Chapter 7

Hazard Characterization and Exposure-Response Assessment Using Epidemiology Data

Draft June 7, 2011

DO NOT CITE OR QUOTE

Prepared by
Sielken & Associates Consulting, Inc.
3833 South Texas Avenue, Suite 230, Bryan, TX 77802
Tel: 979-846-5175; Fax: 979-846-2671; Email: SielkenAssoc@aol.com

Prepared For
Toxicology Division
Chief Engineer's Office
Work Order No. 582-9-80187-05

Table of Contents

List of Figures

Acronyms and Abbreviations

Acronyms and Abbreviations	Definition
ACGIH	American Conference of Governmental Industrial Hygienists
ADAF	age-dependent adjustment factor
ADI	acceptable daily intake
AEGL	Acute Exposure Guideline Level
AIHA	American Industrial Hygiene Association
AMCV	Air Monitoring Comparison Value
AMCV Odor	$^{acute}ESL_{odor}$
AMCV short-term vegetation	$^{acute}ESL_{veg}$
AMCV long-term vegetation	$^{chronic}ESL_{veg}$
AMCV short-term health	Acute ReV or $^{acute}ESL_{generic}$ or interim ESL
AMCV long-term health	lowest value of the chronic ReV [nonlinear(c)], chronic ReV [nonlinear(nc)], $^{chronic}ESL_{linear(c)}$, or $^{chronic}ESL_{linear(nc), or interim ESL}$
ATSDR	Agency for Toxic Substances and Disease Registry
BMC	benchmark concentration
BMCL	benchmark concentration lower confidence limit
BMD	benchmark dose
BMDL	benchmark dose lower confidence limit
BMDS	benchmark dose software
BMR	benchmark response
C	concentration
Cal EPA	California Environmental Protection Agency
CFR	Code of Federal Regulations
CNS	Central nervous system
COT	Committee on Toxicolgy
CSAF	chemical-specific adjustment factor

Acronyms and Abbreviations	Definition
D	exposure duration, hour per day
da	dalton
DF	deposition fraction in the target region of the respiratory tract
DAF	dosimetric adjustment factor
DOE	Department of Energy
DSD	development support document
E	exposure level or concentration
EC	effective concentration
ET	extrathoracic
ERPG	Emergency Response Planning Guideline
ESL	Effects Screening Level
$^{acute}ESL$	acute health-based Effects Screening Level for chemicals meeting minimum database requirements
$^{acute}ESL_{generic}$	acute health-based Effects Screening Level for chemicals not meeting minimum database requirements
$^{acute}ESL_{odor}$	acute odor-based Effects Screening Level
$^{acute}ESL_{veg}$	acute vegetation-based Effects Screening Level
$^{chronic}ESL_{linear(c)}$	chronic health-based Effects Screening Level for linear dose response cancer effect
$^{chronic}ESL_{linear(nc)}$	chronic health-based Effects Screening Level for linear dose response noncancer effects
$^{chronic}ESL_{nonlinear(c)}$	chronic health-based Effects Screening Level for nonlinear dose response cancer effects
$^{chronic}ESL_{nonlinear(nc)}$	chronic health-based Effects Screening Level for nonlinear dose response noncancer effects
$^{chronic}ESL_{veg}$	chronic vegetation-based Effects Screening Level
F	exposure frequency, days per week
FDA	Food and Drug administration
FEL	frank effect level
GHS	Globally Harmonized System
GLC	ground-level concentration
GLC_{max}	maximum ground-level concentration
h	hour

Acronyms and Abbreviations	Definition
$H_{b/g}$	blood:gas partition coefficient
HEAST	Health Effects Assessment Summary Tables
HEC	human equivalent concentration
HED	human equivalent dose
HQ	hazard quotient
IARC	International Agency for Research on Cancer
ICD	International Classification of Diseases
ICRP	International Commission on Radiological Protection
IDLH	Immediately Dangerous to Life or Health
IPCS	International Programme on Chemical Safety
IRIS	Integrated Risk Information System
K	constant level or severity of response
LC_{50}	concentration producing lethality in 50% of experimental animals
LC_{Lo}	lowest concentration producing lethality
LD_{50}	dose producing lethality in 50% of experimental animals
LEC	effective concentration lower confidence limit
LLNA	local lymph node assay
LOAEL	lowest-observed-adverse-effect-level
LTD	limited toxicity data
m	meter
MAK	Federal Republic of Germany Maximum Concentration Values in the Workplace
MERA	Modeling and Effects Review Applicability
MF	modifying factor
MLE	maximum likelihood estimate
MW	molecular weight
µg	microgram
min	minute
MPPD	multiple pass particle dosimetry
MOA	mode of action
MRL	Minimal Risk Level
NAAQS	National Ambient Air Quality Standards

Acronyms and Abbreviations	Definition
NAC	National Advisory Committee
NATA	National-Scale Air Toxics Assessment
NCEA	National Center for Environmental Assessment
NF	normalizing factor
NIOSH	National Institute for Occupational Safety and Health
N-L Ratio	NOAEL-to LC_{50} Ratio
NLM	National Library of Medicine
NOAEL	no-observed-adverse-effect-level
NOEL	no-observed-effect-level
NRC	National Research Council
NTIS	National Technical Information Service
NTP	National Toxicology Program
OECD	Organisation for Economic Co-operation and Development
OEHHA	Office of Environmental Health Hazard Assessment
OEL	Occupational Exposure Limit
OPPTS	Office of Prevention, Pesticides and Toxic Substances
OSHA	Occupational Safety and Health Administration
PAH	polycyclic aromatic hydrocarbon
PBPK	physiologically-based pharmacokinetic model
PCBs	polychlorinated biphenyls
PEL	Permissible Exposure Limit
PM	particulate matter
POD	point of departure
POD_{ADJ}	point of departure adjusted for exposure duration
POD_{HEC}	point of departure adjusted for human equivalent concentration
POE	portal of entry
PU	pulmonary
ppbv	parts per billion by volume
ppb	parts per billion
ppm	parts per million
QSAR	quantitative structure-activity relationship
RDDR	regional deposited dose ratio

Acronyms and Abbreviations	Definition
REL	Reference Exposure Level (Cal EPA OEHHA)
REL	Recommended Exposure Limit (NIOSH)
ReV	Reference Value
RfC	Reference Concentration
RfD	Reference dose
RGD_A	regional gas dose in animal
RGD_H	regional gas dose in human
RGDR	regional gas dose ratio
RIVM	Rijksinstituut voor Volksgezondheid en Milieu (Dutch National Institute for Public Health and the Environment)
R_{GM}	geometric mean ratio
RPF	relative potency factor
RTECS	Registry of Toxic Effects of Chemical Substances
SAR	structure activity relationship
SCAPA	Subcommittee on Consequence Assessment and Protective Action
STEL	Short-term Exposure Level
T	time or exposure duration
TB	trachiobronchial
TC	toxicity category
TCEQ	Texas Commission on Environmental Quality
2,3,7,8-TCDD	2,3,7,8-tetrachlorodibenzo-p-dioxin
TEEL	Temporary Emergency Exposure Limit
TEF	toxicity equivalency factor
TH	thoracic
THSC	Texas Health and Safety Code
TLV	Threshold Limit Value
TS	Toxicology Section
TOC	threshold of concern
TOXLINE	Toxicology Literature Online
TWA	Time-Weighted Average
TWA-TLV	Time-Weighted Average Threshold Limit Value
UF	uncertainty factor

Acronyms and Abbreviations	Definition
UF_H	interindividual or intraspecies human uncertainty factor
UF_A	animal to human uncertainty factor
UF_{Sub}	subchronic to chronic exposure uncertainty factor
UF_L	LOAEL to NOAEL uncertainty factor
UF_D	incomplete database uncertainty factor
UN	United Nations
URF	Unit Risk Factor
URF (MLE)	Unit Risk Factor, maximum likelihood estimate
URF (95% UCL)	Unit Risk Factor, 95% upper bound
URF (95% LCL)	Unit Risk Factor, 95% lower bound
USEPA	United States Environmental Protection Agency
VE	minute ventilation
VE_{ho}	default occupational ventilation rate for an eight-hour day
VE_h	default non-occupational ventilation rate for a 24-h day
WEEL	Workplace Environmental Exposure Level
WHO	World Health Organization
WOE	weight of evidence

Chapter 7 Hazard Characterization and Exposure-Response Assessment Using Epidemiology Data

7.1 Objectives

This chapter provides guidance for conducting a quantitative hazard characterization and exposure-response assessment using epidemiology data. It is neither an exhaustive text on epidemiology nor a guide to the conduct of an epidemiology study. Rather, it is a guide to the methods for performing quantitative hazard characterizations and exposure-response assessments using existing epidemiology studies that focus on chronic toxicity rather than acute toxicity. The main body of this epidemiology cancer section concerns general guidelines, with specific mathematical models included in Appendix A - Linear Multiplicative Relative Risk Models:

- A.1 Overview of Poisson Regression Models
- A.2 Summary Estimates of Standardized Mortality/Incidence Rates
- A.3 Adjustments for Possible Differences Between the Population Background Cancer Rate and the Cohort's Cancer Rate in the Relative Risk Model
- A.4 Estimating the Slope Parameter, β, in the Relative Risk Model Adjusting for Differences in Background Rates
- A.5 Estimating the Asymptotic Variance for the Slope Parameter in the Relative Risk Model

Exposure-response and dose-response are used synonymously throughout this chapter. Exposure-response is a term preferred by epidemiologists whereas dose-response is commonly used by toxicologists. Special emphasis is placed on deriving carcinogenic inhalation unit risk factors and oral slope factors based on human epidemiology studies. Consequently, the discussion is in terms of a dichotomous response such as the presence or absence of a specified carcinogenic response or development of a specific type of cancer.

In this chapter, epidemiology is considered to be the study of diseases in specified populations of humans. The science of epidemiology was first developed to discover and understand possible causes of contagious diseases such as smallpox, typhoid and polio among humans. It has expanded to include the study of factors associated with non-transmissible diseases like cancer, and of potential adverse health effects caused by environmental exposures. Some epidemiology studies are mostly qualitative, primarily descriptive, and focus on determining what factors are associated with diseases (risk factors) and the associated distribution of the disease among the members of the population. More quantitative epidemiology studies attempt to quantify the exposure and the relationships between exposure characteristics (duration, intensity, timing, co-exposures, etc.) and response characteristics (frequency, probability, standardized mortality rates (SMRs), relative risk rates (RRs), odds ratios (ORs), timing, severity, etc.). Some of these more quantitative epidemiology studies can be used for exposure-response modeling and may provide

useful information for extrapolating from relatively high exposure scenarios to lower environmental exposure scenarios.

Properly conducted epidemiology studies (i.e., a proper study design, confounding factors accounted for, Bradford Hill Criteria considered, etc.) can be useful tools. Epidemiology studies can provide evidence (sometimes strong evidence) concerning risk factors; however, they cannot "prove" that a specific risk factor actually causes the disease being studied. In contrast to a cohort or case-control study, an ecological study is an epidemiology study wherein the unit of analysis is a group rather than an individual and, as such, is not suitable for TCEQ's dose-response modeling. The "ecological fallacy" occurs because statistics that accurately describe group characteristics are not necessarily applicable to individuals within that group (e.g., Pearce 2000). Please refer to Appendix C Glossary for definitions of terms used in these guidelines.

7.2 Published Hazard Characterizations and Exposure-Response Assessments

It is generally recognized that human epidemiology data are preferred over data from animal studies as the basis for dose-response modeling. Although the TCEQ does not conduct epidemiology studies, the TCEQ does review such studies and any dose-response modeling therein. If the study is of suitable quality and the necessary data are available, the TCEQ may perform its own dose-response modeling following these guidelines. When quantitative hazard or exposure-response characterizations using epidemiology data are identified in the scientific literature or databases, they are reviewed by the TCEQ to determine whether the approach used to develop these characterizations (and resultant toxicity values) is appropriate. Many published characterizations are not appropriate for use by the TCEQ because procedures other than those recommended in this guidance document were used to derive toxicity factor (e.g., URF, SFo) values. Due to time and resource constraints, the TCEQ considers the published values and their respective key studies as a starting place for gathering information on hazard or exposure-response characterizations. However, because the characterizations may be outdated, the TCEQ also evaluates peer-reviewed studies available after the date these characterizations were published to ensure that the latest data are considered prior to developing a hazard or exposure-response characterization. The TCEQ also reviews other published hazard or exposure-response characterizations from organizations that specifically address susceptibility of children. In addition, the TCEQ considers adoption of a published hazard or exposure-response characterization when the risk assessment procedures used to develop such factors are similar to those described in this guidance. Preference will be given to values that have undergone an external peer review and public involvement process.

The evaluation and selection of suitable epidemiology studies (especially for the purposes of exposure-response modeling) is discussed in the scientific literature (e.g., Federal Focus 1995, Graham 1995, Hertz-Picciotto 1995, WHO Working Group 2000, Goldbohm et al. 2006). Important topics include evaluating weight-of-evidence that a chemical causes a specific cancer(s), estimation of exposure-response regression models, estimation of uncertainty introduced by potential biases and missing information, calculation of excess lifetime risk through a life table to take into account competing risks, and sensitivity analyses focusing on the impact of assumptions made and the variability of the underlying data.

7.3 Components of the Exposure Response Assessment

Whereas Section 7.2 addresses the quality of epidemiology studies, Sections 7.4 to 7.13 focus on a selection of the most important components of the exposure-response assessment and addresses them from a top-down perspective. The purpose of the discussions in the following subsections is to guide TCEQ staff in determining the utility of specific components of an epidemiology study for the purposes of TCEQ's exposure-response assessment.

7.4 Endpoint Selection

The toxicity endpoint for hazard characterization and exposure-response assessment using epidemiology data needs to be explicitly specified *a priori* or after a carefully conducted WOE demonstrating causality based on the Bradford Hill Criteria (Hill 1965, Höfler 2005, Howick et al. 2009, Phillips and Goodman 2006, Ward 2009). If a common name for the toxicity endpoint is specified (e.g., leukemia), then the intended diseases should be more precisely defined (e.g., specific International Classification of Diseases (ICD) codes for mortality) in the epidemiology study. If ICD codes are specified, the ICD revision should also be noted in the study as well as how earlier revision codes can be transformed to the specified revision.

The EPA 2005 Guidelines for Carcinogen Risk Assessment (USEPA 2005a) discuss the role of the mechanism of action and MOA in the dose-response models. EPA clarifies the difference between mechanism of action and MOE by stating the following:

> The term 'mode of action' is defined as a sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in cancer formation. A 'key event' is an empirically observable precursor step that is itself a necessary element of the mode of action or is a biologically based marker for such an element. Mode of action is contrasted with 'mechanism of action,' which implies a more detailed understanding and description of events, often at the molecular level, than is meant by mode of action. The toxicokinetic processes that lead to formation or distribution of the active agent to the target tissue are considered in estimating dose but are not part of the mode of action as the term is used here. There are many examples of possible modes of carcinogenic action, such as mutagenicity, mitogenesis, inhibition of cell death, cytotoxicity with reparative cell proliferation, and immune suppression.

Mechanistic or biologically-based exposure-response models are most appropriate when there is a common mechanism of action or at least a common MOA for the specified toxicity endpoint. There are limitations to assessments of toxicity endpoints that involve multiple MOAs. For example, lymphohematopoietic cancers can have multiple MOAs so that, the appropriate method for low-exposure extrapolation cannot necessarily be defensibly determined. Thus, the TCEQ takes special care when the toxicity endpoint involves grouping of responses with possibly multiple target tissues, mechanisms of action, MOAs, and severities.

