach (Section 5.7.1).

Page 170, lines 11-13:
The Surveillance Epidemiology and End Results (SEER) database is mentioned on pages 199 and 227. TCEQ should also mention it here as a source from which to obtain background cancer rates.

Response: No additional discussion/text was added.

Page 173, lines 15-28:
TCEQ is correct that there is already substantial evidence of nonlinearity of the cancer response of potentially mutagenic chemicals. Additional references in this regard are Williams et al. (1996, 1999), Fukushima et al. (2002), Waddell (2003), Waddell et al. (2006), and Swenberg et al. (2011).

As noted on lines 20-21, the linear no-threshold assumption has been used for the dose-response of genotoxic agents that produce cancer. TCEQ should be commended for inclusion of this balanced and open-minded discussion in its guidance. TCEQ should consider including some of the information in the general comments above to support the choice to take a careful look at low dose linear extrapolation as a default procedure.

Response: No additional discussion/text was added.

Page 174, lines 2-8:
The idea for increased early life susceptibility to cancer came from studies conducted by Maltoni at the Ramazzini Institute on vinyl chloride. In 2011, EPA and the National Toxicology Program (NTP) released preliminary findings of a review of pathology and animal husbandry procedures at the Ramazzini Institute. Their report recommended a quality-assurance review and a pathology working group (PWG) review of rodent bioassays of methyl-tertiary butyl ether, methanol, ethyl-tertiary butyl ether, vinyl chloride, and acrylonitrile. These PWG reviews have not yet been completed.

Hence, TCEQ should closely monitor emerging issues in the evaluation of early life exposure to carcinogens (Page 174, lines 6-7).

Response: No response required.

Page 174, lines 10-13:
The narrative currently reads: "A mutagenic MOA is one that produces cancer via irreversible changes to DNA, a determination that is to be reached by a WOE approach as described below and in additional detail in Section 2.3.5 of the Cancer Guidelines (USEPA 2005a)."

TCEQ should substitute the word "initiates" for "produces" and insert the words "early" and "primary" and "sequence." The narrative should read: "… mutagenic MOA is one that initiates cancer via early irreversible changes in primary DNA sequence … "

It is important to clarify that a mutagenic MOA must be driven specifically by an <u>early</u> and <u>causal</u> mutation. This is made clear in later discussion but should also be clear here, as all cancers require mutagenic events at some point. DNA replication is necessary for mutations to manifest. It is not just

that replication occurs before repair, but also that the repair itself can be faulty and result in a heritable change.

Dr. Toby Rossman, one of the peer reviewers of the 2011 draft of the TCEQ Guidelines, who also served as a peer reviewer for EPA's Toxicological Review of Hexavalent Chromium, wrote:

> *Genotoxic is not the same as mutagenic, … Standard genotoxicity assays were not designed to inform specific modes of tumor induction. With the exception of mutagenesis, these other assays (non-mutagenic assays) do not measure heritable events, but rather measure evidence of DNA damage or its repair.*
>
> *Non-mutagenic assays include chromosome aberrations, micronuclei, comet assays, DNA lesion measurements, and DNA repair assays. These assays are useful for hazard identification or as biomarkers of exposure. They provide only supportive evidence that mutagenesis might be a MOA. DNA damage per se does not inform us about eventual heritable change, which is the true issue. Most (but not all) mutagens cause heritable changes in DNA sequences by causing damage to DNA (pre-mutagenic lesions) that is converted to mutation after cell division.*

This is consistent with the narrative on page 175 discussing the difference between genotoxicity and mutagenicity.

Hence, TCEQ should remove the citations to McCarroll et al. (2008) and McCarroll et al. (2010). Both these papers confuse genotoxicity evidenced by DNA damage with mutagenicity evidenced by heritable changes in DNA sequence. The assays cited in Tables 1 and 2 of McCarroll et al. (2008) all measure DNA damage not mutation with a single exception—Ammenheuser et al. (1988) measured mutations in the HPRT locus in lymphocytes from multiple sclerosis patients. McCarroll et al. (2008) cited a review article by Anderson et al. (1995) instead of the primary reference. The assays cited in Table 2 of McCarroll et al. (2010) were all genotoxicity assays, as noted in the table heading, as opposed to mutation assays.

