# CASE STUDY SUMMARY
## Methods for Deriving Inhalation Effect Levels for Comparison to Health-Protective Values (Workshop IV)

**Roberta L. Grant, Allison Jenkins, Joseph (Kip) Haney, Toxicology Division, Texas Commission on Environmental Quality, Austin, TX**

1. **Provide a few sentences summarizing the method illustrated by the case study.**

Risk managers as well as the general public may want information on the air concentration where health effects would be expected to occur in some individuals of the population (i.e., an air concentration adverse effect level). This case study reviews the procedures used to set acute and chronic adverse effect levels for chemicals with a threshold mode of action (MOA) and a nonthreshold MOA based on guidance in Appendix A of this case study summary (Section 3.13 of TCEQ Guidelines to Develop Toxicity Factors, TCEQ 2012).

Generally for this case study, an effects level based on human dose-response data is an estimate of the lowest points of departure human equivalent concentration ($POD_{HEC}$) that may be expected to cause an adverse response in some humans exposed over a similar or longer duration. Potential human effects levels based on animal dose-response data are estimated by the range of $POD_{HEC}$ values for an endpoint within which it is possible some humans exposed over a similar or longer duration may have an adverse response (assuming no sufficient data on interspecies variability are available). Additionally, the probability of response associated with the POD(s) used to estimate an effects level (e.g., $BMC_{10}$ or % response at the $LOAEL_{HEC}$) may be informative as to the probability of response in similarly-exposed individuals, depending upon the relative sensitivities of the study population compared to the environmentally-exposed population.

As effect levels should be based on concentrations demonstrated to be causally associated with a probability of effects occurring (i.e., actual effect levels), PODs associated with effects should generally not be divided by uncertainty factors (UFs) or duration adjusted since these procedures often have an unknown effect on the probability of a response actually occurring (e.g., unless predictive chemical-specific "n" values, PBPK models or CSAFs are available). However, if animal studies are used, adjustments designed to be predictive in nature (as opposed to simply conservative) such as animal-to-human dosimetry are performed to derive a $POD_{HEC}$ when possible as these procedures themselves should not appreciably affect the expectation of a response (although, for example, interspecies differences in sensitivity may exist). If duration adjustments believed to be toxicologically predictive for the chemical and endpoint in question cannot be performed to the exposure duration of interest (e.g., an adjustment from a 6-hour exposure to a 1-hour exposure cannot be performed because no chemical-specific "n" value for the endpoint is available), the estimated effect level is tied to the exposure scenario under which adverse effects were observed (e.g., 6-hour exposure).

Effects levels have appropriate and often unavoidable caveats because an actual "bright line" cannot be accurately predicted (i.e., an accurate lower bound on the lowest level at which individuals in sensitive subpopulations will respond is often unknown, may not be accurately predicted, and other uncertainties exist such as interspecies differences). The caveats associated with an effect level will vary based on the data available (e.g., if amenable to BMC modeling, sensitivity of the study population, percentage of humans responding at the LOAEL).

These case studies review the procedures used to set acute and chronic effect levels for four example chemicals illustrating development of different types of air concentration adverse effect levels based on threshold and nonthreshold MOAs:

- Methylene Chloride (Case Study A) illustrates development of a carcinogenic effect level based on an animal study (nonthreshold MOA).
- 1,3-Butadiene (BD) (Case Study B-1) illustrates development of an acute effect level using benchmark dose modeling to derive a point of departure (POD) and the use of a sensitive animal species and a species thought to be similar to humans (threshold MOA) (Case Study B-1). Case Study B-2 illustrates a chronic carcinogenic effect level for BD based on an occupational epidemiological study (nonthreshold MOA)
- Benzene (Case Study C) illustrates development of a subacute effect level using a sensitive animal species (threshold MOA) and also a chronic carcinogenic effect level based on an occupational epidemiological study (nonthreshold MOA)
- Acrolein (Case Study D) illustrates development of an acute effect level using a human clinical study (threshold MOA) and also a subchronic/chronic noncarcinogenic effect level based on an animal study

The purpose of this case study is to obtain comments from the panel on procedures to develop effect levels, not on procedures to calculate the health-protective reference values.