There are also other important issues that the TCEQ considers when selecting the toxicity endpoint:

- Potential double counting of individuals with the toxicity endpoint should be addressed and avoided if possible.

- While data derived from death certificates can provide invaluable information, the usefulness of such data depends on the completeness of records, the accuracy in assigning underlying causes of death, etc. (e.g., Bonita et al. 2000, Manos et al. 2008). Thus, there are limitations of death certificates or other means of identifying the presence or absence of a specified toxicity endpoint in the epidemiology study.

- Epidemiology studies usually are based on mortality or incidence data. The severity of the response (e.g., mortality, incidence) should be clearly defined and should be the same severity for both the exposure-response modeling and the risk characterization.

7.5 Exposure Characterization

Exposure characterization is performed by the epidemiologist. The manner in which an individual's exposure is characterized can be important and should be discussed in the epidemiology study. The exposure metric used in the exposure-response model is also important and is discussed in the next section. The TCEQ considers the following potential exposure characterization issues when reviewing epidemiology studies for potential use in deriving toxicity factors:

- temporality,
- measurements,
- models,
- reasonableness of underlying modeling assumptions,
- exposure estimation errors,
- grouped versus continuous exposure values, and
- biomonitoring.

Exposure in epidemiology studies generally occurs over a period of time and has a temporal profile. Exposure characterizations that capture the changes in this profile over calendar years and changes in the epidemiology setting (e.g., occupational setting) generally support age-dependent, exposure-response modeling and calculation of excess risk (i.e., BEIR IV modeling discussed in Section 7.6) and are generally preferred. For example, exposure histories in an occupational epidemiology study can be generated from job/task histories combined with calendar-year and job/task specific exposure characterizations (job-exposure matrices (JEMs)). Macaluso et al. (2004) provides a good example for the characterization of 1,3-butadiene in the styrene-butadiene-rubber cohort.

Ideally, the exposure concentrations in an exposure history are based on accurate analytical measurements (e.g., biomonitoring where the relationship between the exposure concentration and the biomonitored value is known) with personal measurements being better than area measurements (Hays et al. 2007). Sometimes, measurements (e.g., industrial hygiene measurements) are designed for a different purpose than individual exposure measurements and, thus, have limitations (e.g., unrepresentative sampling, changing analytical methods, incomplete documentation, and sparseness over jobs, times of the day, and/or calendar years) (e.g., Hewett

2001 and Stewart 1999). Examples using analytical methods are discussed in Moseman and Oswald 1980, Paustenbach 2006, and Hays et al. 2008.

By necessity, exposure concentration estimates from exposure models are frequently used as alternatives to analytical measurements. Numerous individuals and groups (including academics and federal and state agencies) have created exposure models (e.g., EPA Center for Exposure Assessment Modeling (CEAM), EPA Office of Pollution Prevention and Toxics (OPPT), Calendex™ (http://www.exponent.com/practices/foodchemical/calendex.html), Cumulative and Aggregate Risk Evaluation System (CARES™) (http://cares.ilsi.org), LifeLine™ (http://www.thelifelinegroup.org/lifeline/index.htm), McKone 1987, Little and Chiu 1988, Paustenbach 1989, Kim et al. 2004, Ott et al. 2007). An example where estimates from exposure models were used is the characterization of 1,3-butadiene in the styrene-butadiene-rubber cohort (Macaluso et al. 2004 and Sielken and Valdez-Flores 2011). Because not all exposure models are created equal but some are useful when implemented by epidemiologists in the context of their study, the TCEQ considers the reasonableness of the exposure models when evaluating the usefulness of the study for the TCEQ's purposes.

The evaluation of an exposure model's utility should be based on the quality of the information incorporated into the model and the reasonableness of underlying modeling assumptions. Also, some models may seem to include multiple aspects of exposure (e.g., exposure duration, intensity, formulation, method of use, and personal protective equipment) but are really primarily functions of only one input (duration). For example, duration (e.g., years or days) frequently dominate time-dependent exposure models .

All exposure estimates involve errors and a discussion of these errors should be made available in the published epidemiology study and included in the uncertainty section of the DSD. Two examples of such errors are measurement errors of a continuous variable and misclassification errors of a categorical variable (e.g., classifying a job's exposure into a lower category than it should be). It may be important to use sensitivity analyses (where possible) to determine whether these errors may be leading to a directional bias (e.g., over- or under-estimation of exposure). Measurement errors of a continuous variable (the difference between the actual value of a quantity and the value obtained by a measurement) may be random errors that are unbiased (e.g., for additive errors their expected value is zero). However, even these unbiased random errors may lead to a biased estimator depending on what characteristic of the exposure is being estimated. For example, if the estimator is estimating the 95th percentile of a distribution, then the 95th percentile of the sample values with errors is expected to be an over-estimate of the 95th percentile of the sample values without errors. Thus, knowing the directional bias of an exposure estimator may guide the interpretation and use of that estimator by an epidemiologist.

A random variable is a function that associates a unique numerical value with every outcome of an experiment. A continuous variable is a random variable that can take on any value between its minimum value and its maximum value. Continuous variables typically correspond to measurements. For example, body weight is a continuous variable if the weight can be measured to as many decimal points as desired and is not restricted to be a whole number of units. Random variables that are restricted to a countable number of possible values are discrete variables. For example, if body weight is restricted to a whole number of pounds (or kilograms), then body weight is discrete. Variables that are restricted to a finite number of values are categorical

variables. Categorical variables are usually names or labels such as gender, health status, and type of job. Categorical variables may also be labels for groups of discrete or continuous variable values. For example, body weight under 100 pounds might be labeled 1; body weight between 100 and 200 pounds might be labeled 2; and body weight over 200 lbs might be labeled 3; in which case this label would be a categorical variable.

The TCEQ will evaluate how any continuous variables (e.g., cumulative ppm-years) in the epidemiology study have been grouped or partitioned into categories and how the categories are characterized quantitatively (e.g., by the mean or median in a given category) with particular attention to unbounded categories. Categorizing continuous exposure variables can cause several problems (Shaw et al. 1987, Altman et al. 1994, Schulgen et al. 1994, Figueiras and Cadarso-Suárez 2001, Hollander at al. 2004, Richardson and Loomis 2004, Royston et al. 2006, Wainer 2006, Fedorov et al. 2009). For example, categorizing can cause loss of power and loss of precision of estimated means, odds, hazards, etc. This is due to the fact that categorization assumes that the relationship between the predictor (e.g., dose) and the response (e.g., cancer endpoint) is flat within each exposure interval, which is an assumption far less reasonable than a linearity assumption in most cases. There are other potential issues as well. For example, researchers seldom agree on the choice of category cutpoints. Thus, there is a potentially severe study interpretation problem among researchers and across studies. Also, because of sample size limitations in the very high range of the exposure variable, there will be significant heterogeneity of subjects within those intervals and residual confounding. Categorization assumes that there is a discontinuity in response as interval boundaries are crossed, and categorization that is not blinded to the response variable (i.e., when response is considered in deriving categories) can result in biased effect estimates. For example, cutpoints are arbitrary and can be manipulated, which can result in either positive or negative associations within the same study depending upon the choice of cutpoints. If a confounder is adjusted for by data categorization, there may be residual confounding that can be explained away by inclusion of the continuous form of a predictor (e.g., dose) in the model in addition to the categories.

Biomonitoring can be useful as either an indicator of the presence or absence of exposure or as a quantitative measure of the magnitude of exposure. In addition, biomonitoring may also be useful as a means of creating a relative ranking of exposures and as a tool for validating other exposure characterizations. However, there are some potential limitations associated with biomonitoring such as its relation to dose and temporal integration (Hays et al. 2007). For example, the relationship between the dose at the target tissue and the amount in the biomonitoring medium (e.g., blood or urine) may be unknown. It may also be unclear as to how the observed biomonitoring value has integrated the preceding exposures over time. For example, did the urine concentration reflect exposures over the last 3 hours, 24 hours, 48 hours, etc. For such reasons, biomonitoring data may be of limited value to retrospective epidemiology studies.

7.6 Exposure Metric

The exposure metric, used in exposure-response modeling or as dose in dose-response modeling, is a critical component of both the modeling and the risk characterization. As mentioned previously, epidemiologists prefer the term exposure metric versus dose metric. The exposure metrics used for exposure-response modeling are evaluated by the TCEQ along with the

following information when reviewing epidemiology studies for potential use in deriving toxicity factors.

The most commonly reported exposure metric, which is frequently the only reported exposure metric, in epidemiology studies is cumulative exposure (e.g., cumulative ppm-years). Cumulative exposures can either incorporate or not incorporate simple lags where exposures in a specified number of preceding years are excluded. Cumulative exposures can also be restricted to an exposure window where exposures in a specified number of preceding years are excluded as well as excluding exposures that occurred more than a specified number of years into the past. Simple lags and/or exposure windows have been included in the risk assessment of several substances (Shore et al. 1992, Steenland et al. 1998, Steenland et al. 2001, Crump et al. 2003, Agalliu et al. 2005). Cumulative exposures may also be weighted. Weighted cumulative exposures are an alternative to unweighted cumulative exposures. For example, the exposure in a year can be weighted on the basis of its relative importance with respect to age or distance into the past. The TCEQ evaluated whether the specific form of the cumulative exposure is biologically and statistically defensible. In addition, when evaluating the appropriateness of specific exposure metrics, the TCEQ considerd the sensitivity of the derived toxicity factors to the exposure metrics.

It is important to note that the use of the simplest form of cumulative exposure (i.e., without lags, windows, weights, etc.) as the exposure metric makes several implicit assumptions. It assumes that cumulative exposure is more biologically relevant than other aspects of exposure such as duration and intensity. Cumulative exposure also does not differentiate between high intensity exposures for short durations and low intensity exposures for long durations. In other words, cumulative exposure assumes that the temporal pattern of exposure magnitudes within a specified exposure duration is not important to the toxic response. For example, the cumulative exposure by age 50 could be the same numerical value if all of the exposure occurred between ages 20 to 30, all of the exposure occurred between ages 40 and 50, or the exposure was at a constant level for 50 years - ignoring the temporal pattern of exposure levels. As another example, using cumulative exposure as the exposure metric assumes that a 10 ppm exposure has the same impact on the likelihood of a response today if it occurred yesterday, 5 years ago, or 50 years ago. Limitations of cumulative exposure have been widely discussed in the literature (e.g., Copes et al. 1985, ten Berge 1986, Checkoway et al. 1992, Cox et al. 1996, USEPA 1998, Weller et al. 1999, Murdoch et al. 1992, Goddard et al. 1995, Miller et al. 2000, Evans et al. 2002, Buchanan et al. 2003, Collins et al. 2003, Ginsberg 2003, Boyes et al. 2005, and Shusterman et al. 2006). After reviewing the limitations of cumulative exposure, EPA's 2005 Risk Assessment Guidelines (page 3-4) concludes that cumulative exposure or potential dose may be replaced by a more appropriate dose metric when indicated by the data.

Alternatives to cumulative exposure include those based on metrics that place greater emphasis on intensity or duration, such as $(C-C_0)^n \times (T-T_0)^m$ where C is the concentration intensity, C_0 is a concentration threshold, T is exposure duration, and T_0 is a duration threshold, and n and m are parameters (specified or estimated) (e.g., ten Berge 1986, Vacek et al. 1991, Smith 1992, Schnatter et al. 1996, Rozman 2000, Blankenship and Stefanski 2001, Bunce et al. 2003, Kriebel et al. 2007).

Additionally, exposure metrics do not necessarily have to be cumulative. For example, the exposure concentration at the time the observation is made, the exposure duration, years since hire, and the average exposure intensity are non-cumulative exposure metrics.

With several possibly relevant exposure metrics available for exposure-response assessment and because a single exposure metric (e.g., cumulative exposure) captures only one part of the exposure scenario, an epidemiologist may evaluate more than one exposure metric, either separately or together. For example, it may be worthwhile to consider other characteristics of the exposure to the toxicant under consideration as well as other potentially confounding co-exposures to other toxicants. Many of these exposure metrics may be correlated or otherwise dependent upon one another. However, their individual or joint impacts may still be worth investigating by the epidemiologist and should not be ignored *a priori*. When several exposure metrics are available to characterize exposure-response, the TCEQ will evaluate the relevance of each exposure metric being considered for selection in terms of the specified toxic response, the chemical, and its mechanism(s) and mode(s) of action.

In addition to the considerations discussed above, the TCEQ evaluates the exposure metric from the perspective of any differences between the modeling scenario and the inference scenario (i.e., the exposure scenario being extrapolated to which is the exposure to the general population). For example, the exposures in the modeling scenario might be sporadic, high-intensity exposures and the inference scenario might be continuous, low-intensity exposures. Such differences add to the uncertainty in using a particular exposure-response model from an epidemiology study to predict risks to the general population.

Ultimately, the exposure metric used in the dose-response modeling discussed in the next section must be an exposure metric reported in the epidemiology study.

7.7 Dose-Response Models

As mentioned previously, mechanistic or biologically-based dose-response models are most appropriate when there is a common mechanism of action or a common MOA for the specified toxicity endpoint. Care should be taken when the toxicity endpoint involves grouping of responses with possibly multiple target tissues, mechanisms of action, MOAs, and severities.

The EPA(2005a) recognizes that there is rarely sufficient information about the MOA to scientifically justify a specific dose-response model. In the absence of a scientifically defensible and biologically-based dose-response model, the TCEQ makes reasonable health-protective assumptions about the relationship between exposures and toxicologic endpoints in epidemiology data by assuming that the relationship conforms to linearity at low doses of exposure.

Epidemiology data can be analyzed in diverse ways and for several different purposes. The analyses of epidemiology data for the purpose of risk assessment requires that the data be modeled in such a way that the model can be used to evaluate the risk for a target population different than the population included in the epidemiology study. A model based on epidemiology data describes the mathematical relationship between the observed mortality or disease incidence and the exposures. The form of the dose-response model should enable the separation of the effect of the agent being evaluated on the toxicity endpoint from the effects that

other factors may have on that endpoint. For example, the effect of background hazard rates and co-exposures in the epidemiology data should be part of the dose-response model when fitting or describing the epidemiology data but should be excluded from the model when evaluating risks for a target population with different background hazard rates and not exposed to other agents.

Obviously, the models that can be used for epidemiology studies depend on the availability of the data. The following subsections discuss various models and then provide guidelines indicating the different alternative models that can be used for different epidemiology data, depending on the type of data that were collected, evaluated, and reported. The TCEQ uses these guidelines to identify possible modeling approaches to investigate available, chemical-specific epidemiology data.

7.7.1 Individual Epidemiology Data

The full potential and best use of dose-response modeling of epidemiology data are possible only when the information is available at the individual (person) level rather than a group level. Information such as time-dependent exposure history, demographic characteristics (e.g., gender, race), lifestyle habits (e.g., smoking) relevant to the health endpoint under investigation, time-dependent co-exposures history to other potential agents that cause or may affect the health endpoint, etc., allow the researcher to use exposure-response models to best describe the relationship between health endpoints and potential explanatory variables. The TCEQ rarely has access to this type of data and relies on published results from epidemiologists (e.g., modeling for 1,3-butadiene by Cheng et al. 2007) or dose-response modelers who have been able to obtain the raw data (e.g., modeling for 1,3-butadiene by Sielken and Valdez-Flores 2011).

Ideally, a dose-response model should be such that the relationship between the health endpoint and the explanatory variables are biologically defined. Most biological processes that give rise to health endpoints, however, are complex in nature and generally not fully understood or developed and cannot be summarized via simple mathematical models. Researchers usually must rely on statistical methods to fit mathematical exposure-response models to the observed epidemiology data. As indicated above, in the absence of scientific justification to use a biological-based dose response model, the TCEQ makes the reasonable health-protective assumption that the relation between the dose and the health endpoint is linear at low doses.