Response: To greater emphasize that a mutagenic MOA must be driven specifically by an early and causal mutation, language was added to indicate that since the determination that a carcinogen is capable of producing mutation is insufficient alone to conclude that it operates via a mutagenic MOA, to demonstrate a mutagenic MOA the WOE must further indicate that it induces an early mutation in target tissue that causes cancer (i.e., the ability to produce mutagenicity alone or mere genotoxicity does not lead to a presumption of a mutagenic MOA as this only supports that mutagenesis might be an MOA). Additionally, the referenced citations were removed from this section.

Page 184, line 17:
The heading for section 5.7.4.4 would be better as: "Carcinogens with a Non-Mutagenic MOA or an Unknown MOA."

Response: The TCEQ agrees and the heading was revised accordingly.

3.6 Chapter 6

Page 186, lines 4-14:
Depending on the endpoints chosen by TCEQ for the toxicity factors for TCDD, the toxicity equivalency factor (TEF) values used in Van den Berg et al. (2006) may or may not be applicable. These TEFs are based on a subjective evaluation, albeit by experts, of both in vivo and in vitro effects (Haws et al., 2006; Van den Berg et al., 2006).

Hence, the subjective consensus values of the TEFs represent a mixed bag of effects, some measured in vivo, which include TK and TD effects, and some measured in vitro, which include only TD effects.

For example, in human primary hepatocytes in vitro, the estimates of the TEFs for 4-PeCDF (2,3,4,7,8-pentachlorodibenzofuran) and for TCDF (2,3,7,8-tetrachlorodibenzofuran) differ considerably depending on whether gene expression or enzyme induction is used as the response (Budinsky et al., 2010). From in vivo studies in mink, TCDF is metabolized very quickly and the consensus TEF of 0.1 does not reflect this rapid metabolism (Simon, T. et al., 2011).

These same considerations may also apply to other chemical classes. For example, Fisher et al., (2011) recently developed a TEF scheme for the acute toxicity of PAHs in sediment. Simon et al. (2007) and Yang et al. (2010) have developed TEF schemes for non-dioxin-like PCBs. TCEQ appears to be prepared to evaluate other TEF schemes (e.g., Collins et al., 1998) and possibly develop its own TEF schemes (Page 186, lines 11-14).

Response: No response required.

3.7 Chapter 7

Page 197, lines 30-33:
The last sentence of section 7.7.1 indicates TCEQ will adopt a linear low dose extrapolation model for epidemiologic data in cases where a BBDR cannot be justified. What about the rare but possible case where a biomarker is involved or no effect is consistently observed at low exposures in well-conducted epidemiologic studies?

There may also be a biomarker that shows no change below a particular dose level. For example, dioxin-like chemicals induce CYP1A in humans, non-human primates and rodents. In humans, no induction above background is observed at blood levels below 1000 ppt-lipid which would correspond to a dose in humans of about 400 pg/kg/d.

Each chemical or chemical class is different, but the available information may be sufficient to support a threshold dose-response but not a BBDR. TCEQ should acknowledge this possibility in the narrative.

Response: No changes were made. The second paragraph in Section 7.7 already states that if MOA information is available to indicate a threshold dose-response (but not a BBDR) than the TCEQ will consider the information: "In the absence of a scientifically defensible and biologically-based dose-response model or sufficient MOA information indicating a nonlinear dose-response relationship at low doses, . . ."

Page 212, line 212:
The narrative includes "Texans of a specified race" as a possible inference population. TCEQ should remove this statement unless race corresponds to a specific biological factor that alters susceptibility. It would be more appropriate to consider ethnic characteristics in development support documents (DSDs) for specific chemicals but only when there are reliable scientific data on differences in sensitivity or susceptibility conferred by ethnic characteristics.