2. **Describe the problem formulation(s) the case study is designed to address. How is the method described in the case useful for addressing the problem formulation?**

The following case study concerns the development of air concentration adverse effect levels for comparison to lower health-protective air concentrations (e.g., USEPA Reference Concentrations (RfC) or TCEQ Reference Values (ReVs)), to help inform risk managers, assessors, and the general population. For example, in regard to TCEQ health-protective air comparison values, short-term ReVs are developed to evaluate acute exposures of one hour (hr) whereas long-term ReVs are developed to evaluate annual or longer average air monitoring data (i.e., chronic exposures). If air monitoring data are at or below health-based ReVs, then adverse health effects are not expected to occur. If air monitoring data are above ReVs, it doesn't necessarily mean that adverse effects would occur, but further analyses are required.

However, the general public typically believes that if an exceedance of a health-protective comparison value occurs, then adverse health effects will occur (i.e., believes health-protective levels are "bright lines" between safe levels and those at which adverse effects will occur). An evaluation of health-protective comparison values (e.g., ReVs) and the levels where adverse effects would actually be expected in some individuals based on dose-response data in the observable range of the data provides useful information and important context for risk managers and the general population. This information is an important part of the risk communication process. In addition, this information is helpful to risk assessors for performing health effects reviews when air monitoring data exceed health-protective levels.

The methods described in the case study are simple and straight forward and are useful for addressing the problem formulation because they present guidelines to calculate effect levels based on the dose-response relationship and the MOA. The procedures in Appendix A are part of updated guidelines (TCEQ 2012) that have been peer-reviewed (TERA 2011).

3. **Comment on whether the method is general enough to be used directly, or if it can be extrapolated, for application to other chemicals and/or problem formulations. Please explain why or why not.**

Although the case examples are for specific chemicals and are written specifically for inhalation exposure data, analogous procedures could be used for oral data and other chemicals. This method can be used by others who

need to communicate health risks with managers and the general public when exceedances of health-protective comparison values occur.

To the extent possible, determinations of actual effect levels should have a reasonable degree of certainty associated with them, and therefore should be based on what is known (i.e., be founded in actual dose-response data). This method is not useful for chemicals with limited toxicity information. An adverse effect level cannot be determined using toxicity studies where only a free-standing NOAEL is identified, although providing information on the free-standing $NOAEL_{HEC}$ without duration or uncertainty factor adjustments may be useful.

4. **Discuss the overall strengths and limitations of the methodology.**

There are several overall strengths to this methodology. The procedures in Appendix A are part of updated guidelines (TCEQ 2012) that have been peer-reviewed (TERA 2011).

The purpose is to report an air concentration effect level for the critical adverse effect (i.e., the effect that occurs at the lowest concentration in the most sensitive species relevant to humans) which is associated with some probability/expectation (or at least possibility) of a response in some humans (exposed over a similar or longer duration) based on available dose-response data. The methods and approaches used to develop inhalation effect levels are simple and straight-forward. If a health-protective comparison value (e.g., RfC, REV) has been determined, then it is relatively simple to calculate the effect level. The methods used to derive effect levels generally do not incorporate UFs or duration adjustments as these procedures often result in an unknown effect on the probability of a response. If the adverse effect is observed in animals, then the LOAEL (for example) can be converted to a $LOAEL_{HEC}$ concentration. If a health-protective comparison value (e.g., RfC, ReV) has been determined, then it is relatively simple to calculate the effect level (i.e., choose the appropriate POD identified in the critical study used to derive the RfC or ReV, and perform animal to human dosimetric adjustments and any predictive duration adjustments). When effect levels are developed, the associated uncertainties should be discussed in a narrative (e.g., potential interspecies and intraspecies differences in sensitivity).

Effect levels are based on dose-response data in the observable range using the critical effect (the adverse health effect that occurs at the lowest concentration) – values are not extrapolated below known effect levels.

As with most methodologies there are also limitations. These effect levels are primarily for informational purposes and may have significant caveats depending upon the available information. For example, the general public may want information on effect levels for sensitive subpopulations, although an accurate lower bound on the lowest level at which individuals in sensitive subpopulations will respond is often unknown and may not be accurately predicted. There are other potentially significant uncertainties.