7.7.2 Multiplicative Background Hazards Models

The risk of a specified health endpoint increases proportionally to the background rate for a specified value of the dose metric and values of the covariates in multiplicative background hazards models. These models are also known as relative risk, proportional hazards, or multiplicative background risk models. The multiplicative background models have been used extensively in modeling epidemiology data because of their flexibility, robustness and the assumption that the increase in risk is proportional to the background hazard rate.

Multiplicative background hazards models have been useful in modeling the incidence of cancer in humans exposed to radiation (Törnqvist and Ehrenberg 1994). According to Törnqvist and Ehrenberg (1994), a multiplicative model for cancer incidence is expected if the agent is an initiator and cause irreversible damage. The validity of multiplicative background hazards models is supported by analyses of epidemiology data of cancers related to ethylene oxide

(Törnqvist and Ehrenberg 1992). Multiplicative background hazards models have been used by USEPA (e.g., 1986, 2001, 2006); TCEQ (e.g., 2007, 2008); and several others (e.g., Harris, 1983, Crump 1994, Cheng et al. 2007, Sielken and Valdez-Flores 2009 and 2010). The National Research Council (BEIR V 1990) adopted the multiplicative background hazards model for radiation-induced cancers.

7.7.3 Additive Background Hazards Models

In additive background hazards models, the risk of a specified health endpoint increases by the same amount, regardless of the size of the background rate, for a specified value of the dose metric and values of the covariates. These models are also known as absolute risk models. Additive background models have not been used as often by USEPA as multiplicative background models in epidemiological risk assessment.

Additive background hazards models have been used only in limited occasions. Oftentimes additive models have been used in conjunction with multiplicative background hazards models (e.g., EPA 1986). Publications comparing additive and multiplicative models have not made any recommendations for using one model over another (e.g., Stayner et al. 1995), while some others have recommended using the multiplicative model (e.g., Törnqvist and Ehrenberg 1994).

7.7.4 Adjustment of Background Hazard Rate

The multiplicative or additive background hazards models discussed in Sections 7.7.2 and 7.7.3, respectively, used to model epidemiology data usually include a factor to account for the potential differences between a target population background rate and the underlying background rate in the cohort of the epidemiology study. For example, if the models were linear in the dose (i.e., the hazard rate in the dose-response model is a linear function of the dose), then they would be as follows:

$$\text{Additive Rate } \lambda(d) = \lambda_0 + \alpha + \beta \times d \text{ for the additive background rate model}$$

$$\text{Relative Rate } \lambda(d) = \lambda_0 \times \alpha \times (1 + \beta \times d) \text{ for the multiplicate background rate model}$$

where λ_0 is the cohort's estimated background hazard rate of the endpoint being analyzed, d is the dose measure (e.g., cumulative exposure in ppm-years), β is the slope (i.e., the change in the rate per unit increase in the dose), and α is the adjustment of the background hazard rate.

The target-population specific background hazard rates substitute for the estimated $\lambda_0+\alpha$ in the additive models and the $\lambda_0 \times \alpha$ in the multiplicative models in the evaluation of risks. For example, if a model is fit to the mortality of lung cancer in an epidemiology cohort, the risks for the population of Texas based on the model should be calculated using the Texas population background hazard rates of lung cancer mortality instead of the background hazard rate and adjustments estimated for the cohort. Age- and calendar-year- dependent background hazard mortality and incidence rates are regularly published by federal and state agencies (e.g., Surveillance Epidemiology and End Results of the National Cancer Institute at seer.cancer.gov).

7.7.5 Cox Regression

The Cox regression model fits a family of multiplicative background hazards models to epidemiology data. Cox regression is the preferred modeling methodology for health endpoints of epidemiology studies because of its statistical properties and widespread availability in software packages. Cox regression, as opposed to other methods, is more robust and does not require any assumptions about the underlying background hazard rates of the health endpoint. In addition, Cox regression can readily incorporate time-dependent covariates as well as fixed covariates (Cox 1972, Allison 2010). The covariate effects in Cox regression can be modeled as parametric or nonparametric effects. A parametric model assumes a specified functional form (e.g., linear or log-linear), and a nonparametric model does not assume a specified functional form. For example, regression models assume specified functional forms (e.g., linear or polynomial) and hence are parametric models. The magnitudes of the independent variable (e.g., dose) have an impact on the estimation of the model parameters. On the other hand, nonparametric models do not assume a functional relationship between the independent variable (e.g., treatments) and the response. The labels for the different treatments do not have a numerical significance and do not have an impact on the results.

Nonparametric modeling of covariate effects in Cox regression is especially useful when the effects do not have a clearly defined functional form (which happens frequently). The effects of fixed covariates can also be included by using stratified Cox regression whereby different strata are formed for each combination of the values of the stratifying covariates.

Cox regression has several other advantages. For example, because age is usually the main factor in the increased incidence of carcinogenic endpoints, it is of paramount importance to closely control the effect of age on health endpoints of epidemiology studies. Cox regression uses age as the index variable, adjusting for age in an optimal way. Also, Cox regression has an advantage over other methods in that exposure and other time-dependent covariates are treated as continuous variables that can take on any real value and do not have to be discrete values or group values. This feature of the Cox model avoids making extra assumptions that may increase the uncertainty of exposure estimates and of other measured or estimated covariates.

7.7.6 Poisson Regression

The Poisson regression methodology fits either multiplicative or additive background hazards models to epidemiology data. Poisson regression models require that individual person-years at risk be partitioned into different risk groups because these models operate on grouped data (Crump and Allen 1985). The effect of dose, for example, has to be modeled by creating dose intervals where the background hazard rates are approximately constant through the interval. These models (as opposed to the Cox proportional hazards models that use continuous measures of exposures or other time-dependent covariates) use groups of person-years at risk, grouped averages of exposure, and groupings of other time-dependent covariates to fit dose-response models. Poisson regression models assume that the hazard rate in any specific group is a constant through the intervals of time defining the group.

The number of observed individuals with the health endpoint under investigation in each group of person-years and each combination of time-dependent and fixed covariates characteristics is assumed to follow a Poisson distribution with a group-specific rate. Specifically, the number of

responses occurring in a particular group of exposure and a particular group of other covariates is assumed to take the values of r=0, 1, 2, … with the probability given by:

$$p(R=r) = (\lambda n)^r \times e^{-\lambda n} / r!$$

where p(R=r) is the probability that r is observed, r is the number of responses occurring in the group, n is the number of person-years in the group, λ is the unknown rate of occurrence of the response per person-year at risk (i.e., λn is the expected number of responses in the group). The parameter λ can be modeled using an additive or multiplicative background hazard dose-response model that depends on the dose and the covariates defining the different groups. For example, if each group of person years were defined by a dose interval, an age group and sex, a multiplicative model could be:

$$\lambda = \lambda(d,age,sex) = \lambda_0 \times \text{Effect of Age} \times \text{Effect of Sex} \times (1 + \beta \times d)$$

where the rate λ depends on the dose d, the age and the sex of the group. The "Effect of Age" can be represented by a parametric function or nonparametric estimates. The "Effect of Sex" is a nonparametric estimate and accounts for the difference between males and females in the response rate. The parameters λ_0, "Effect of Age", "Effect of Sex" and β are unknown and need to be estimated from the data. (See Appendix A for more details.)

The form of the group-specific rate λ is given by the specified model, and its numerical value is the expected number of responses observed in each group. The expected number of responses in a group is the product of the group-specific rate and the group-specific number of person-years at risk.

When individual (person) data are available, Cox regression methods are preferable over Poisson regression methods. When only grouped data are available, Cox regression methods cannot be used, but Poisson regression models can be used.

7.7.7 Parametric Dose-Response Models

Dose-response models for epidemiology studies usually incorporate a parameter that estimates the underlying background hazard rate for the unexposed individuals included in the study. Epidemiological dose-response models used for risk assessment are a parametric function relating the health endpoint being investigated and a measure of the dose from the carcinogenic agent being evaluated. As discussed in Section 7.7.4, for the calculation of excess risks, the underlying background hazard rate estimated for the epidemiology study is replaced by the underlying background hazard rate in the target population for whom risks of the health endpoint are to be estimated (e.g., the general public). The same exposure-response relationship estimated from the epidemiology study is used in the estimation of risks for the inference population (e.g., the Texas population, the US population, etc.).

Regardless of whether the data are individual or grouped or the background hazards are additive or multiplicative, the dose-response model is of paramount importance because it defines the shape of the curve that describes the relationship between a dose metric and a health outcome. The shape of the dose-response model has an important impact on the estimation of risks at low doses.

7.7.7.1 Linear Dose-Response Models

Generally, the dose-response model can be assumed to be a polynomial function of the dose metric. In the absence of mechanistic information about the carcinogen and the health endpoint, the linear exposure-response model is the most parsimonious and simplest polynomial that should be used to fit epidemiology data. There are at least three reasons for using linear exposure-response models for epidemiology data (Crump and Allen 1985):

1. A linear model is biologically plausible for carcinogens, particularly for genotoxic carcinogens or those acting at a site where cancers occur spontaneously (i.e., in the absence of the carcinogen of interest; Crump et al. 1976).

2. A linear model is considered to be conservative in the sense that other biologically plausible dose-response models would generally imply lower risks.

3. A linear model usually fits data adequately.

Although some authors have used more sophisticated exposure-response models for epidemiology data, there has not yet been any statistical evidence showing any superiority of these models over the linear model in describing the relationship between exposure or dose and cancer endpoints. Biological justification for nonlinear models may exist on a case-by-case basis.

The linear dose-response multiplicative background hazard rate model can be written as follows:

$$\lambda = \lambda_0 \times \text{Covariate Effects} \times (1 + \beta \times d)$$

where λ_0 is the background hazard rate, d is the value of the dose-metric and the "Covariate Effects" can be parametric functions or nonparametric estimates that describe the effect of variables other than the dose metric on the hazard rate. The dose-response model is said to be linear because the rate ratio is equal to:

$$\text{Rate Ratio} = \lambda / [\lambda_0 \times \text{Covariate Effects}] = 1 + \beta \times d.$$

7.7.7.2 Log-Linear Dose-Response Models

An alternative model that has been used extensively in modeling epidemiology data is the log-linear model. The log-linear dose-response model has very similar characteristics, advantages, shape, and statistical behavior as the linear models do, especially at low doses. The functional form of the log-linear model is such that the logarithm of the hazard rate of the health endpoint is linearly related to the dose metric. The log-linear model can be used in conjunction with Poisson regression but is especially useful in conjunction with Cox regression.

The log-linear dose-response multiplicative background hazard rate model can be written as follows:

$$\lambda = \lambda_0 \times \text{Covariate Effects} \times e^{\beta \times d}$$

where λ_0 is the background hazard rate, d is the value of the dose-metric and the "Covariate Effects" can be parametric functions or nonparametric estimates that describe the effect of

variables other than the dose metric on the hazard rate. The dose-response model is said to be log-linear because the logarithm of the rate ratio is linearly related to the logarithm of the product of the slope and the dose; that is:

$$\text{Ln(Rate Ratio)} = \text{Ln}(\lambda / [\lambda_0 \times \text{Covariate Effects}]) = \text{Ln}(\beta \times d)$$

where Ln(x) is the natural logarithm of x. The above equation can also be written as:

$$\text{Rate Ratio} = \lambda / [\lambda_0 \times \text{Covariate Effects}] = e^{\beta \times d}.$$

7.7.7.3 Log-transformed Dose and Supra-Linear Models

As indicated by Crump and Allen (1985), linear exposure-response models are "considered conservative in the sense that other biologically plausible dose-response models would generally imply lower risks." Some researchers have published dose-response models that are inherently supra-linear at low exposures (e.g., Steenland et al. 2003 and 2004). The increase of the hazard rate or relative risk of a supra-linear exposure-response model is faster at lower exposures than at higher exposures. These types of models are generally not biologically plausible and tend to grossly exaggerate the estimation of risks at low exposures (Crump 2005, Valdez-Flores et al. 2010, Ginevan and Watkins 2010). A power model is another name used for a log-transformed dose model.

An example of a multiplicative hazard background log-transformed dose model is given as follows:

$$\lambda = \lambda_0 \times \text{Covariate Effects} \times e^{\beta \times \text{Ln}(1+d)}$$

or, equivalently:

$$\lambda = \lambda_0 \times \text{Covariate Effects} \times (1+d)^{\beta}.$$

The value of 1 (or some other positive value) is usually added to the value of the dose d to avoid having an undefined logarithm when the dose is equal to zero.

Unrealistic supra-linear exposure-response models oftentimes result from exposure transformations that automatically render a supra-linear shape of the relationship between cancer incidence and dose measures. Inappropriately, dose metric transformations like the square root of cumulative exposure or the logarithm of cumulative exposure have sometimes been used in modeling epidemiology data even though linear dose-response models fit the data as well as models with transformed doses. Crump (2005) showed that even when the true dose-response relationship is linear in dose, the models based on log-transformed doses fit the data as a supra-linear function of dose in the presence of random error in the estimation of exposure. Using supra-linear exposure-response models can only be justified if there is sufficient biological or mechanistic data to support their application.

7.7.7.4 Splines and Nonparametric Estimates

Splines and nonparametric estimates of health endpoint incidences and exposure are useful techniques for purposes of preliminary evaluation or exploration of epidemiology data (e.g.,

Steenland and Deddens 2004). However, the TCEQ will typically not use these two techniques as exposure-response models for risk characterization based on the following reasons.

There has been some recent research promoting the use of splines in modeling epidemiology data (e.g., Steenland and Deddens 2004). Splines are mathematical functions that try to provide a piecewise description of the shape and behavior of observed epidemiology data. There are several assumptions that have to be made when fitting a spline to epidemiology data and the resulting spline depends on these assumptions. The simplest splines are given by piecewise linear dose-response models but could include piecewise polynomial models of any order. Splines, although useful, can be misinterpreted and misused as a guiding tool in model selection and model formulation. Splines should not be used as surrogates for biological or mechanistic interpretations of exposure-response relationships. For example, a threshold in a dose-exposure relationship should be included only if there is evidence that exposure below a specific value does not increase the risk of the incidence of the health effect being investigated.

A more basic technique than splines, but frequently just as useful as splines, is nonparametric estimation of the relationship between exposure and incidence, which also helps in inferring the shape of the exposure-response relationship. Also, model selection can be guided by the models ability to reflect the nonparametric estimates. Both splines and nonparametric estimates cannot be used for risk characterization and are useful only as tools in the exploratory analyses of epidemiology data sets.

7.7.8 Grouped Epidemiology Data

Individual worker information in epidemiology studies published in the literature is not available in most cases. Information for groups of person-time at risk, however, is sometimes reported for epidemiology studies in the open literature. Depending on the level of detail of summary epidemiology data published, exposure-response models can be fit to these summary data using Poisson regression. For example, if the number of individuals observed with a specified health endpoint and the corresponding standardized mortality ratio (SMR) for several intervals of an exposure metric are reported, a multiplicative exposure-response model can be fit using Poisson regression. If the information is further split into other categories (e.g., sex, race, plant, year of hire, smoking, etc.), then these data can be used to adjust for the effect of those factors on the incidence of the health effect. For example, the DSD for Arsenic and Inorganic Arsenic Compounds (Erraguntla et al. 2010) fit multiplicative relative risk models adjusting for year of hire to summary epidemiology data of arsenic exposures.