Response: It appears this comment is actually referring to line 41. This sentence was revised to exclude examples of specific groups other than Texans, since that is the inference population generally used by the TCEQ.

Page 213, lines 12-16:
The first sentence in section 7.9.1 is awkward and too long. TCEQ should rewrite for clarity.

Response: The first sentence was rewritten to improve clarity.

Page 220, line 7:
The narrative should read "… is or is not …" Please add the word "not".

Response: The suggested change was made.

Page 221, line 27:
Presumably (TCEQ, 2012) refers to the arsenic DSD. It is not included in the reference list.

Response: The TCEQ (2012) reference has been included in the references.

Pages 226-2267:
The section on "Reality Checks" is excellent.

Response: No response required.

Page 227, lines 37-41:
In this section on Uncertainty Analysis, the sentence reading "Variability refers to differences that cannot be controlled by statistical modeling" is not accurate taken out of context. The context here is a comparison of uncertainty and variability. In that regard, uncertainty also cannot be controlled by statistical modeling. The section would read better if this sentence were removed.

Response: The suggested change was made.

4 References

Ammenheuser, M.M., Ward Jr., J.B. , Whorton Jr., E.B. , Killian, J.M., and Legator, M.S. (1988) Elevated frequencies of 6-thioguanine-resistant lymphocytes in multiple sclerosis patients treated with cyclophosphamide: a prospective study. Mutation Research/Genetic Toxicology, Volume 204, Issue 3, March, Pages 509-520

Anderson, E. L. (1983), Quantitative Approaches in Use to Assess Cancer Risk. Risk Analysis, 3: 277–295. doi: 10.1111/j.1539-6924.1983.tb01396.x

Beamer P.I., Canales R.A., Ferguson A.C., Leckie J.O., and Bradman A., (2012) Relative pesticide and exposure route contribution to aggregate and cumulative dose in young farmworker children. Int J Environ Res Public Health 9, 73-96.

Bell, I. R., Schwartz, G. E., Peterson, J. M. and Amend, D. (1993). Self-reported illness from chemical odors in young adults without clinical syndromes or occupational exposures. Arch Environ Health 48, 6-13.

Bell, I. R., Miller, C. S., Schwartz, G. E., Peterson, J. M. and Amend, D. (1996). Neuropsychiatric and somatic characteristics of young adults with and without self-reported chemical odor intolerance and chemical sensitivity. Arch Environ Health 51, 9-21.

Bhalli J.A., Pearce M.G., Dobrovolsky V.N., and Heflich R.H., (2011) Manifestation and persistence of Pig-a mutant red blood cells in C57BL/6 mice following single and split doses of N-ethyl-N-nitrosourea. Environ Mol Mutagen 52, 766-773.

Boekelheide K. and Andersen M.E., (2010) A mechanistic redefinition of adverse effects - a key step in the toxicity testing paradigm shift. ALTEX 27, 243-52.

Bogen K.T. (2011) Generic Hockey-Stick Model for Estimating Benchmark Dose and Potency: Performance Relative to BMDS and Application to Anthraquinone. Dose Response 9, 182-208.

Bol S.A., van Steeg H., van Oostrom C.T., Tates A.D., Vrieling H., de Groot A.J., Mullenders L.H., van Zeeland A.A., and Jansen J.G., (1999) Nucleotide excision repair modulates the cytotoxic and mutagenic effects of N-n-butyl-N-nitrosourea in cultured mammalian cells as well as in mouse splenocytes in vivo. Mutagenesis 14, 317-22.