When human data are available for determination of effect levels, the effect levels are more predictive of effect levels where adverse effects will occur in some individuals. When animal data are used as the basis of effect levels, there is uncertainty that the effect levels are relevant and predictive for humans (e.g., interspecies differences in sensitivity). As previously mentioned, when effect levels are developed, a narrative that discusses the uncertainties associated with the effect levels should be included. For example, in chemical-specific Case Studies A-D, a brief summary of the effect level values and a narrative explaining the uncertainties in the effect levels is included. The narrative discusses such issues

as whether the effect level is based on the most sensitive species or multiple species exhibiting a range of sensitivity, MOA information pertinent to human relevance (Boobis et al. 2006, 2008), the associated response level and exposure scenario, potential interspecies and intraspecies differences in sensitivity, etc.

These effect levels are not designed to be used for evacuation or other emergency response activities - but as information to provide perspective on the RfC or ReV air concentration.

5. **Outline the minimum data requirements and describe the types of data needed.**

Effect levels should be provided for those chemicals with adequate toxicity information, not for chemicals with limited toxicity data. Appropriate PODs for the critical effects should be available (i.e., the NOAEL, LOAEL or other appropriate points of departure ($BMCL_{10}$ and BMCL)). If an animal study is used, then data should be available to perform conversion of the $POD_{animal}$ to the $POD_{HEC}$, and to evaluate whether the effect in animals is relevant to humans. A free-standing NOAEL should not be used to predict effect levels. If a health-protective comparison value (e.g., RfC, REV) has been determined, then it is relatively simple to calculate the effect level.

**HOW THIS ASSESSMENT ADDRESSES ISSUES RAISED IN SCIENCE & DECISIONS:**

A. **Describe the dose-response relationship in the dose range relevant to human exposure**?

Yes, to the extent possible. Procedures for calculation of effect levels for threshold chemicals as well as nonthreshold chemicals are provided and are based on the observed dose-response relationship in the low dose range most relevant to human exposure (e.g., lowest human exposure associated with increased cancer risk). When human data are available for determination of effect levels, the effect levels are levels where adverse effects were found to occur in some individuals. When animal data are used as the basis of effect levels, there is uncertainty that the effect levels are relevant and predictive of when effects will occur in humans. Guidance discussed as part of an IPCS framework (e.g., MOA information, species sensitivity) should be considered to determine the extent to which effect levels from animal studies are relevant and predictive for humans (Boobis et al. 2006, 2008). If MOA information is not available, then it is assumed as a default that responses in animals are relevant to humans.

B. **Address human variability and sensitive populations?**

Yes, to the extent possible. If human data are available in known or potentially sensitive subpopulations, those data should be used for determining effect levels. Otherwise, the effect levels are applicable to individuals in the general population, not sensitive subpopulations. Determinations of actual effect levels should have a reasonable degree of certainty based on what is known (i.e., be founded in actual dose-response data). UFs (e.g., intrahuman UF) are not applied because they are based in uncertainty and applying a UF interjects uncertainty about (i.e., essentially negates) the expectation of a human response occurring in some individuals based on the dose-response data (i.e., has an unknown effect on the probability of a response).

C. **Address background exposures and responses?**

These methods do not directly address background exposures or responses in people, but indirectly reflect background exposures and responses to the extent that they contributed to the effects

observed in the key studies. Air concentration adverse effect levels are well above background exposures so extrapolation below the observed data is not performed.

D. **Address incorporation of existing biological understanding of the likely mode of action (MOA)?**

Yes, MOA data can be used to more fully understand the relevance and/or predictiveness of the effect level. When animal data are used as the basis of effect levels, MOA information should be considered to determine the extent to which effect levels from animal studies are relevant to humans (Boobis et al. 2006, 2008). MOA information is useful to understand the relevance and/or predictiveness of the effect level when animal data from different species are available (e.g., Case Study B-1 as an example). Different procedures for developing effect levels based on threshold and nonthreshold MOAs are provided.