There are instances when odds ratios (ORs) or rate ratios (RRs) are reported (instead of SMRs) for the number of individuals with the health endpoint for groups of workers exposed to different exposure intervals. These data cannot be used for exposure-response modeling using Poisson regression because ORs and RRs do not include sufficient information to estimate the expected number of individuals with a specific health effect in each exposure interval. If there are no better epidemiology data, the ORs can be used as surrogate estimates of the rate ratios or relative risks (RRs) and fit an exposure-response model using least squares methodology. For example, the DSD for Nickel and Inorganic Nickel Compounds (McCant et al. 2009) fit a least squares linear dose-response model to summary epidemiology RRs of lung cancer (in addition to other analyses).

It is important to note that the OR and the RR are, by definition, equal to one for the reference or "control" group. The reference or "control" group may or may not be exposed to any agent causing the health effect under investigation. In addition, the underlying background hazard rate for the health effect in the epidemiology study is likely different than the background hazard rate for the health effect in the target population for whom the risks are to be estimated (e.g., the general public). It is thus important, when fitting a dose-response model to RRs or ORs, to include an intercept in the model with a parameter to estimate the underlying background hazards rate of the health endpoint being investigated. The estimate of the intercept (the estimate of the underlying background hazards rate in the epidemiology cohort) is replaced by the age-dependent and population-specific background hazards rate in the target population for the characterization of risks. For example, the DSD for Nickel and Inorganic Nickel Compounds (McCant et al., 2009) fit a least squares linear dose-response model with a multiplicative intercept to summary epidemiology RRs of lung cancer.

7.7.9 Limited Epidemiology Data

Sufficient epidemiology data to fit a dose-response model to the observed health endpoint of interest are not always available. Oftentimes only minimal summary data are reported in published articles. When only limited information is available, the researcher has to make the best use of these limited data to develop a dose-response relationship. Regulatory agencies have often relied on simple linear regression models whenever there are only limited data (e.g., USEPA 1986). One such example is when results (e.g., SMRs) are published only for the epidemiology group as a whole and hopefully an estimated average exposure for the whole group is given. A simple linear model that goes through one at zero exposure and through the SMR at the average exposure can define the linear exposure-response. This limited estimate of a model is justified only if there are no better data that can be obtained. The maximum likelihood estimate of the slope would just be equal to the SMR divided by the average of the exposure. The slope of the linear multiplicative dose-response model is then equal to:

$$\beta = (SMR - 1) / d$$

where d is the average value of the dose metric for all the individuals in the epidemiology study and the SMR is the ratio of the hazard rate in the individuals in the epidemiology study and the hazard rate of a reference population , such as the U.S. or Texas population. This type of model has been referred as an average relative risk model because is based on an average estimate of risk (SMR) and an average estimate of the dose.

7.7.10 Covariate Effects

Dose-response modeling of health endpoints in epidemiology studies should be adjusted by the epidemiologist for the effects of any relevant factors. These factors are usually called covariates because they vary from one individual to another. The epidemiologist should provide a discussion of covariate effects in the published paper, and covariate effects should be discussed in the uncertainty section of the DSD (see Cheng et al. 2007 for an example). Ideally, the covariates included in the model should be selected based on biological or mechanistic arguments. However, if there is no biological or mechanistic information that dictates which covariates should be included in the dose-response model, the impact of the covariates in the health effect being investigated can be assessed using statistical methodology. Likelihood ratio

tests can be used to determine whether a covariate contributes significantly to the explanation of the relationship between exposure and incidence of a health endpoint.

The covariates can be time independent (e.g. race, gender) or time dependent (e.g. age, years since hire, co-exposures to other agents, jobs, or plants). The effect of the covariates on the model fit to the data can be evaluated, with the one covariate with the most significant impact on the likelihood included first. This process can be repeated including one covariate at a time until all covariates that make a significant improvement in the likelihood of the fit are included in the exposure-response model. Sielken and Valdez-Flores (2011) provide an example of this type of analysis for 1,3-butadiene. There are additional procedures for selecting which covariates to include.

Statistically-based criteria to include covariates lead to more robust and less subjective modeling. However, if there is any biological or mechanistic indication that a covariate should be part of the dose-response model, then the covariate should be included in the analysis. It is better to risk increasing the uncertainty in the estimate of the dose-response relationship by adding potentially unnecessary covariate effects than to introduce the bias resulting from excluding important covariate effects (Checkoway et al. 1989, Rothman 1986, Breslow and Day 1980 and 1987).

The estimates of the underlying background hazard rate and of the covariate effects and other adjustments to the underlying background hazard rate are replaced by the age-dependent background hazard rates observed in the target population for purposes of risk characterization. In the risk characterization step of the risk assessment, only the dose-response component of the model is used because the target population's underlying age-dependent background hazard rates are intrinsically adjusted for any other relevant factors, or are not part of the target population's experience of the health endpoint being studied. For example, if the following multiplicative background dose-response model were used:

$$\lambda = \lambda_0 \times \text{Covariate Effects} \times (1 + \beta \times d),$$

then only the slope parameter (β) estimated from the model is needed for the characterization of risks in a specified population. In addition, the estimated background hazard rate ($\lambda_0 \times$Covariate Effects) is replaced by the target population hazard rate in the characterization of risks for such population. Sielken and Valdez-Flores (2009a, 2009b) are examples of how to calculate population risks from parameters of dose-response models fitted to epidemiology data.

7.7.11 Goodness of Fit

Goodness of fit of dose-response models fit to epidemiology study data involve groupings of the numbers of individuals with the health endpoint under investigation (e.g., Breslow and Day 1980 and 1987). The goodness of fit of the model is then judged by the lack of fit of the model that compares the likelihood of the model's fit to the data with the likelihood of the observed data. The difference in likelihoods is then compared with a chi-square distribution with the appropriate degrees of freedom.

Because Poisson regression requires data to be grouped, goodness of fit tests for models based on Poisson regression are usually easy to perform. Methods similar to assessing the goodness of fit based on Poisson regression models can be used in assessing the goodness of fit based on other

models but the data has to be grouped first in order to perform the test. One procedure that has been used is to group the data into risk groups based on the risk predicted by the model and then compare the number of predicted cases with the number of observed health endpoint cases in each group (Lemeshow and Hosmer 1982).

Graphical display of ORs and RRs cannot be used to judge goodness of fit or to compare alternative models fit to epidemiology data. Nonparametric estimates of ORs and RRs are relative to an estimate of the underlying background hazard rates and oftentimes are adjusted for covariate effects. The estimated background hazard rates and estimated covariate effects for nonparametric ORs and RRs are different than the estimated background hazard rates and estimated covariate effects for parametric dose-response models. Thus, plots of nonparametric estimates of ORs and RRs cannot be compared to fitted dose-response models.

7.8 Quantitative Cancer Exposure-Response Characterizations

The TCEQ derives URF and SFo values consistent with the following information. Some definitions of quantitative exposure-response characterizations in the context of the presence or absence of a response like cancer (either cancer mortality or cancer incidence) are as follows (although the following sections refer to effective concentration for calculation of an inhalation URF, it is also applicable for effective dose for calculation of an oral SFo).

URFs express cancer potency in terms of risk per unit air concentration (e.g., risk per $\mu g/m^3$) assuming continuous environmental lifetime exposure. They are calculated using linear low-dose extrapolation when the carcinogenic MOA is mutagenic or the MOA is unknown. When a dose-response curve is modeled for tumor or cancer mortality data (Figure 7.1), the URF is the slope of a straight line from the POD to the origin, with the POD being the lowest tumor response or cancer mortality response supported by the study data. (Specifically, the curve is a graph of the frequency (or probability) of the specified response versus exposure concentration.)

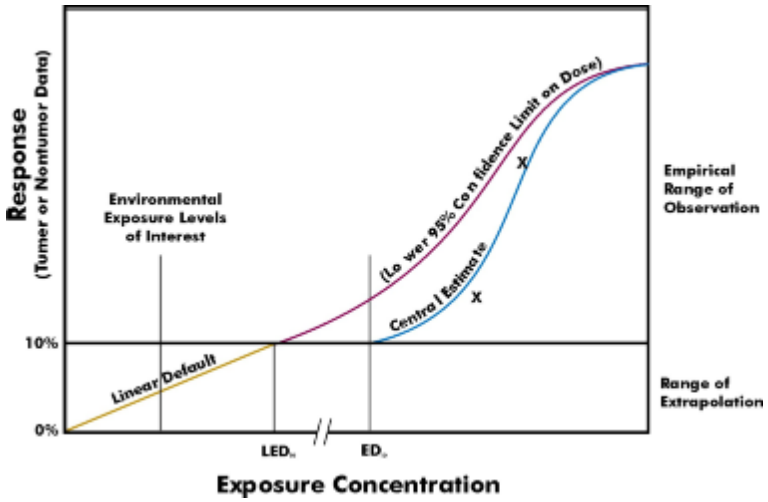


Figure 7-1 Example of a linear approach to extrapolate to lower exposures

The terms “ED10 and LED10” refer to dose but are analogous to the terms “EC10 and LEC10”, respectively, which refer to exposure concentrations (Exhibit 12-3A of USEPA 2004).

The effective concentration (EC) is defined as the best estimate of the exposure concentration in the inference population corresponding to a specified excess risk of the specified response (either cancer mortality or cancer incidence). The exposure concentration is generally thought of as being at a constant level for a lifetime. Strictly speaking, the exposure scenario in the inference situation (public environmental exposure in this case) is assumed to either be the same as the exposure scenario in the exposure-response modeling or adjusted to be the same as the exposure scenario in the exposure-response modeling (e.g., adjusting from an environmental scenario to an occupational scenario). For example, an EC_{001} is the concentration in a specified inference population corresponding to an excess risk of 0.001 (1 in one-thousand).

Two definitions of excess risk are commonly used. One definition of excess risk is added risk where:

added risk = P(C) – P(0)

and

P(C) = the probability of the specified response when the exposure concentration is C units greater than the background concentration level, and
P(0) = the probability of the specified response when the exposure concentration is at the background concentration level.

A second definition of excess risk is extra risk where:

extra risk = [P(C) – P(0)] / [1 – P(0)] = added risk / [1 – P(0)].

Added risk is the absolute increase in the probability of the specified response. For example, an added risk of 0.10 means that in an inference population of 100,000 individuals, it is expected that there will be 0.10 × 100,000 = 10,000 more individuals with the specified response if the exposure concentration is C units greater than the background concentration level compared to the expected number of individuals with the specified response if exposure concentration is the background concentration level. That is, an increase of 10,000 individuals with the specified response. On the other hand, the meaning of extra risk (which is always greater than or equal to the added risk) depends on P(0), the probability of the specified response when the exposure concentration is at the background concentration level. For example, if 25% of the inference population is expected to have the specified response when the exposure concentration is at the background concentration level (i.e., P(0)=0.25), then an extra risk of 0.10 implies that, of the 75% (i.e., 100% - 25%) of the inference population that is not expected to have the specified response when the exposure concentration is at the background concentration level, 10% are expected to have the specified response when the exposure concentration is C units greater than the background concentration level. That is, an increase of 0.10 × 0.75 × 100,000 = 7,500 individuals with the specified response. Although the USEPA generally defines their excess risk as extra risks, when they communicate excess risk to the public it is generally interpreted by the public as if it were added risk. This potential for misinterpretation would be avoided if excess risk were defined as added risk. For rare responses, added and extra risks are very similar. Currently, the TCEQ generally uses estimates of extra risk when deriving carcinogenic toxicity factors (e.g., URF, SFo).

In the definitions of excess risk, "0" refers to the background concentration level which may or may not be zero. Similarly "C" refers to the concentration above and beyond the background concentration level and not the total concentration of exposure.

Returning to the concept of an EC, the 95% lower confidence limit (lower bound) on the EC is denoted by LEC. The numerical value of the LEC depends on the method of calculating the lower bound. Different computer software programs use different methods, and no one method is universally acknowledged as best. The LEC reflects only part of the uncertainty regarding the inference population's exposure-response relationship. In some circumstances (e.g., when the observed dose-response data is nonlinear with no increase in the number of observed responses at the lower doses), the LEC is much less responsive to the observed epidemiology data than the EC, and the LEC is determined more by study designs and statistical assumptions than the observed data. Consequently, for the purposes of comparing the potency of two toxicants or substances, the EC or the corresponding URF (MLE) (rather than the LEC or the corresponding URF (95% UCL)) is the best basis for comparison (USEPA 1995, 2000a, b). The TCEQ reports EC values as well as LEC and 95% upper confidence limit (upper bound or UEC) values in the DSDs, as recommended by USEPA (2005a).

The specified excess risk level (BMR) in the definition of an EC (or LEC) can be any value between zero and one, although in practice it is generally the lowest BMR supported by the data. The BMR is a probability for a dichotomous endpoint like cancer mortality or incidence (the BMR for a continuous endpoint such as a measurement may be defined differently). For example, EC_{10}, EC_{01}, EC_{001}, etc., correspond to excess risks of 0.10, 0.01, 0.001, etc. When the EC (LEC) is used as a POD for extrapolation to low-exposure levels, then the BMR should be chosen so that the EC (LEC) is within the observed data and should not be so small that it is unnecessarily dependent upon the assumed shape of the exposure-response model.

If the probability of the specified response in the exposure-response model includes time (age), then the excess risk and the definition of the EC also includes a specified time (age). For example, most exposure-response models used for epidemiology data incorporate the time at which a response is observed. In these models, the excess risk and the definition of the EC also include a specified time (age). In the calculation of excess risk, it is assumed that the exposure scenario remains unchanged up to that time (age) (i.e., a constant exposure concentration up to that age is presupposed). Furthermore, it is assumed that the estimated exposure-response model is appropriate up to that time (age). For example, if an exposure-response model is estimated using occupational epidemiology data that only includes workers up to age 65 years, then calculating an excess risk up to age 70 years or higher involves an extrapolation over age that may or may not be warranted. Also, because the excess risks and ECs are often heavily dependent upon the specified time (age), it is important to consider what the specified time (age) is when interpreting the results. For example, time might be age and the specified time be set to 70 years. In which case, the excess risk refers to the excess risk by age 70 years. As is common in regulatory risk assessment, the TCEQ uses a default exposure duration of 70 years as discussed in Chapter 1. Another reason to use an exposure duration of 70 years for calculation of excess risk using epidemiology data is that the background rates of the disease and survival rates for a population used in the life-table analysis (BIER IV approach) discussed in the following section are more uncertain after 70 years.

Some exposure-response models (e.g., the multistage model) do not include time or age but rather consider the presence or absence of the specified response during the lifetime of the individual. For these models, extra risks and ECs refer to lifetimes rather than a specified time (age).

Returning to definitions, the BMD (BMC) in the context of a dichotomous response (e.g., the presence or absence of a specified tumor) is analogous to the EC. The 95% lower confidence limit (lower bound) on the BMD (BMC) is denoted by BMDL (BMCL). The meaning of the BMR is the same for BMD (BMC) as it is for EC.

7.9 Excess Risk Calculations for the General Population

Some definitions of quantitative dose-response characterizations in the context of the presence or absence of a response like cancer (either cancer mortality or cancer incidence) are given in Section 7.8. The calculation of excess risk for the inference population (i.e., the Texas general population) is the focus of the current section. The TCEQ uses these calculations to derive toxicity factors (e.g., URF, SFo).

Dose-response models that are multiplicative background response models (or additive background response models) characterize the effect of exposure as a multiplier of (or in addition to) the background hazard rate. When the dose-response model is being estimated (and the estimate of the slope, β, multiplying dose in the linear portion of the model is being determined), the background hazard rate should correspond to the epidemiology study cohort. Because quantitative risk characterizations are for a specific inference population, when the hazard is being characterized for a specific inference population (e.g., the Texas general population), the background hazard rate should correspond to that specific inference population (e.g., the general public in Texas). An inference population can be relatively general group (the entire US) or more specific group (e.g., Texans, male Texans, female Texans, Texans of a specified race, all residents within 100 miles of a specified location). The general population of Texas is the inference population generally evaluated by the TCEQ.