Budinsky R.A., LeCluyse E.L., Ferguson S.S., Rowlands J.C., and Simon T., (2010) Human and rat primary hepatocyte CYP1A1 and 1A2 induction with 2,3,7,8-tetrachlorodibenzo-p-dioxin, 2,3,7,8-tetrachlorodibenzofuran, and 2,3,4,7,8-pentachlorodibenzofuran Toxicol Sci 118, 224-35

Calabrese E.J. (2009) The road to linearity: why linearity at low doses became the basis for carcinogen risk assessment Arch Toxicol 83, 203-25.

Calabrese E.J. (2011a) Key studies used to support cancer risk assessment questioned Environ Mol Mutagen

Calabrese E.J. (2011b) Muller's Nobel lecture on dose-response for ionizing radiation: ideology or science? Arch Toxicol

Cammerer Z., Bhalli J.A., Cao X., Coffing S.L., Dickinson D., Dobo K.L., Dobrovolsky V.N., Engel M., Fiedler R.D., Gunther W.C., Heflich R.H., Pearce M.G., Shaddock J.G., Shutsky T., Thiffeault C.J., and Schuler M., (2011) Report on stage III Pig-a mutation assays using N-ethyl-N-nitrosourea-comparison with other in vivo genotoxicity endpoints. Environ Mol Mutagen 52, 721-30.

Chen J.B., Dass S.B., Burkhart J.G., and Heflich R.H., (1998) Sensitivity of the phiX174 am3 allele in relation to the endogenous Hprt gene for detecting mutation in transgenic mice. Environ Mol Mutagen 32, 229-35.

Collins, J. F., Brown, J. P., Alexeeff, G. V., & Salmon, A. G. (1998). Potency equivalency 12 factors for some polycyclic aromatic hydrocarbons and polycyclic aromatic 13 hydrocarbon derivatives. Regul Toxicol Pharmacol, 28(1), 45-54. doi: 14 10.1006/rtph.1998.1235

Crump K.S., Guess H.A., and Deal K.L., (1977) Confidence intervals and test of hypotheses concerning dose response relations inferred from animal carcinogenicity data. Biometrics 33, 437-451.

Crump K.S., Hoel D.G., Langley C.H., and Peto R., (1976) Fundamental carcinogenic processes and their implications for low dose risk assessment Cancer Res 36, 2973-9.

Crump K.S. (1984) An improved procedure for low-dose carcinogenic risk assessment from animal data J.Environ Pathol.Toxicol Oncol. 5, 339-348.

Doty R.L. and Cameron E.L., (2009) Sex differences and reproductive hormone influences on human odor perception. Physiol Behav 97, 213-28.

Elsner R.J. (2001) Environment and medication use influence olfactory abilities of older adults. J Nutr Health Aging 5, 5-10.

Eriksson L., Jaworska J., Worth A.P., Cronin M.T., McDowell R.M., and Gramatica P., (2003) Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. Environ Health Perspect 111, 1361-75.

European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) Technical Report No. 109. High information content technologies in support of read-across in chemical risk assessment at http://www.ecetoc.org/technical-reports

Environmental Protection Agency (1994) Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry. (EPA/600/8-90/066F). Washington, D.C.

Environmental Protection Agency (2001) Risk Assessment Guidance for Superfund. Volume III. Probabilistic Risk Assessment

Environmental Protection Agency (2005a) Guidelines for Carcinogen Risk Assessment EPA/630/P-03/001B

Environmental Protection Agency (EPA) (2005b) Toxicological Review of n-Hexane (CAS No. 110-54-3). EPA/635/R-03/012.

Environmental Protection Agency (EPA) (2009) Risk Assessment Guidance for Superfund

Volume I: Human Health Evaluation Manual (Part F, Supplemental Guidance for Inhalation Risk Assessment). EPA-540-R-070-002.

Environmental Protection Agency (EPA) (2011) DRAFT Guidance on Applying Quantitative Data to Develop Data-Derived Extrapolation Factors for Interspecies and Intraspecies Extrapolation. EPA/100/J-11/001.