E. **Address other extrapolations, if relevant – insufficient data, including duration extrapolations, interspecies?**

Yes, the applicability of such extrapolations is considered and discussed. More specifically, if data are not available for an exposure duration adjustment that is predictive of toxicity (as opposed to an adjustment that is merely conservative), exposure duration adjustments will not be performed. For example, for an acute 1-h ReV, an "n" of 3 is used with Haber's law as modified by ten Berge et al. (1986) to perform exposure duration adjustments from a longer exposure duration study to 1-h because it is generally considered to be conservative, not because the duration adjustment accurately predicts a 1-h level associated with the same probability of a response. For interspecies uncertainty, UFs are inapplicable because they are based in uncertainty and applying a UF has an unknown effect on the probability of a response, that is, interjects uncertainty about the expectation of a human response occurring in some individuals based on the dose-response data. Effect levels should not be developed for chemicals with insufficient toxicity data so a database UF is not applicable.

F. **Address uncertainty.**

UFs are not used to determine effect levels since effect levels are intended to predict concentrations where an adverse effect in the general population would be expected based on known effects levels. That is, the application of UFs would have an unknown effect on the probability of response observed in the study, which is the focus of determining an actual effect level (i.e., based on dose-response data what is a level associated with a probability of effect). Effect levels are based levels in the observable range of the dose response curve.

G. **Allow the calculation of risk (probability of response for the endpoint of interest) in the exposed human population?**

Yes. Risk estimates (probability of response) corresponding to the effect levels for threshold and nonthreshold chemicals are possible based on available dose-response data (e.g., at the $LOAEL_{HEC}$ a certain percentage of individuals responded, at a given air concentration workers experienced a certain increased cancer risk). However, the extent to which the probability of response (risk) observed in an animal or human study is predictive of the probability of response in the general population depends upon the relative sensitivities of these populations, which may be largely unknown and a caveat of the associated effect level as discussed in the narrative.

**Threshold Chemicals**

For threshold chemicals, for example, the percentage of workers or volunteers affected at the $LOAEL_{HEC}$ is the risk level. If the data are amenable to BMC modeling and the $BMC_{HEC}$ does not require significant extrapolation below the range of the data (i.e., is well founded in the dose-response data) and is used as the effect level, the benchmark response level that is considered adverse may be considered as the risk level (typically 5-10% response level).

**Nonthreshold Chemicals**

For animal studies, air concentrations corresponding to the detected increase in cancer incidence/mortality over background can be used (i.e., the $EC_{10}$ converted to an HEC).

For epidemiology studies, an air concentration corresponding to the excess risk level detected by the key epidemiological study (e.g., $10^{-3}$), preferably based on the statistical best estimate of the potency factor since this may be most predictive (i.e., central estimate or maximum likelihood estimate), can be considered the lowest level for which effects in some individuals in the human population would be expected with reasonable certainty if exposed over a similar (or longer) exposure duration than those in the epidemiological study. Alternatively, a range of values corresponding to air risk-specific concentrations of around $10^{-3}$ (e.g., $10^{-4}$ to $10^{-3}$) could be used since it may be possible to detect an increase in background cancer incidence/mortality at the $10^{-3}$ risk level for a well-conducted epidemiology study with adequate number of subjects and statistical power.

**H. Work practically? If the method still requires development, how close is it to practical implementation?**

The procedures for calculation of inhalation effect levels are included in updated TCEQ Guidelines to Develop Toxicity Factors (TCEQ 2012) and have undergone a peer review (TERA 2011). They are practical and readily implemented by trained risk assessors if an RfC or ReV has been developed. However, no effect levels have been included in Development Support Documents developed by the TCEQ as of this time. This case study is designed to provide effect levels for 1,3-butadiene, benzene, acrolein, and methylene chloride as example chemicals to demonstrate the practical implementation of the method.

## References

Boobis, A. R., Cohen, S. M., Dellarco, V., et al. (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. Crit Rev Toxicol, 36(10), 781-792.

Boobis, A. R., Doe, J. E., Heinrich-Hirsch, B., et al. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. Crit Rev Toxicol, 38(2), 87-96.