Quantitative risk characterizations for different inference populations are usually different, although the differences in the derived URF values are often slight. For example, the URF for 1,3-butadiene based on Texas background rates was 1.097×10^{-6} per ppb whereas the URF based on US background rates was 1.062×10^{-6} per ppb (Grant et al. 2009). Similar small differences in the URFs based on US background rates as opposed to Texas background rates were obtained for nickel (TCEQ 2011), silica (TCEQ 2009), and arsenic (TCEQ 2010).

7.9.1 Life-Table Calculations for Excess Risks

When epidemiology data are modeled (Section 7.7) and the corresponding dose-dependent adjustment to the background hazard rate is identified (Section 7.7.4) in preparation for the calculation of excess risks, the maximum likelihood estimate (MLE) of the slope, β, multiplying dose in the linear portion of the model is obtained as well as the SE of the estimate of β. Using a standard normal distribution, the 95% LCL and 95% UCL on the slope β are calculated as follows:

$$\beta\ (95\%\ \text{LCL}) = \beta - (1.645 \times \text{SE}), \text{ and}$$

$$\beta\ (95\%\ UCL) = \beta + (1.645 \times SE).$$

These characterizations (MLE, 95% LCL, and 95% UCL) of the slope estimate β are used to calculate an air concentration (or oral dose) corresponding to a known risk level using a life-table calculation for excess risk. From this air concentration (or oral dose), the URF or SFo value (i.e., increase in risk for the general population per ppb or $\mu g/m^3$ or per mg/kg-day) can then be determined. For example, if the LEC_{10} is used as the POD, then the slope of the line from the LEC_{10} to the origin yields the inhalation URF (95% UCL), that is, the upper-bound excess lifetime cancer risk estimated to result from continuous lifetime exposure to an agent at a concentration of 1 $\mu g/m^3$ in air:

$$URF\ (95\%\ UCL) = 0.10 / LEC10.$$

Life-table calculations are sequential calculations that follow an individual from birth to a specified age. The life-table method is sometimes called the actuarial method. In the life-table method, for each year of an individual's lifetime (year 1 from birth to the age 1 birthday, year 2 from the age 1 birthday to the age 2 birthday, etc.) the life-table calculation incorporates the age-specific values of the individual's exposure (the exposure metric in the exposure-response model), the background (all-cause) survival probability, the background probability of the specified response, the effect of the exposure on the probability of the specified response, and any ADAFs. The life-table method of calculating excess risks is described in the BEIR IV report (NRC 1988). Computational details of the BEIR IV methodology are also described in Sielken and Valdez-Flores (2009b). Sielken & Associates have prepared an EXCEL implementation of the BEIR IV methodology for both incidence and mortality responses for the TCEQ (Valdez-Flores and Sielken 2010). Because life-table calculations are computationally intensive, there have been simpler alternatives to the BEIR IV methodology proposed in the early literature (e.g., USEPA 1986, Gail 1975). However, high speed computers have eliminated the need for such approximations.

When calculating excess risk for the inference population (e.g., the Texas general public), the portion of the dose-response model fit to epidemiology data corresponding to the estimate of the background hazard rate is replaced by the inference population background hazards rates. That is, inference population background hazards rates are combined with the dose-response model fit to the epidemiology data (excluding the estimated background hazard rate estimated for the epidemiology data) in the characterization of risks. Population-specific background hazard rates depend on age, sex, race and other factors that need to be incorporated into the characterization of risks. The TCEQ uses appropriate methodology (e.g., life-table calculations) to take into account all these factors when characterizing lifetime risks to the extent possible and necessary, which is usually accomplished by incorporating rates for the general population of Texas that inherently reflect these factors.

7.9.2 Characterizing Risks for Older Ages and Different Health Endpoints

There are several alternatives in the calculation of excess risks that would increase the uncertainty associated with the corresponding calculated excess risks. Methods to account for this uncertainty are discussed in Section 7.10. Characterizing risks at low environmental doses usually well below the range of doses in the epidemiology study used to fit the model increases the uncertainty of the estimates.

Similarly, using dose-response models to characterize risks at ages other than those observed in the epidemiology data adds uncertainty to the risk estimates. Because many cancer responses have a background hazard rate that increases greatly at older ages, the choice of the terminal age in a life-table calculation from birth to a specified (terminal) age can have a substantial impact on calculated excess risks. Any comparisons of excess risk across chemicals should reflect any differences in the specified terminal age. The choice of the terminal age for a specific case should reflect the reasonableness of the assumption that individuals in the inference population would be exposed at the older ages. In addition, the choice should reflect the reasonableness of extrapolating from the exposure ages in the epidemiology study to the exposure ages being assumed for the inference population. As mentioned previously, consistent with standard risk assessment practice, the TCEQ uses 70 years as the default exposure duration to calculate URF and SFo values.

Uncertainty is also increased if the endpoint used in calculating excess risks is different than the endpoint used in the dose-response modeling. For example, in USEPA (2006) one health endpoint was used in the dose-response model fitting and a different health endpoint was used to calculate excess risks. There the dose-response model and the estimated β slope used mortality as the health endpoint (i.e., death with the specified cancer), but the health endpoint used to calculate excess risks was incidence (presence of the cancer but not necessarily death with the cancer). It is most appropriate, when excess risks for the inference population are being calculated, for the health endpoint to be the same health endpoint as was used in the dose-response modeling. Here, mortality refers to death from or with the disease whereas incidence is the onset or diagnosis of the disease that may or may not result in death. Similarly, an exposure-response model that has estimated a β slope using an incidence response is appropriate when excess risks using a life-table analysis are being calculated for that same response (i.e., incidence of the specified response as opposed to the mortality with/from the specified response). The TCEQ does not generally use a mortality-based exposure-response model as the basis for the calculation of excess risks for an incidence response, or vice versa. The computational details of the BEIR IV methodology are different for incidence and mortality as shown in Sielken and Valdez-Flores (2009b).

As a general rule, the health endpoint used for dose-response modeling and excess risk calculation should match. However, there are instances (i.e., exceptions) in which a model based on mortality can be used to approximate (represent) the risk for incidence (e.g., TCEQ 2009, 2009). For example, in the silica DSD, TCEQ (2009) justified using the mortality-based lung cancer model to characterize risk of lung cancer incidence because “mortality rates for lung cancer are high and correlate well with incidence.” Another example of using a model for one health endpoint to characterize risks for a different health endpoint is given in the Arsenic DSD (TCEQ 2010). There, the TCEQ used models fit to respiratory cancer mortality to characterize the risk of lung cancer incidence. TCEQ (2010) stated that “respiratory cancer mortality data …. are a reasonable surrogate for lung cancer as most (96 %) of the observed deaths …. were due to lung cancer,” adding “lung cancer mortality, and consequently respiratory cancer mortality, are reasonably predictive of lung cancer incidence.” Consequently, for arsenic, where respiratory cancer mortality was used to "represent" lung cancer mortality, lung cancer background mortality rates were used in the life-table calculation of excess risks in lieu of respiratory cancer mortality rates.

7.9.3 Dosimetric Adjustments

As indicated in Section 7.6, the exposure or dose metric used in the dose-response model and the excess risk calculation using the life-table analyses should be the same. For example, if the dose metric is cumulative exposure, then the method of calculating cumulative exposure should be exactly the same in both the dose-response model and the excess risk calculation. This includes any weightings, lags, windows of exposure, etc.

The dose-response model used in the BIER IV life-table analysis should be the same as that used to fit the epidemiology data. For example, if the dose-response model were the multiplicative background linear model given by

$$\lambda = \lambda_0 \times \text{Covariate Effects} \times (1 + \beta \times d),$$

then the slope β estimated by fitting the epidemiology data should be used in the life-table analysis to calculate excess risks using the same dose-response model (after some adjustments to reflect the differences between the exposure profiles in the epidemiology study and the inference population).

In order to calculate excess risks for environmental exposures, the units of exposure in the inference situation (e.g., environmental exposures to the Texas general public) need to be converted to the units of exposure in the estimation situation (e.g., the occupational exposures in the dose-response model estimated using the epidemiology data). Environmental concentrations for the general population ($\text{Concentration}_{HEC}$) are converted to Occupational concentrations ($\text{Concentration}_{OC}$) using the following equation:

$$\text{Concentration}_{OC} = \text{Concentration}_{HEC} \times (VE_h/VE_{ho}) \times (\text{days per week}_{res}/\text{days per week}_{oc})$$

where:

VE_h = non-occupational ventilation rate for a 24-h day (20 m^3/day),
VE_{ho} = occupational ventilation rate for an 8-h day (10 m^3/day),
$\text{days per week}_{res}$ = residential weekly exposure frequency (7 days per week), and
$\text{days per week}_{oc}$ = occupational weekly exposure frequency (default of 5 days per week).

7.9.4 Adjustments for Early-Age Exposures

USEPA's Supplemental Guidance (USEPA 2005b, Sections 5 and 6) documented their procedure for incorporating ADAFs into lifetime excess risk calculations if the chemical acts through a mutagenic MOA. As detailed in Sielken and Valdez-Flores (2009a), USEPA's first attempt to implement an ADAF when the dose–response model had a cumulative dose metric failed to successfully follow USEPA's own guidelines (USEPA 2005b). The failure overstated the impact of ADAFs by approximately 8,000 fold. Because cumulative exposure is a common dose metric in dose-response models of epidemiology data, if it is decided to incorporate ADAFs following EPA guidelines, the TCEQ will incorporate ADAFs into the life-table analyses using the BIER IV approach using procedures outlined in Sielken and Valdez-Flores (2009a).

7.10 Determination of URFs and SFo Values from Dose-Response Modeling

As indicated in Section 7.9, the risk of an adverse health endpoint for a specific population exposed to a specified dose can be calculated using life-table calculations once a dose-response model has been fit to the epidemiology data. Dose-response models are fit to epidemiology data with individuals usually exposed to high doses. Several different shapes of dose-response models can fit the same epidemiology data equally well in the observed range of the data but may have very different behavior at doses below the range of the observed data. That is, for example, the risks predicted by different models at doses in the observed range can be similar but risks predicted at doses much lower than the observed doses may be very different.

The observed doses in epidemiology studies are usually much greater than the doses of interest in risk characterization (i.e., typical environmental doses). Risk estimates at doses much lower than the doses in epidemiology studies are subject to potentially substantial uncertainty. In the interest of accounting for the effect of uncertainty on risk estimates below the dose range in epidemiology studies, the TCEQ uses default, health-protective methods to calculate low-dose risks. The TCEQ uses procedures consistent with those discussed below when deriving carcinogenic toxicity factors (e.g., URF, SFo).

7.10.1 Linear Model

If the MOA is mutagenic or the MOA is unknown, then the default is to determine a POD based on the observed data and perform a linear extrapolation from the POD to determine the URF or SFo (Chapters 3 and 5).

The 2005 USEPA Guidelines for Cancer Risk Assessment defines a POD as marking

> "... the beginning of extrapolation to lower doses. The POD is an estimated dose (usually expressed in human-equivalent terms) near the lower end of the observed range, without significant extrapolation to lower doses."

This particular definition of the POD is trying to ensure that the POD reflects the observed exposure-response information without extrapolating beyond the observed data and without having to unduly depend on the assumptions or choices underlying the estimated dose-response model. Depending on the characteristics of the epidemiology study, the selected endpoint, and the underlying exposure-response relationship, the POD of an epidemiology study can be the dose corresponding to an excess risk of 1/100 or 1/1,000 or in some cases as low as 1/100,000 or 1/1,000,000 (e.g., ethylene oxide (Valdez-Flores and Sielken 2010) and butadiene (Sielken and Valdez-Flores 2011)). In contrast, the POD from an animal study is typically the dose associated with an excess risk of 1/10. The important point is that the POD should be in the range of the observed data -- "near the lower end of the observed range, without significant extrapolation to lower doses" (USEPA 2005a, page 1-13).

Given the intent of the POD to reflect the observed data without over-dependence on the exposure-response modeling, the time/age in the definition of the EC, BMD, etc. should be within the range of the observed data. Most epidemiology studies follow-up with workers even after retirement age and several individuals live past the age for which the POD is to be estimated. However, if workers in an epidemiology study are observed only until a limited age

(e.g., prior to retirement at age 65 years), then for the purpose of establishing a POD, the corresponding exposure-response model should not be extrapolated substantially beyond the age of 65 years. Extrapolations beyond the range of observation (i.e., below the observed exposure levels, above the observed exposure levels, or beyond the ages observed in the study) should be discussed in an uncertainty analysis.

Again, if the MOA is mutagenic or the MOA is unknown, then the default approach would be to determine a POD and perform a linear extrapolation from the POD (Chapters 3 and 5 in this guidance document). If the excess risk is to be linearly extrapolated below the POD, then the URF is defined as:

$$\text{URF} = (\text{excess risk at POD}) / \text{POD}$$

For example, if the POD is the concentration corresponding to an excess risk of 1 in 1,000 (i.e., $EC_{0.001}$), then:

$$\text{URF (MLE)} = 0.001 / EC_{0.001}$$

which is an assumed rate of increase (slope per unit concentration) between zero excess risk at concentration zero and an excess risk of 1/1,000 at concentration $EC_{0.001}$.

The URF can be denoted on whether it is a best estimate or a bound. For example, URF (MLE) is based on the maximum likelihood estimate of the concentration (e.g., EC or BMD) with the specified excess risk. On the other hand, URF (95% UCL) is based on the lowest concentration that has a 95% upper confidence limit on the excess risk equal to the specified excess risk. The URF (LEC) is the lowest concentration that has a 95% upper confidence limit on the excess risk equal to the specified excess risk. Thus, for example,

$$LEC_{0.001} < EC_{0.001}, \text{ and}$$

$$\text{URF (95\% UCL)} = 0.001/ LEC_{001}$$

is an upper bound on the URF, and URF(95% UCL) is greater than

$$\text{URF (MLE)} = 0.001/ EC_{0.001}.$$

The lower bounds (LEC or BMDL) on the concentration with a specified excess risk are not very responsive to the observed dose-response data. That is, very different observed dose-response data (for the same study design) may result in very similar LEC and BMDL values. Given the non-responsiveness of the LEC and BMDL and the overestimation of the likely true low-dose risk when the POD is a lower bound instead of a best estimate, URF (95% UCL) values are a poorer basis for comparing the potency of different chemicals than URF (MLE) values.

URF (95% UCL) values reflect the uncertainty present in the dose-response data. However, because they are statistical bounds and not maximum likelihood estimates, URF (95% UCL) values derived from epidemiology data may not be the best estimates for risk management decisions based on the above discussion (see also Section 7.8 and USEPA 2000a,b).

The TCEQ will provide the URF (MLE), as well as the URF (95% LCL) and the URF (95% UCL). The URF (MLE) is preferred because it is, by definition, the estimate that maximizes the likelihood of the observed data, and therefore, the best estimate to be used. This is especially true in situations where URFs from different studies are combined. However, ultimately, scientific judgment is used to decide what estimate of the URF is most applicable based on these considerations, MOA information, and other chemical-specific information.

The basic procedures used to derive URFs are also used to derive SFo values. Both are the slopes in an assumed linear extrapolation between the excess risk at a POD and zero excess risk at zero concentration. As discussed previously, URFs generally refer to studies involving inhalation exposure where the POD is a concentration. SFo values generally refer to studies involving oral exposure where the POD is a dose. For example, a SFo has a POD in units of mg/kg-day and is the rate of increase (slope per unit dose) between zero excess risk at zero dose (mg/kg-day) and a specified risk at the ED, BMD, LED or BMDL (assuming a linear extrapolation below the POD).