Fisher, T. T., Law, R. S., Rumney, H. S., Kirby, M. F. and Kelly C. (2011) Towards a scheme of toxic equivalency factors (TEFs) for the acute toxicity of PAHs in sediment. Ecotoxicol Environ Saf 74, 2245-51

Fukushima S., Wanibuchi H., Morimura K., Wei M., Nakae D., Konishi Y., Tsuda H., Uehara N., Imaida K., Shirai T., Tatematsu M., Tsukamoto T., Hirose M., Furukawa F., Wakabayashi K., and Totsuka Y., (2002) Lack of a dose-response relationship for carcinogenicity in the rat liver with low doses of 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline or N-nitrosodiethylamine Jpn J Cancer Res 93, 1076-82.

Gocke E. and Wall M., (2009) In vivo genotoxicity of EMS: statistical assessment of the dose response curves. Toxicol Lett 190, 298-302.

Grant, R. L., Kadlubar, B. J., Erraguntla, N. K., and Honeycutt, M. (2007) Evaluation of acute inhalation toxicity for chemicals with limited toxicity information. Regul Toxicol Pharmacol 47, 261-273.

Grant, R. L., Haney, J., Curry, A. L., & Honeycutt, M. (2009). Development of a unit risk 22 factor for 1,3-butadiene based on an updated carcinogenic toxicity 23 assessment. Risk Anal, 29(12), 1726-1742. doi: 10.1111/j.1539-24 6924.2009.01302.x

Guess H., Crump K., and Peto R., (1977) Uncertainty estimates for low-dose-rate extrapolations of animal carcinogenicity data Cancer Res 37, 3475-83.

Hartley H.O. and Sielken R.L., (1977) Estimation of "safe doses" in carcinogenic experiments. Biometrics 33, 1-30.

Haws L.C., Su S.H., Harris M., DeVito M.J., Walker N.J., Farland W.H., Finley B., and Birnbaum L.S., (2006) Development of a refined database of mammalian relative potency estimates for dioxin-like compounds Toxicol.Sci. 89, 4-30.

Honma, T., & Suda, M. (1998). Correlation of lethal doses of industrial chemicals between oral or intraperitoneal administration and inhalation exposure. Ind Health, 36(3), 273-281.

Jarabek, A.M., Pottenger, L.H., Andrews, L.S., Casciano, D., Embry, M.R., Kim, J.H., Preston, R.J., Reddy, M.V., Schoeny, R., Shuker, D., Skare, J., Swenberg, J., Williams, G.M., Zeiger, E. (2009). 2009. Creating context for the use of DNA adduct data in cancer risk assessment: I. Data organization. Crit. Rev. Toxicol. 39: 659-678.

Jones, T. F., Craig, A. S., Hoy, D., Gunter, E. W., Ashley, D. L., Barr, D. B., Brock, J. W. and Schaffner, W. (2000). Mass psychogenic illness attributed to toxic exposure at a high school. N Engl J Med 342, 96-100, 10.1056/NEJM200001133420206.

Judson R.S., Kavlock R.J., Setzer R.W., Hubal E.A., Martin M.T., Knudsen T.B., Houck K.A., Thomas R.S., Wetmore B.A., and Dix D.J., (2011) Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. Chem Res Toxicol 24, 451-62.

Lutz W.K. and Lutz R.W., (2009) Statistical model to estimate a threshold dose and its confidence limits for the analysis of sublinear dose-response relationships, exemplified for mutagenicity data. Mutat Res 678, 118-22.

Lyman, W.J., Rosenblatt, D.H., and Reehl, W.J. (1990) Handbook of Chemical Property Estimation Methods: Environmental Behavior of Organic Compounds. American Chemical Society.

Mantel N., Bohidar N.R., Brown C.C., Ciminera J.L., and Tukey J.W., (1975) An improved Mantel-Bryan procedure for "safety" testing of carcinogens. Cancer Res 35, 865-72.

Mantel N., Heston W.E., And Gurian J.M., (1961) Thresholds in linear dose-response models for carcinogenesis. J Natl Cancer Inst 27, 203-15.