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

Texas Commission on Environmental Quality. TCEQ 2007. Development support document Benzene CAS registry numbers: 71-43-2.Texas Commission on Environmental Quality, Toxicology Division, Chief Engineer's Office, available at http://www.tceq.state.tx.us/implementation/tox/dsd/final.html

Texas Commission on Environmental Quality. TCEQ 2008. Development support document 1,3-Butadiene CAS registry numbers: 106-99-0.Texas Commission on Environmental Quality, Toxicology Division, Chief Engineer's Office, available at http://www.tceq.state.tx.us/implementation/tox/dsd/final.html

Texas Commission on Environmental Quality. TCEQ 2010. Development support document Acrolein CAS registry numbers: 107-02-8.Texas Commission on Environmental Quality, Toxicology Division, Chief Engineer's Office, available at http://www.tceq.state.tx.us/implementation/tox/dsd/final.html

Texas Commission on Environmental Quality. TCEQ 2011. Development support document Methylene Chloride CAS registry numbers: 75-09-2.Texas Commission on Environmental Quality, Toxicology Division, Chief Engineer's Office, available at http://www.tceq.state.tx.us/implementation/tox/dsd/final.html

Texas Commission on Environmental Quality. TCEQ 2012. TCEQ Guidelines to develop toxicity factors, RG-442 Revised DRAFT. Texas Commission on Environmental Quality, available at http://m.tceq.texas.gov/toxicology/esl/guidelines/about.html

Toxicology Excellence for Risk Assessment. 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. Final Report August 31, 2011. Available at http://www.tera.org/peer/tceqesl/TCEQ%20ESL%20Report%20Final%208%2031%2011.pdf

# Appendix A
# Identification of Inhalation Effect Levels (TCEQ 2012)

## 3.13 Identification of Inhalation Effect Levels

Toxicity factors and their corresponding risk-specific concentrations are considered "safe levels" because they are set below levels where adverse health effects are expected to occur. Risk managers as well as the general public may want information on the air concentrations where health effects would be expected to occur (i.e., an air concentration adverse effect level) in some individuals of the population. Thus, when adequate data exist for inhalation, the TCEQ will provide effect levels in DSDs for comparison to safe levels. Although written specifically for inhalation exposure data, analogous procedures could be used for oral data. These effect levels are primarily for informational purposes.

## 3.13.1 Chemicals with a Threshold MOA

For noncarcinogens or carcinogens with a threshold MOA, the $LOAEL_{HEC}$ from the study that identified the critical effect can be considered the lowest documented level where effects in the human population could be expected to occur in some members of the population. If BMC modeling is conducted, the central estimate $BMC_{HEC}$ corresponding to a BMR of concern for adverse effects (e.g., $BMC_{10\text{-}HEC}$ for decreased body weight) which does not require significant extrapolation below the range of the data is used as the lowest level where effects in the human population could be expected to occur.
More specifically, an $LOAEL_{HEC}$ determined from human studies, where effects occurred in some individuals, represents a concentration at which it is probable that similar effects could occur in some individuals exposed to this level over the same duration as used in the study or longer. Importantly, effects are not a certainty due to potential intraspecies differences in sensitivity. Conversely, the $NOAEL_{HEC}$ from a human study is the highest concentration known (based on dose-response data) which may not be expected to result in adverse effects in humans similar to the study-exposed population (e.g., workers, adult volunteers) exposed over the same (or shorter) duration, although this is not a certainty (e.g., study power considerations). In addition, other subpopulations could be more sensitive than the study-exposed population.

For an estimated $LOAEL_{HEC}$ extrapolated from animal studies in the most sensitive species, as effects occurred in some animals in the most sensitive species, the $LOAEL_{HEC}$ represents a concentration at which it is possible that similar effects could occur in some individuals exposed to this level over the same duration as used in the study or longer, assuming no available data on the sensitivity of animals versus humans, although effects are not a certainty (e.g., potential species differences in sensitivity). If laboratory animal data are relied upon and there is no information on the sensitivity of animals versus humans, the determination of effect levels needs to be put into the context of a discussion of studies in other species which did not show effects at similar or higher levels/durations.