7.10.2 Nonlinear Models

If there is sufficient MOA information indicating a nonlinear dose-response relationship at low doses, then a nonlinear approach can be used to extrapolate risks to doses below the POD. EPA (2005a) guidelines state that "the linear approach is used when: (1) there is an absence of sufficient information on modes of action or (2) the mode of action information indicates that the dose-response curve at low dose is or is expected to be linear. Where alternative approaches have significant biological support, and no scientific consensus favors a single approach, an assessment may present results using alternative approaches. A nonlinear approach can be used to develop a reference dose or a reference concentration."

Nonlinear low-dose extrapolation can refer to different types of low-dose extrapolations (See Pottenger et al. 2011 for a general discussion.). For instance, nonlinear often refers to dose-response relationships that have a threshold, in which case a reference dose or a reference value can be developed based on procedures in Chapters 3 and 5 in this guidance document. In other circumstances, nonlinear refers to dose-response relationships which are not linear throughout the range of doses, although risk may be linear at lower doses. For example, Kirman et al. (2004) estimated the dose-response relationship for ethylene oxide using a quadratic relationship between concentration and cancer risk and used this fitted model to determine a POD. Then, Kirman et al. compared low-dose quadratic and low-dose linear extrapolations below the POD. This comparison followed the general guidance assumptions that linear extrapolations of risks for doses below the POD should be presented alongside nonlinear extrapolations even if there is sufficient biological information justifying the low-dose nonlinear extrapolation relationship.

7.11 Meta-Analyses

Meta-analysis is a technique used to combine and summarize results from several different independent analyses. The TCEQ may use meta-analyses as appropriate when several epidemiology studies are being used to derive a carcinogenic toxicity factor, depending on time and resource constraints. There are two types of meta-analyses; qualitative and quantitative. Qualitative meta-analyses include little to no quantitative manipulation of the results from individual analyses. Qualitative meta-analyses are for the purpose of summarizing, comparing

and contrasting results from different sources. Quantitative meta-analyses, on the other hand, are statistical methods used to combine individual results into a summary value. The focus of these guidelines is on evaluating and conducting quantitative meta-analyses of epidemiology studies published in the scientific literature for the purpose of risk estimation. In 1995, the ILSI Risk Science Institute and the Office of Research and Development, Office of Health and Environmental Assessment, and USEPA funded a group of scientists to develop guidelines for the application of meta-analysis in epidemiological assessments (Blair et al. 1995). The application of meta-analyses to epidemiology studies has evolved and their application has been expanded since that time. The guidelines presented here summarize and supplement the guidelines published by Blair et al. (1995). The TCEQ may perform meta-analysis of risk measures (e.g., URFs) (Section 7.11.3 Meta-Analyses of Risk Measures) or meta-analysis of slope estimates (e.g., β values) (7.11.4 Meta-Analyses of Slope Estimates) depending on the availability of data and resources.

The traditional notion of quantitative meta-analysis is that of calculating a single risk measure from a set of risk measures from different independent individual studies. Epidemiology studies with agents and health endpoints thoroughly researched offer more data than simple estimates of risk. Meta-analyses based on these data-rich studies can be performed at a higher level by modeling combined summary exposure-response data (rather than combining summary results as is done in the traditional meta-analyses) to estimate a risk that is based on the combined dose-response data rather than the combined individual measures of risk.

There are several necessary steps in the performance of a quantitative meta-analysis of epidemiology studies. First, the epidemiology studies relevant to the agent of concern and the specific health endpoint have to be identified. Then, the studies that meet qualitative inclusion criteria set *a priori* need to be identified and selected to be part of the quantitative meta-analysis. Depending on the information available for each of the selected studies, the results are either used as reported or re-calculated/verified. Then, the selected epidemiology studies can be combined in several different ways using meta-analysis techniques. The next sections will discuss in detail guidelines to follow in the performance of meta-analyses given different alternative levels of information available. Appendix B provides a case study in the context of arsenic (see also TCEQ 2010).

7.11.1 Identification of Individual Studies

In this step, all the literature corresponding to the chemical of concern should be identified. The most reliable and updated results should be preferred over outdated and less relevant analyses. The studies should include published and unpublished results and data. The studies identified can be used for a WOE narrative in a risk assessment document even if they do not meet the selection criteria to be included in the meta-analysis.

7.11.2 Selection of Individual Studies for Quantitative Meta-Analysis

Once all studies relevant to the agent of concern have been identified and a WOE assessment has been performed, studies for inclusion in the meta-analysis should be selected. The studies selected for meta-analysis are a subset of the studies identified for WOE. In the process of evaluating the WOE, it should become clear what health endpoint(s) is the main concern. The studies selected should meet specified criteria to be included in the meta-analysis. These

selection criteria should include, but are not limited to: risk measure, endpoint consistency, dose or exposure metric consistency, exposure- or dose-response modeling, quality of data, quality of analyses, and data availability.

Although different criteria for selection of individual studies should be tailored to the problem at hand, the criteria should follow some pre-established guidelines to be valid. The criteria should be developed *a priori*, before a meta-analysis is started, to avoid selection bias. Different aspects to consider in study selection for quantitative meta-analysis are as follows:

1. Health endpoint. The studies selected for a meta-analysis should, ideally, be based on the same health endpoint. For example, if the health endpoint is lung cancer mortality, then all the studies selected for meta-analysis should be based on lung cancer mortality. There are instances in which epidemiology studies with different endpoints can be combined in a meta-analysis. For example, if necessary, a meta-analysis can combine results from epidemiology studies that report lung cancer mortality and lung cancer incidence because lung cancer incidence is reasonably predictive of lung cancer mortality. Lung cancer and respiratory cancer are another example of different endpoints that could be combined in a meta-analysis if most of the respiratory cancers are lung cancers and the mortality/incidence of both endpoints are similar.

2. Study design. The design of the study should be such that the health endpoint investigated was selected *a priori* and not an endpoint that came out as a result of exploratory analyses.

3. Study quality. The epidemiology studies selected for the meta-analysis should have comparable and reliable exposure estimates. Preferably, the exposure estimates will include quantitative rather than qualitative estimates. In addition, exposure estimates of selected epidemiology studies should meet quality criteria that make them credible. Epidemiology studies selected for meta-analyses should also be studies with a high degree of health endpoint ascertainment. In other words, there should be a high percentage of workers with the health endpoint of interest ascertained. The magnitude of the risk or statistical the significance of the findings is not an indicator of the quality of the study.

4. Data reliability. The source of the data should be reliable. Epidemiology data that have not been peer-reviewed and gone through some scientific scrutiny should be considered with caution or excluded from the meta-analysis.

5. Data availability. Depending on the reliability of the results and the extent of information reported in the open literature for the epidemiology studies selected, there may be a need for more data to estimate model parameters. The more accessible the individual data, the more valuable the study is because a meta-analysis using individual epidemiology data has much more modeling flexibility and potential control of study heterogeneity and other statistical issues.

6. Dose measure. The dose metric should, preferably, be the same for all the epidemiology studies selected for a meta-analysis. For example, if cumulative exposure is used as the

dose metric, then all the studies included in the meta-analysis have to use cumulative exposure as their dose metric. If one epidemiology study uses lagged cumulative exposure, for example, then that study cannot be combined in a quantitative meta-analysis with studies that used un-lagged cumulative exposures as the dose metric. Meta-analyses of risk measures where the dose metrics of the individual studies are different can be performed but with careful consideration to the potential heterogeneity of the individual risk estimates.

7. Risk measure. Risk in epidemiology studies is reported several different ways. The studies selected for a meta-analysis should all report the same measure of risk that the meta-analysis is intended to report. For example, if a meta-analysis to estimate the odds ratio is to be performed, then all the selected epidemiology studies should report the odds ratio for the health endpoint of interest or provide enough information to calculate the odds ratio. Epidemiology studies that qualify for inclusion in a meta-analysis are to be included whether their findings are positive or negative and regardless of the magnitude of the risk estimates.

8. Reproducibility of results. Epidemiology studies selected for inclusion in a meta-analysis should include enough information to corroborate or reproduce the results used for the meta-analysis. Studies that only include summary data without enough data to support the reported results should be seriously considered for exclusion from the meta-analysis.

9. Methodology. Studies selected for a meta-analysis should use the same, or at least similar, methodology to derive the individual risk estimates. There are several modeling issues that can affect the potential summarization of different results into one meta-analysis. The studies could adjust for different covariates using different types of adjustment (e.g., parametric or nonparametric). The individual risk estimates may have been derived using different statistical techniques (e.g., Poisson regression modeling, Cox proportional hazards modeling). The models fit to the epidemiology data of the individual studies may incorporate different assumptions (e.g., multiplicative background rates, additive background rates). Results derived from different epidemiology studies with different methodologies can be combined in a meta-analysis as long as the combination takes into account those differences. For example, the URF derived from an epidemiology study where the model was a polynomial in dose can be combined with a URF derived from an epidemiology study where the model was linear in dose. This can be done because the URFs already incorporate all the assumptions made in the derivation of these values. In contrast, for example, the URF derived using the U.S. population background hazard rates should not be combined with a URF derived using China population background hazard rates (unless the U.S. and China population background rates prove to be sufficiently similar).

The epidemiology study selection criteria for inclusion in a meta-analysis listed above is for guidance purposes. TCEQ staff may supplement or adjust these selection criteria guidelines according to their needs. There are circumstances in which a particular study report does not satisfy the selection criteria, but the authors can be contacted and are willing to share information that is not necessarily available in the open literature, making the study usable for the meta-analysis.

7.11.3 Meta-Analyses of Risk Measures

If estimates of risk (e.g., URFs) are the only available data, then a meta-analysis that combines risk estimates into a single risk estimate is the only possible option of estimating a single summary risk estimate. Such meta-analyses estimate the summary risk as a weighted average of the individual risk estimates. The weights of the individual study estimates are usually the inverse of the variance of the risk estimates (e.g., the inverse of the variance of the URF). The standard errors of the risk estimates are frequently reported in published studies or they can be back-calculated from reported confidence intervals, upper bounds or lower bounds.

Meta-analyses of estimates of risk have the advantage that they can accommodate results from individual epidemiology studies with differences in several characteristics of the risk characterization process. Risks based on different dose metrics, different dose-response models, etc. can be combined provided that the final risk estimates are based on the same risk metric (e.g., URFs in the same units), same health endpoint, same risk estimation methodology (e.g., the method of incorporating ADAFs, using life-table analyses, and using the same method of low-dose extrapolation), etc. The flexibility offered by meta-analyses based on individual study risk estimates is also a potential weakness in that different studies infrequently present estimates of risk using the same methodology, the same target population at risk and the same risk metric.

The standard error of the meta-analysis summary risk estimate can be similarly calculated from the standard errors of the risk estimates from the individual studies.

Other weighting factors that reflect the precision of the estimates can also be considered as alternatives to, or to supplement, the standard errors of the estimates of the individual studies. For example, if the standard errors are not available, the numbers of person-years at risk or the numbers of workers in the study are some alternative weighting factors to consider.

7.11.4 Meta-Analyses of Slope Estimates

A meta-analysis that combines slope (β) estimates, as opposed to final risk estimates (e.g., URFs), of individual epidemiology studies is more reliable than a meta-analysis that combines final risk estimates. A meta-analysis combining slope (β) estimates requires that the slopes of the individual studies be available and that the units of the slope be identical. That is, for example, if the slope (β) is in terms of risk increase per unit cumulative exposure, then all the slopes (β values) have to be in terms of risk increase per unit cumulative exposure and the cumulative exposure has to have been calculated in a similar way. Cumulative exposures could have been calculated un-weighted, weighted, lagged, etc., but must be calculated the same way in all studies. Such a meta-analysis estimates the summary slope (β) as a weighted average of the individual slope estimates. The weights of the individual study slope estimates are usually the inverse of the variance of the estimates. The standard errors of the slope (β) estimates are frequently reported in published studies or they can be back-calculated from reported confidence intervals, upper bounds or lower bounds.

Meta-analyses based on slopes (β values) of individual studies have the advantage that the estimates of risk can be calculated from the slope (β) determined from the meta-analysis, thereby avoiding any potential heterogeneity in the methods used to calculate excess risks from estimated slopes. There is, however, the disadvantage that individual studies based on different dose

metrics or different models cannot be combined to calculate a single slope. There is always the possibility, if the slopes are available, of estimating risks for each individual study using a standard methodology and then using a meta-analysis to combine individual risks as described above.

The standard error of the meta-analysis summary slope estimate can be similarly calculated from the standard errors of the slope (β) estimates from the individual studies.

Other weighting factors that reflect the precision of the estimates can also be considered as alternatives to, or to supplement, the standard errors of the estimates of the individual studies. For example, if the standard errors are not available, the numbers of person-years at risk or the numbers of workers in the study are some alternative weighting factors to consider.

7.11.5 Meta-Analyses of Individual Data

A meta-analysis based on all the individual data from the individual epidemiology studies is most desirable. This meta-analysis can control for all possible covariates and sources of heterogeneity. Alternatively, a meta-analysis that is based on summary (e.g., grouped) characterizations of exposures, observed number of events, expected number of events, standardized mortality or incidence ratios, etc., from individual studies can be used to perform a meta-analyses that can control for some sources of heterogeneity.

A meta-analysis based on the combined individual data or summary characterizations of the individual study data can be modeled. That is, a dose-response model can be fit to all the individual data combined (or all of the summary characterizations combined). This modeling can be done using a consistent methodology and adjusting for potential covariate effects like study, plant, co-exposures, sub-cohorts, etc. In such cases, the results of the meta-analysis include estimates of the standard errors of the model parameters and the risk estimates, and the individual studies are intrinsically weighted by the number of person-years, number of cases, etc. (see e.g., Valdez-Flores et al. 2010).

7.11.6 Heterogeneity and Uncertainty Analyses

The presentation of a meta-analysis should include the presentation and discussion of the individual study estimates. In addition, an uncertainty analysis for the sensitivity of the summary estimate to including and excluding each individual study result should be performed.

Implicit in a meta-analysis is the assumption that the individual study results are homogeneous with respect to the effect being estimated. Significant departures of individual study results from the expected results should be noted and, when possible, accounted for by discussing differences in study designs, methods of analysis, etc. The meta-analysis should include a detailed evaluation of the homogeneity of the individual study results, and if there is any heterogeneity detected, a discussion of the reasons for such heterogeneity and a justification for inclusion of the individual study in the meta-analysis.

7.12 Reality Checks

Reality checks on the predictions of risk based on the estimated toxicity factors (e.g., URF) are frequently worthwhile. They can be used to at least partially evaluate the reasonableness of dose-

response modeling assumptions and resulting estimates and bounds. Upper bounds can substantially overestimate their targets, and lower bounds can substantially underestimate their targets. Calculating both upper and lower bounds on the same target provides some measure of the uncertainty involved in the bounding methods.

Reality checks on bounds are also useful. For example, if the number of specified responses expected in a study cohort using a particular URF value is statistically significantly greater than the actual observed number of specified responses in that cohort, then the URF value is unrealistically high (i.e., results in unrealistically high risk estimates). In addition to making reality checks in terms of the study population, reality checks can be based on another population (e.g., another study or data source other than the epidemiology study used for dose-response modeling). For example, the estimated toxicity factor from the epidemiology study used for dose-response modeling can be used to predict the number of responses in the other population and the reasonableness of this prediction evaluated.