Mantel N. and Schneiderman M.A., (1975) Estimating "safe" levels, a hazardous undertaking. Cancer Res 35, 1379-86.

Martin M.T., Knudsen T.B., Reif D.M., Houck K.A., Judson R.S., Kavlock R.J., and Dix D.J., (2011) Predictive model of rat reproductive toxicity from ToxCast high throughput screening. Biol Reprod 85, 327-39.

Masaoka Y., Satoh H., Kawamura M., and Homma I., (2008) Respiratory responses to olfactory stimuli in Parkinson's disease. Respir Physiol Neurobiol 161, 136-41.

Matsunaga M., Isowa T., Yamakawa K., Kawanishi Y., Tsuboi H., Kaneko H., Sadato N., Oshida A., Katayama A., Kashiwagi M., and Ohira H., (2011) Psychological and physiological responses to odor-evoked autobiographic memory. Neuro Endocrinol Lett 32, 774-80.

McCarroll N., Keshava N., Cimino M., Chu M., Dearfield K., Keshava C., Kligerman A., Owen R., Protzel A., Putzrath R., and Schoeny R., (2008) An evaluation of the mode of action framework for mutagenic carcinogens case study: Cyclophosphamide. Environ Mol Mutagen 49, 117-31.

McCarroll N., Keshava N., Chen J., Akerman G., Kligerman A., and Rinde E., (2010) An evaluation of the mode of action framework for mutagenic carcinogens case study II: chromium (VI). Environ Mol Mutagen 51, 89-111

Moberg P.J., Arnold S.E., Doty R.L., Kohler C., Kanes S., Seigel S., Gur R.E., and Turetsky B.I., (2003) Impairment of odor hedonics in men with schizophrenia. Am J Psychiatry 160, 1784-9.

Moffatt, S., Mulloli, T. P., Bhopal, R., Foy, C. and Phillimore, P. (2000). An exploration of awareness bias in two environmental epidemiology studies. Epidemiology 11, 199-208.

Murrell J.A., Portier C.J., and Morris R.W., (1998) Characterizing dose-response: I: Critical assessment of the benchmark dose concept Risk Anal 18, 13-26.

National Research Council (NRC) (2007b) Toxicity Testing in the 21st Century: A Vision and A Strategy. ISBN: 0-309-10993-0.

Niimura Y. (2009) Evolutionary dynamics of olfactory receptor genes in chordates: interaction between environments and genomic contents. Hum Genomics 4, 107-18.

Organization for Economic Co-operation and Development (OECD) (2007) Guidance Document on the Validation of (Quantitative)Structure-Activity Relationships [(Q)SAR] Models at http://www.oecd.org/officialdocuments/displaydocumentpdf?cote=env/jm/mono(2007)2&doclanguage=en accessed March 30, 2012.

Olender T., Lancet D., and Nebert D.W., (2008) Update on the olfactory receptor (OR) gene superfamily. Hum Genomics 3, 87-97.

Page, L. A., Keshishian, C., Leonardi, G., Murray, V., Rubin, G. J. and Wessely, S. (2010). Frequency and predictors of mass psychogenic illness. Epidemiology 21, 744-7, 10.1097/EDE.0b013e3181e9edc4.

Poncelet J., Rinck F., Bourgeat F., Schaal B., Rouby C., Bensafi M., and Hummel T., (2010) The effect of early experience on odor perception in humans: psychological and physiological correlates. Behav Brain Res 208, 458-65.

Pottenger L.H., Schisler M.R., Zhang F., Bartels M.J., Fontaine D.D., McFadden L.G., and Bhaskar Gollapudi B., (2009) Dose-response and operational thresholds/NOAELs for in vitro mutagenic effects from DNA-reactive mutagens, MMS and MNU. Mutat Res 678, 138-47.

Pottenger L.H. and Gollapudi B.B., (2010) Genotoxicity testing: moving beyond qualitative "screen and bin" approach towards characterization of dose-response and thresholds. Environ Mol Mutagen 51, 792-9.