## 3.13.2 Application of UFs or Duration Adjustments

To the extent possible, determinations of actual effect levels should have a reasonable degree of certainty associated with them, and therefore should be based on what is known (i.e., be founded in actual dose-response data). Consequently, UFs are inapplicable because they are based in uncertainty and applying a UF interjects uncertainty about (i.e., essentially negates) the expectation of a human response occurring in some individuals based on the dose-response data. Additionally, if data are not available for an exposure duration adjustment that is predictive of toxicity (as opposed to an adjustment that is merely conservative), exposure duration adjustments will not be performed. For example, for an acute 1-h ReV, an "n" of 3 is used to perform exposure duration adjustments from a longer exposure duration study to 1-h because it is generally considered to be conservative, not because the duration adjustment accurately predicts a 1-h level associated with the same probability of a response. Using UFs and such duration adjustments results in a value with an unknown ability to predict the probability of a response and arguably some conservatism. This is contrary to the purpose of, to the extent possible based on available dose-response data, identifying a level where with a reasonable degree of certainty, a response in some individuals may be expected.

## 3.13.3 Chemicals with a Nonthreshold MOA

For carcinogenic effects (or noncarcinogens with a nonthreshold MOA), the risk-specific dose for the $^{chronic}ESL_{linear(c)}$ is set at the no significant excess risk level associated with a theoretical excess lifetime cancer risk of one in 100,000 ($1 \times 10^{-5}$ or simply $10^{-5}$). USEPA's acceptable risk range is $10^{-6}$ to $10^{-4}$ (USEPA 2000e). When tumor data are used, a POD is obtained from the modeled tumor incidences. Conventional cancer bioassays, with approximately 50 animals per group, generally can support modeling down to an increased incidence of 1–10% ($10^{-2}$ to $10^{-1}$ risk); epidemiologic studies, with larger sample sizes, below 1% ($10^{-2}$ risk) (USEPA 2005a). For a well-conducted epidemiology study with adequate number of subjects and statistical power, it may be possible to detect an increase in background cancer incidence/mortality at the $10^{-3}$ risk level (Grant *et al.* 2007) or lower. Seiler and Alvarez (1994) determined that for radiation carcinogenesis, the minimum significant risk for the model is considerable larger than $10^{-3}$ and for the usual confidence limits, the minimum significant risk exceeds $10^{-2}$:

> "Whereas a more careful error analysis may yield lower limits, it is unlikely that they will lie below $1 \times 10^{-3}$. Thus, even though risk values below this limit can be calculated, they are not meaningful because they are smaller than their total standard errors, and are thus not compatible with finite risks."

An air concentration corresponding to the excess risk level detected by the key epidemiological study (e.g., $10^{-3}$), preferably based on the statistical best estimate of the potency factor since this may be most predictive (i.e., central estimate or maximum likelihood estimate), can be considered the lowest level for which effects in some individuals in the human population would be expected with reasonable certainty if exposed over a similar (or longer) exposure duration than those in the epidemiological study. Alternatively, a range of values corresponding to air risk-specific concentrations of around $10^{-3}$ (e.g., $10^{-4}$ to $10^{-3}$) could be used.

However, calculating these risk-specific values based on assumed lifetime environmental exposure (e.g., using duration adjustments) may be inappropriate for this purpose. For example, dose rate may have at least enhanced carcinogenesis or the carcinogenic MOA in workers (e.g., the metabolic pathways responsible for carcinogenesis in workers). That is, an exposure duration similar to that in the

epidemiological study may need to be used to calculate these risk-specific values so that the predictiveness of the values is reasonably certain to the extent possible (e.g., dose rate effects and dose-related changes in metabolic pathways are not potential issues).

For animal studies, air concentrations corresponding to the detected increase in cancer incidence/mortality over background or the $EC_{10}$ can be used. The $EC_{10}$ should be converted to an HEC. The considerations discussed above would still apply (e.g., interspecies sensitivity, duration adjustments). For example, if there is no information on the sensitivity of animals versus humans, the determination of cancer effect levels should be put into the context of studies in other species which did not show effects at similar or higher levels/durations. Ultimately, the response level used to calculate an air concentration where adverse effects would be expected to occur (based on dose-response data) in some individuals of the population will be based on best scientific judgment on a case-by-case basis and justified in the DSD.