The utility of the toxicity factor in setting reasonable and/or meaningful risk management goals may also be evaluated using reality checks of the feasibility of attaining health-protective environmental (and other media) levels which would result from the toxicity factor. One such example is to compare the $LEC_{0.00001}$ (i.e., the lower bound on the concentration of a chemical corresponding to an increase of 1/100,000 in the excess risk) to the observed background environmental levels of that chemical or the background endogenous levels of the chemical in the human body. If this LEC is at or below these background levels (e.g., ambient air), this LEC may be an unattainable and unreasonable risk management goal. The corresponding toxicity factors represent an unrealistic characterization of environmental risk (at least from a regulatory perspective). The possibilities of such unreasonableness should be evaluated based on all relevant information (e.g., MOA, background rates of response).

Another form of a reality check is to estimate the number of response mortalities in the entire U.S. (or other specified population) in a year (e.g., 2010) that would be eliminated if the chemical's exposure via ambient air were reduced from a specified value to zero. If that number is less than 1 or an extremely small fraction of the background response rate, then any such reduction in the chemical's exposure may not be a meaningful risk management goal in terms of significantly reducing risk relative to total risk. This can happen simply because the chemical may not be associated with a significant environmental risk due to the presence of other substantial risk factors for the response in the population (e.g., other exposures, lifestyle choice factors).

7.13 Uncertainty Analysis

Risk characterizations based on epidemiology data contain an inherent degree of uncertainty and variability. Variability refers to differences that cannot be controlled by statistical modeling. The susceptibility to a specific agent of different persons, differences in lifestyle habits (e.g., smoking, drinking), and the differences in dose received by different individuals are examples of variability. Although all relevant variability may not be able to be controlled, it should be characterized. Uncertainty, on the other hand, refers to gaps in knowledge. The form of the dose-response model, the dose metric, the estimates of dose, job exposure histories, job classifications, and identification of causes of death are examples of uncertainty. Although uncertainties can be

reduced with more research, a risk characterization based on the available epidemiology data should recognize and characterize the uncertainties.

USEPA (2005a) and the National Research Council (BEIR V 1990 and BEIR VI 1999) emphasize that uncertainty analysis is an essential part of a risk characterization based on epidemiology data. The exclusion of an uncertainty analysis from a risk assessment prevents decision makers from taking well-informed actions in setting health-protective standards for chemicals that cause adverse health effects.

The uncertainty may refer to the model, the model parameters, the endpoint selected, the modeling methodology, the exposure estimation, etc. and TCEQ (2008; Grant et al. 2009), for example, include an extensive uncertainty analysis characterizing the impact of several alternative risk assessments. In their analysis, TCEQ (2008; Grant et al. 2009), considered the uncertainty of their risk characterization with respect to:

1. Population at risk: the effect of having only adult males in the epidemiology cohort and characterizing risk for the general population that includes females and young individuals.

2. Exposure estimation: The effect of occupational exposure estimation error when compared with actual measurements obtained for validation of the exposure estimates.

3. Statistical Methodology: the effect of using Cox proportional hazards modeling as opposed to Poisson regression modeling.

4. Dose-response modeling: The effects of including/excluding exposure peaks from the model.

5. Dose-response modeling: The effect of using individual exposure estimates instead of mean-scored deciles.

6. Dose-response modeling: The effect of including all the person-years or workers in the cohort as opposed to excluding the person-years or individuals with the highest doses.

7. Endpoint selection: The effect of the endpoint being mortality as opposed to incidence of the health effect.

In addition, TCEQ (2008; Grant et al. 2009), presented the uncertainty in the model's parameter estimates by presenting 95% confidence intervals in their risk characterization of 1,3-butadiene exposures.

The extent of the uncertainty analysis depends on and should be tailored to, the particular data and model(s) available for use. For example, if the risk characterization is based on a group of different epidemiology studies, the uncertainty about the inclusion/exclusion of some of the data from the risk characterization should be discussed and evaluated. The results of the individual epidemiology studies in addition to the results of a meta-analysis can be used to characterize the distribution of the uncertainty related to the selection of epidemiology studies.

References for Chapter 7

Agalliu, I., D. Kriebel, M. M. Quinn, D. H. Wegman, E. A. Eisen. (2005). Prostate Cancer Incidence in Relation to Time Windows of Exposure to Metalworking Fluids in the Auto Industry. Epidemiology 16: 664-671.

Allison, P.D. *Survival Analysis Using SAS: A Practical Guide, Second Edition*. SAS Institute, 2010.

Altman, D. G., B. Lausen, W. Sauerbrei, and M. Schumacher. (1994). Dangers of Using "Optimal" Cutpoints in the Evaluation of Prognostic Factors. Journal of the National Cancer Institute 86: 829-835.

BEIR IV. *Health Risks of Radon and Other Internally Deposited Alpha-Emitters.* Committee on the Biological Effects of Ionizing Radiations. Board on Radiation Effects Research Commission of Life Sciences. National Research Council. National Academy Press, Washington, DC, 1988.

BEIR V. *Health Effects of Exposures to Low Levels of Ionizing Radiation.* Committee on the Biological Effects of Ionizing Radiations. Board on Radiation Effects Research Commission of Life Sciences. National Research Council. National Academy Press, Washington, DC, 1990.

BEIR VI. *Health Effects of Exposure to Radon.* Committee on Health Risks of Exposure to Radon (BEIR VI). Board on Radiation Effects Research. Commission of Life Sciences. National Research Council. National Academy Press, Washington, DC, 1999.

Blair, A., J. Burg, J. Foran, H. Gibb, S. Greenland, R. Morris, G. Raabe, D. Savitz, J. Teta, D. Wartenberg, O. Wong, and R. Zimmerman, 1995. Guidelines for Application of Meta-Analysis in Environmental Epidemiology. ILSI Risk Science Institute. Regulatory Toxicology and Pharmacology 22:189-197.

Bonita, R., R. Beaglehole, and T. Kjellström. (2006). Basic Epidemiology, 2nd Edition. World Health Organization, Geneva, Switzerland.

Boyes, W. K., M. Evans, C. Eklund, P. Janssen, and J. E. Simmons. (2005). Duration Adjustment of Acute Exposure Guideline Values for Trichloroethylene Using a Physiologically-Based Pharmacokinetic Model. Risk Analysis 25: 677-686.

Breslow, N. E. and N. E. Day. *Statistical Methods in Cancer Research: Volume I – The Analysis of Case-Control Studies.* International Agency for Research on Cancer, Lyon, France, 1980.

Breslow, N. E. and N. E. Day. *Statistical Methods in Cancer Research: Volume II – The Design and Analysis of Cohort Studies.* International Agency for Research on Cancer, Lyon, France, 1987.

Buchanan, D., B. G. Miller, and C. A. Soutar. (2003) Quantitative Relations Between Exposure to Respirable Quartz and Risk of Silicosis. Occupational and Environmental Medicine 60: 159-164.

Checkoway, H., N. E. Pearce, and D. J. Crawford-Brown. *Research Methods in Occupational Epidemiology.* Oxford University Press, Inc., New York, NY, 1989.

Checkoway, H. and C. H. Rice. (1992). Time-Weighted Averages, Peaks, and Other Indices of Exposure in Occupational Epidemiology. American Journal of Industrial Medicine 21: 25-33.

Cheng, H., N. Sathiakumar, J. Graff, R. Matthews, and E. Delzell, 2007. 1,3-Butadiene and leukemia among synthetic rubber industry workers: Exposure–response relationships. Chemico-Biological Interactions 166: 15-24.

Collins, J.J., Ireland, B., Buckley, C.F., and Shepperly, D. (2003). Lymphohaematopoeitic Cancer Mortality Among Workers with Benzene Exposure. Occupational and Environmental Medicine 60: 676-679.

Copes R., D. Thomas, M. R. Becklake. (1985). Temporal Patterns of Exposure and Nonmalignant Pulmonary Abnormality in Quebec Chrysotile Workers. Arch Environ Health 40: 80-87.

Cox, D. R., (1972). Regression Models and Life Tables (with discussion). Journal of the Royal Statistical Society B34: 187-220.

Cox, L. A., Jr., M. G. Bird, and L. Griffis (1996). Isoprene Cancer Risk and the Time Pattern of Dose Administration. Toxicology 113(: 263-272.

Crump, C., K. Crump, E. Hack, R. Luippold, K. Mundt, E. Liebig, J. Panko, D. Paustenbach, and D.Proctor. (2003). Dose-Response and Risk Assessment of Airborne Hexavalent Chromium and Lung Cancer Mortality. Risk Analysis 23: 1147-1163.

Crump, K. S., D. G. Hoel, C. H. Langley, and R. Peto. (1976). Fundamental Carcinogenic Processes and Their Implications to Low Dose Risk Assessment. Cancer Research 36: 2973-2979.

Crump, K. S. and B. C. Allen. (1985). Methods for Quantitative Risk Assessment Using Occupational Studies. The American Statistician 39: 442-450.

Crump K. S. (1994). Risk of Benzene-Induced Leukemia: A Sensitivity Analysis of the Pliofilm Cohort with Additional Follow-up and New Exposure Estimates. Journal Toxicology and Environmental Health 42: 219-242.

Crump, K. (2005). The Effect of Random Error in Exposure Measurement Upon the Shape of the Exposure Response. Dose-Response (formerly Nonlinearity in Biology, Toxicology, and Medicine) 3: 456-464.

Enterline, P. E., R. Day, and G. M. Marsh. (1995). Cancers Related to Exposure to Arsenic at a Copper Smelter. Occupational and Environmental Medicine, 52: 28-32

Evans, M. V., W. K. Boyes, J. E. Simmons, D. K. Litton, and M. R. Easterling. (2002). A Comparison of Haber's Rule at Different Ages Using a Physiologically Based Pharmacokinetic (PBPK) Model for Chloroform in Rats. Toxicology 176: 11-23.

Federal Focus (1995). Principles for Evaluating Epidemiologic Data in Regulatory Risk Assessment. Developed by an Expert Panel at a Conference in London, England, October 1995. Federal Focus, Inc., 11 Dupont Circle, Suite 700, Washington, DC 20086.

Fedorov, V., F. Mannino, and R. Zhan. (2009). Consequences of Dichotomization. Pharm Stat 8: 50-61.

Figueiras, A. and C. Cadarso-Suárez. (2001). Application of Nonparametric Models for Calculating Odds Ratios and Their Confidence Intervals for Continuous Exposures. American Journal of Epidemiology 154: 264-275.

Gail, M. (1975). Measuring the Benefit of Reduced Exposure to Environmental Carcinogens. Journal of Chronic Diseases 28: 135-147.

Ginevan, M. E. and D. K. Watkins (2010). Lagarithmic Dose Transformation in Epidemiologic Dose-Response Analysis: Use with Caution. Regulatory Toxicology and Pharmacology 58: 336-340.

Ginsberg, G. L. (2003). Assessing Cancer Risks from Short-Term Exposures in Children. Risk Analysis 23: 19-34.

Goddard, M. J., D. J. Murdoch, and D. Krewski. (1995). Temporal Aspects of Risk Characterization. Inhal Toxicol 7: 1005–1018.

Goldbohm, R. A., Tielemans, E. L. J. P., D. Heederik, C. M. Rubingh, S. Dekkers, M. I. Willems, and E. D. Kroese. (2006). Risk Estimation for Carcinogens Based on Epidemiological Data: A Structured Approach, Illustrated by an Example on Chromium. Regulatory Toxicology and Pharmacology 44: 294-310.

Graham, J. P. (Editor). 1995. The Role of Epidemiology in Regulatory Risk Assessment. Proceedings of the Conference on the Proper Role of Epidemiology in Risk Analysis, Boston, MA, U.S.A. 13-14 October 1994. Elsevier Science B.V., New York.

Grant, R. L., J. Haney, A. L. Curry, and M. Honeycutt. (2009). Development of a Unit Risk Factor for 1,3-Butadiene Based on an Updated Carcinogenic Toxicity Assessment. Risk Analysis 29: 1726-1742.

Harris, J. E. (1983). Response: Diesel Emissions and Lung Cancer Revisited. Risk Analysis 3: 139-146.

Hays, S. M., R. A. Becker, H. W. Leung, L. L. Aylward, and D. W. Pyatt. (2007). Biomonitoring Equivalents: A screening approach for interpreting biomonitoring results from a public health perspective. Regulatory Toxicology and Pharmacology, 47: 96-109.

Hays, S. M., L. L. Aylward, and J. S. LaKind. (2008). Introduction to the Biomonitoring Equivalents Pilot Project: Development of Guidelines for the Derivation and Communication of Biomonitoring Equivalents. Regulatory Toxicology and Pharmacology, Vol. 51, Number 3, Supplement 1: S1-S2. Foreword and Supplement Articles: S3-S77.

Hertz-Picciotto, I. (1995). Epidemiology and Quantitative Risk Assessment: a Bridge from Science to Policy. American Journal of Public Health 85: 484–491.

Hewett, P. (2001). Misinterpretation and Misuse of Exposure Limits. Applied Occupational and Environmental Hygiene 16: 251-256.

Hill, B. A. (1965). The Environment and Disease: Association or Causation?. Proceedings of the Royal Society of Medicine 58: 295–300.

Höfler, M. (2005). The Bradford Hill considerations on causality: a counterfactual perspective?. Emerging themes in epidemiology 2: 11.

Hollander, N., W. Sauerbrei, and M. Schumacher. (2004). Confidence Intervals for the Effect of a Prognostic Factor after Selection of an ‘Optimal’ Cutpoint. Stat Med 23: 1701-1713.

Howick J., P. Glasziou, J. K. Aronson. (2009). The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute?. Journal of the Royal Society of Medicine 102: 186–94.

Kim, E., J. C. Little, and N. Chiu. (2004). Estimating Exposure to Chemical Contaminants in Drinking Water. Environ Sci Technol 38: 1799-1806.

Kirman, C. R., L. M. Sweeney, M. J. Teta, R. L. Sielken, C. Valdez-Flores, R. J. Albertini, and M. L. Gargas (2004). Addressing Nonlinearity in the Exposure-Response Relationship for a Genotoxic Carcinogen: Cancer Potency Estimates for Ethylene Oxide. Risk Analysis 24: 1165-1183.

Lemeshow, S. and D. W. Hosmer. (1982). A Review of Goodness of Fit Tests for use in the Development of Logistic Regression Models. American Journal of Epidemiology, 115: 91-106.

Little, J. C. and N. Chiu. (1998). Exposure to Contaminant in Drinking Water: Estimating Uptake Through the Skin and by Inhalation. ILSI, Risk Science Institute Working Group. Olin ed. page 93.

Manos M. M., W. A. Leyden, R. C. Murphy, N. A. Terrault, and B. P. Bell. (2008). Limitations of Conventionally Derived Chronic Liver Disease Mortality Rates: Results of a Comprehensive Assessment. Hepatology 47: 1150-7.

McKone, T. (1987). Human Exposure to Volatile Organic Compounds in Household Tap Water: The Indoor Inhalation Pathway. Environ Sci Technol 21: 1194-1201.

Miller, F. J., P. M. Schlosser, and D. B. Jansen. (2000). Haber's Rule: A Special Case in a Family of Curves Relating Concentration and Duration of Exposure to a Fixed Level of Response for a Given Endpoint. Toxicology 149: 21-34.

Moseman, R. F. and E. O. Oswald. (1980). Development of Analytical Methodology for Assessment of Human Exposure to Pesticides, Chapter 13 in Pesticide Analytical Methodology. ACS Symposium Series, Vol. 136, American Chemical Society.

Murdoch, D. J., D. Krewski, and J. Wargo. (1992). Cancer Risk Assessment with Intermittent Exposure. Risk Analysis 12: 569-577.

Ott, W. R., A. C. Steinemann, and L. A. Wallace. Exposure Analysis. CRC Press, 2007.

Paustenbach, D. J. The Risk Assessment of Environmental Hazards: A Textbook of Case Studies. John Wiley & Sons. New York, 1989.