Press B. and Di Grandi D., (2008) Permeability for intestinal absorption: Caco-2 assay and related issues. Curr Drug Metab 9, 893-900.

Rhomberg L.R. (2011a) Practical Risk Assessment and Management Issues Arising were we to Adopt Low-Dose Linearity for all Endpoints. Dose Response 9, 144-57.

Rhomberg L.R., Goodman J.E., Haber L.T., Dourson M., Andersen M.E., Klaunig J.E., Meek B., Price P.S., McClellan R.O., and Cohen S.M., (2011b) Linear low-dose extrapolation for noncancer heath effects is the exception, not the rule Crit Rev Toxicol 41, 1-19.

Simon, G. E., Daniell, W., Stockbridge, H., Claypoole, K. and Rosenstock, L. (1993). Immunologic, psychological, and neuropsychological factors in multiple chemical sensitivity. A controlled study. Ann Intern Med 119, 97-103.

Simon T., Britt J.K., and James R.C., (2007) Development of a neurotoxic equivalence scheme of relative potency for assessing the risk of PCB mixtures. Regul Toxicol Pharmacol 48, 148-170

Simon, T, Rowlands JC, Budinsky RA (2011) The Toxic Equivalence Factor for 2,3,7,8-Tetrachlorodibenzofuran (TCDF): Toxicokinetic and Toxicodynamic Considerations. Organohalogen Compounds

Spiegelhalter, D. Pearson, M, and Short, I. (2011) Visualizing Uncertainty About the Future. Science 333, 1393-400.

Staudenmayer, H., Binkley, K. E., Leznoff, A. and Phillips, S. (2003a). Idiopathic environmental intolerance: Part 1: A causation analysis applying Bradford Hill's criteria to the toxicogenic theory. Toxicol Rev 22, 235-46, 2245 [pii].

Staudenmayer, H., Binkley, K. E., Leznoff, A. and Phillips, S. (2003b). Idiopathic environmental intolerance: Part 2: A causation analysis applying Bradford Hill's criteria to the psychogenic theory. Toxicol Rev 22, 247-61, 2246 [pii].

Swenberg J.A., Lu K., Moeller B.C., Gao L., Upton P.B., Nakamura J., and Starr T.B., (2011) Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment Toxicol Sci 120 Suppl 1, S130-45

Tarlo, S. M., Poonai, N., Binkley, K., Antony, M. M. and Swinson, R. P. (2002). Responses to panic induction procedures in subjects with multiple chemical sensitivity/idiopathic environmental intolerance:

understanding the relationship with panic disorder. Environ Health Perspect 110 Suppl 4, 669-71, sc271_5_1835 [pii].

ten Berge, W. F., Zwart, A., & Appelman, L. M. (1986). Concentration-time mortality response relationship of irritant and systemically acting vapours and gases. Journal of Hazardous Materials, 13(3), 301-309.

TriServices Risk Assessment Work Group (2009) DOD Vapor Intrusion Handbook

Van den Berg M., Birnbaum L.S., Denison M., De V.M., Farland W., Feeley M., Fiedler H., Hakansson H., Hanberg A., Haws L., Rose M., Safe S., Schrenk D., Tohyama C., Tritscher A.,

Tuomisto J., Tysklind M., Walker N., and Peterson R.E., (2006) The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicol.Sci. 93, 223-241.

van Zeeland A.A., de Groot A.J.L., Mohn G.R., van Steeg H., van Oostrom C., van Duijn-Goedhart A.-M., Mullenders L.F.H., and Jansen J.G., (2008) Reduced methylation-induced mutagenesis in rat splenocytes in vivo by sub-chronic low dose exposure to N-metyl-N-nitrosourea. Mutat Res 640, 131-8.

Vink R., Mikkers J., Bouwman T., Marquart H., and Kroese D., (2010) Use of read-across and tiered exposure assessment in risk assessment under REACH – A case study on a phase-in substance. Regulatory Toxicology and Pharmacology 58, 64-71

Volpe D.A. (2011) Drug-permeability and transporter assays in Caco-2 and MDCK cell lines. Future Med Chem 3, 2063-77.

Waddell W.J. (2003) Thresholds in chemical carcinogenesis: what are animal experiments telling us? Toxicol Pathol 31, 260-2.

Waddell W.J., Fukushima S., and Williams G.M., (2006) Concordance of thresholds for carcinogenicity of N-nitrosodiethylamine Arch Toxicol 80, 305-9.

Waszak SM, Hasin Y, Zichner T, Olender T, Keydar I, et al. (2010) Systematic Inference of Copy-Number Genotypes from Personal Genome Sequencing Data Reveals Extensive Olfactory Receptor Gene Content Diversity. Systematic Inference of Copy-Number Genotypes from Personal Genome Sequencing Data Reveals Extensive Olfactory Receptor Gene Content Diversity. PLoS Comput Biol 6(11): e1000988.

Williams G.M., Iatropoulos M.J., Jeffrey A.M., Luo F.Q., Wang C.X., and Pittman B., (1999) Diethylnitrosamine exposure-responses for DNA ethylation, hepatocellular proliferation, and initiation of carcinogenesis in rat liver display non-linearities and thresholds Arch Toxicol 73, 394-402.

Williams G.M., Iatropoulos M.J., Wang C.X., Ali N., Rivenson A., Peterson L.A., Schulz C., and Gebhardt R., (1996) Diethylnitrosamine exposure-responses for DNA damage, centrilobular cytotoxicity, cell proliferation and carcinogenesis in rat liver exhibit some non-linearities Carcinogenesis 17, 2253-8

World Health Organization/International Programme on Chemical Safety (WHO-IPCS) (2005) Chemical-Specific Adjustment Factors For Interspecies Differences And Human Variability: Guidance Document For Use Of Data In Dose/Concentration–Response Assessment. WHO, Geneva.

Wu S., Blackburn K., Amburgey J., Jaworska J., and Federle T., (2010) A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. Regul Toxicol Pharmacol 56, 67-81.

Yang, J. M., Salmon, A. G., Marty, M. A. (2010) Development of TEFs for PCB congeners by using an alternative biomarker--thyroid hormone levels. Regul Toxicol Pharmacol 56, 225-36

Public Comment 2

Science Panel of the Alliance for Risk Assessment fourth workshop on *Beyond Science and Decisions: From Problem Formulation to Dose-Response* (May 22-24, 2012, Austin, Texas). The meeting report may be accessed online here: http://www.allianceforrisk.org/ARA_Dose-Response.htm#panel.

Suggestions for improvement of the methodology in Section 3.13

(1) include a graphic presentation to aid in communication;

(2) be clear that the robustness of the effect level for a given chemical depends on that chemical's database;

(3) focus groups can be useful for testing the effectiveness of the approach for risk communication;

(4) clearly communicate data available and data gaps regarding potential sensitive groups;

(5) modify the term to something like "known effect level" or "observed effect level," to show the complementarity to safe doses and communicate that effects may occur at lower levels; and

(6) emphasize the importance of presenting appropriate context of why the health-protective value is lower, which the authors noted would be in the accompanying narrative.

Responses: Regarding (1), graphical depictions of ESLs, PODs, observed effects levels, etc. may be provided in DSDs but this decision will be made on a case-by-case basis.

Concerning comments (2), (4), and (6), the Guidelines were revised to clarify that DSDs should include a narrative to put the observed effect levels and their associated uncertainties and caveats (e.g., data limitations and gaps, potential inter/intraspecies differences in sensitivity), for comparison to corresponding health-protective values, into proper context.

The suggestion in comment (3) is beyond the scope of the Guidance, which is to document the methodology used to derive observed effect levels.

Consistent with comment (5), the term was revised to "observed effect level."