Paustenbach D., D. Galbraith. (2006). Biomonitoring and Biomarkers: Exposure Assessment Will Never Be the Same. Environ Health Perspect 114: 1143-1149.

Pearce, N. (2000). The ecological fallacy strikes back. J Epidemiol Community Health 2000;54:326-327.

Phillips, C. V. and K. J. Goodman. (2006). Causal Criteria and Counterfactuals; Nothing More (or Less) Than Scientific Common Sense?. Emerging Themes in Epidemiology 3: 5.

Pottenger, L. H., R. A. Becker, E. J. Moran, and J. A. Swenberg. (2011). Workshop report: Identifying key issues underpinning the selection of linear or non-linear dose-response extrapolation for human health risk assessment of systemic toxicants. Regulatory toxicology and pharmacology 59: 503-510.

Richardson, D. B., and D. Loomis. (2004). The Impact of Exposure Categorisation for Grouped Analyses of Cohort Data. Occupational and Environmental Medicine 61: 930–935.

Rothman, K. J. *Modern Epidemiology*. Little, Brown and Company, Boston, MA, 1986.

Royston, P., D. G. Altman, and W. Sauerbrei. (2006). Dichotomizing Continuous Predictors in Multiple Regression: A Bad Idea. Stat Med, 25: 127-141.

Schulgen, G., B. Lausen, J. Olsen, and M. Schumacher. (1994). Outcome-Oriented Cutpoints in Quantitative Exposure. Am J Epi, 120: 172-184.

Shaw, D. G., M. D. Huffman, and M. G. Haviland. (1987). Grouping Continuous Data in Discrete Intervals: Information Loss and Recovery. Journal of Educational Measurement, 24: 167-173.

Shore, R. E., V. Iyer, B. Altshuler, and B. S. Pasternack. (1992). Use of Human Data in Quantitative Risk Assessment of Carcinogens: Impact on Epidemiologic Practice and the Regulatory Process. Regulatory Toxicology and Pharmacology 15: 180-221.

Shusterman, D., E. Matovinovic, and A. Salmon. (2006). Does Haber's Law Apply to Human Sensory Irritation? Inhalation Toxicology 18: 457-471.

Sielken, R. L., Jr., and C. Valdez-Flores (2009a). Calculating excess risk with age-dependent adjustment factors and cumulative doses: Ethylene oxide case study. Regulatory Toxicology and Pharmacology 55: 76-81.

Sielken, R. L., Jr., and C. Valdez-Flores (2009b). Life-table calculations of excess risk for incidence versus mortality: Ethylene oxide case study. Regulatory Toxicology and Pharmacology 55: 82-89.

Sielken, R. L., Jr. and C. Valdez-Flores (2011). Butadiene Cancer Exposure-Response Modeling: Based on Workers in the Styrene-Butadiene-Rubber Industry: Total Leukemia, Acute Myelogenous Leukemia, Chronic Lymphocytic Leukemia, and Chronic Myelogenous Leukemia. accepted by Regulatory Toxicology and Pharmacology.

Stayner, L., R. Smith, A. J. Bailer, E. G. Luebeck, and S. H. Moolgavkar (1995). Modeling Epidemiologic Studies of Occupational Cohorts for the Quantitative Assessment of Carcinogenic Hazards.American Journal of Industrial Medicine 27: 155-170.

Steenland, K., J. Deddens, and L. Stayner. (1998). Diesel Exhaust and Lung Cancer in the Trucking Industry: Exposure–Response Analyses and Risk Assessment. American Journal of Industrial Medicine 34:220–228.

Steenland, K., J. Deddens, and L. Piacitelli. (2001). Risk Assessment for 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) based on an Epidemiologic Study. American Journal of Epidemiology 154: 451-458.

Steenland, K., E. Whelan, J. Deddens, L. Stayner, and E. Ward (2003). Ethylene Oxide and Breast Cancer Incidence in a Cohort Study of 7576 Women (Unites States). Cancer Causes and Control 14: 531-539.

Steenland, K. and J. Deddens (2004). A Practical Guide to Dose-Response Analyses and Risk Assessment in Occupational Epidemiology. Epidemiology 15: 63-70.

Steenland, K., L. Stayner, and J. Deddens (2004). Mortality Analyses in a Cohort of 18 235 Ethylene Oxide Workers: Follow up Extended from 1987 to 1998. Occupational and Environmental Medicine 61: 2-7.

Stewart, P. (1999). Challenges to Retrospective Exposure Assessment. Scand J Work Environ Health 25: 505-510.

Texas Commission on Environmental Quality (TCEQ). (2007). Benzene CAS Registry Number 71-43-2, Final Development Support Document, Texas Commission on Environmental Quality (TCEQ), Austin, TX. http://www.tceq.texas.gov/toxicology/dsd/final.html

Texas Commission on Environmental Quality (TCEQ). (2008). 1,3-Butadiene CAS Registry Number 106-99-0, Final Development Support Document, Texas Commission on Environmental Quality (TCEQ), Austin, TX. http://www.tceq.texas.gov/toxicology/dsd/final.html

Texas Commission on Environmental Quality (TCEQ). (2009) Silica, Crystalline Forms CAS Registry Numbers: 14808-60-7 (quartz), 14464-46-1 (cristobalite), 1317-95-9 (tripoli), 15468-32-3 (tridymite). Final Development Support Document, Texas Commission on Environmental Quality (TCEQ), Austin, TX. http://www.tceq.texas.gov/toxicology/dsd/final.html

Texas Commission on Environmental Quality (TCEQ). (2010). Arsenic and Inorganic Arsenic Compounds CAS Registry Numbers: 7440-38-2 (Arsenic). Draft of the Development Support Document, Texas Commission on Environmental Quality (TCEQ), Austin, TX. (Follow-up publication to appear in 2011). http://www.tceq.texas.gov/toxicology/dsd/reviews.html

Texas Commission on Environmental Quality (TCEQ). (2011). Nickel and Inorganic Nickel Compounds CAS Registry Numbers: Nickel: 7440-02-0, Nickel Sulfate: 7786-81-4, Nickel Subsulfide: 12035-72-2, Nickel Oxide: 1313-99-1, Nickel Chloride: 7718-54-9. Final Development Support Document,Texas Commission on Environmental Quality (TCEQ), Austin, TX. (Follow-up publication to appear in 2011). http://www.tceq.texas.gov/toxicology/dsd/final.html

ten Berge, W. F., A. Zwart, L. M. Appelman. (1986) Concentration-Time Mortality Response Relationship of Irritant and Sytemically Acting Vapors and Gases. Journal of Hazardous Materials 13: 301-309

Törnqvist, M. and L. Ehrenberg (1992). Risk Assessment of Urban Air Pollution. Pharmacogenetics 2: 297-303.

Törnqvist, M. and L. Ehrenberg (1994). On cancer Risk Estimation of Urban Air Pollution. Environmental Health Perspectives 102(Suppl 4): 173-181.

USEPA(1986). Health Assessment Document for Nickel and Nickel Compounds. EPA/600/8-83/012FF, September 1986, Final Report. US Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC.

USEPA. (1995) “The Use of the Benchmark Dose Approach in Health Risk Assessment” Risk Assessment Forum, Washington DC. EPA/630/R-94/007.

USEPA. (1998) “Summary of the U.S. Workshop on the Relationship between Exposure Duration and Toxicity” Office of Research and Development, Washington DC. EPA/600/R-99/081.

USEPA. (2000a) “Benchmark Dose Technical Guidance Document” Risk Assessment Forum, Washington DC. EPA/630/R-00/001.

USEPA(2000b). Summary Report of the Workshop to Peer Review the Benchmark Dose Technical Guidance Document. Washington, D.C., December 7-8, 2000.

USEPA(2001). Health Assessment of 1,3-Butadiene. EPA/600/P-98/001B, January 2001, Consensus Review Draft. US Environmental Protection Agency, Office of Research and Development, Washington, DC.

USEPA(2004). Air Toxics Risk Assessment Reference Library. EPA-453-K-04-001A. Office of Air Quality Planning and Standards, Research Triangle Park, N.C.

USEPA(2005). Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001F. US Environmental Protection Agency, Risk Assessment Forum, Washington, D.C., March, 2005.

USEPA(2005). Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens. EPA/360/R-03/003F. US Environmental Protection Agency, Risk Assessment Forum, Washington, D.C., March, 2005.

USEPA(2006). Evaluation of the Carcinogenicity of Ethylene Oxide. EPA/635/R-06/003, August 2006, External Review Draft. US Environmental Protection Agency, Office of Research and Development, Washington, DC.

Valdez-Flores, C., R. L. Sielken, Jr., and M. J. Teta. (2010). Quantitative Cancer Risk Assessment based on NIOSH and UCC Epidemiological Data for Workers Exposed to Ethylene Oxide. Regulatory Toxicology and Pharmacology, 56: 312-320.

Valdez-Flores, C. and R. L. Sielken, Jr. (2010). Actuarial Computation of Excess Risks and Point of Departures for Rate Ratio and Hazard Rate Models Using Population-Based Age-Specific Background Rates and Competing Risks. Sielken & Associates Technical Report.

Wainer, H. (2006). Finding What is not there Through the Unfortunate Binning of Results: The Mendel Effect. Chance 19: 49-56.

Ward,A. C. (2009). The Role of Causal Criteria in Causal Inferences: Bradford Hill's "Aspects of Association". Epidemiological Perspectives and Innovations 6: 2.

Weller, E., N. Long, A. Smith, P. Williams, S. Ravi, J. Gill, R. Hennessey, W. Skornik, J. Brain, C. Kimmel, G. Kimmel and L. Ryan, L. (1999). Dose-Rate Effects of Ethylene Oxide Exposure on Developmental Toxicity. Toxicological Sciences 50: 259-270.

WHO Working Group. 2000. Evaluation and Use of Epidemiological Evidence for Environmental Health Risk Assessment: WHO Guidance Document. Environmental Health Perspectives, Vol. 108, No. 10 (Oct. 2000), 997- 1002.

Appendix A (Chapter 7): Linear Multiplicative Relative Risk Models

A.1 Overview of Poisson Regression Models

Poisson regression models assume that the number of observed events (e.g., number of cancer deaths) for a particular group of individuals or persons years follow a Poisson distribution. The Poisson distribution indicates that the probability of observing r events is given by the following function

$$p(R=r) = (\lambda n)^r \times e^{-\lambda n} / r!$$

where p(R=r) is the probability that r events are observed, r is the number of events occurring in the group, n is the number of individuals or person-years in the group, λ is the unknown rate of occurrence of events per individual or person-year at risk (i.e., λn is the number of events occurring in the group). The expected value (i.e., E[R]) of the Poisson distribution is given by λn.

Tables of summary data from epidemiological studies are often presented in the form of observed and expected number of cancer deaths for different groups. The groups can correspond to combinations of different dose intervals, different sexes, different plants, etc. Poisson regression assumes that the rate of cancer death remains constant within each group defined by a combination of the dose interval, sex, plant, etc. The rate of cancer deaths in a specific group is in terms of the number of cancer deaths per person-year at risk. The number of person-years at risk is the total number of years that different individuals contribute to each different group. For example, consider the following table defining groups according to sex, age and cumulative exposure:

Covariates		Cumulative Exposure (ppm-years)			
Sex	Age	0	0 to 10	10 to 100	100+
Male	< 40				
	40 to 60				
	60 +				
Female	< 40				
	40 to 60				
	60 +				

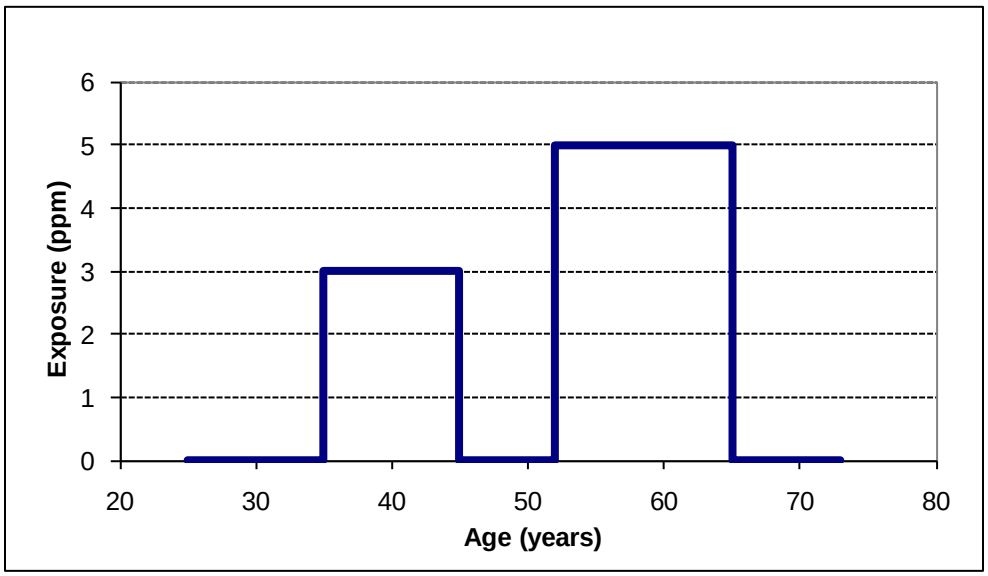
Now consider the following job exposure profile for one male worker who was followed up (i.e., was at risk because he was being observed) from age 25 through age 73 years. His job history indicates that:

He started to be followed up when he was 25

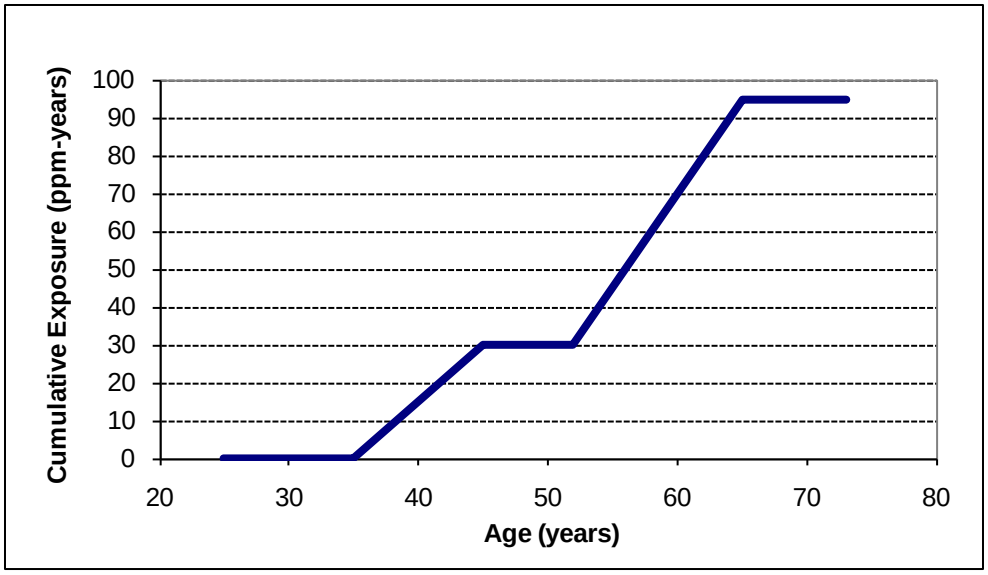
Before age 25 he was not exposed
From age 25 to 35 he was not exposed
From age 35 to 45 he was exposed to 3 ppm on his job
From age 45 to 52 he was not exposed
From age 52 to 65 he was exposed to 5 ppm on his job
From age 65 to 73 he was not exposed

His follow-up ended when he was 73 and he was alive at that time
He was followed up for a total of 48 years (48 person-years)

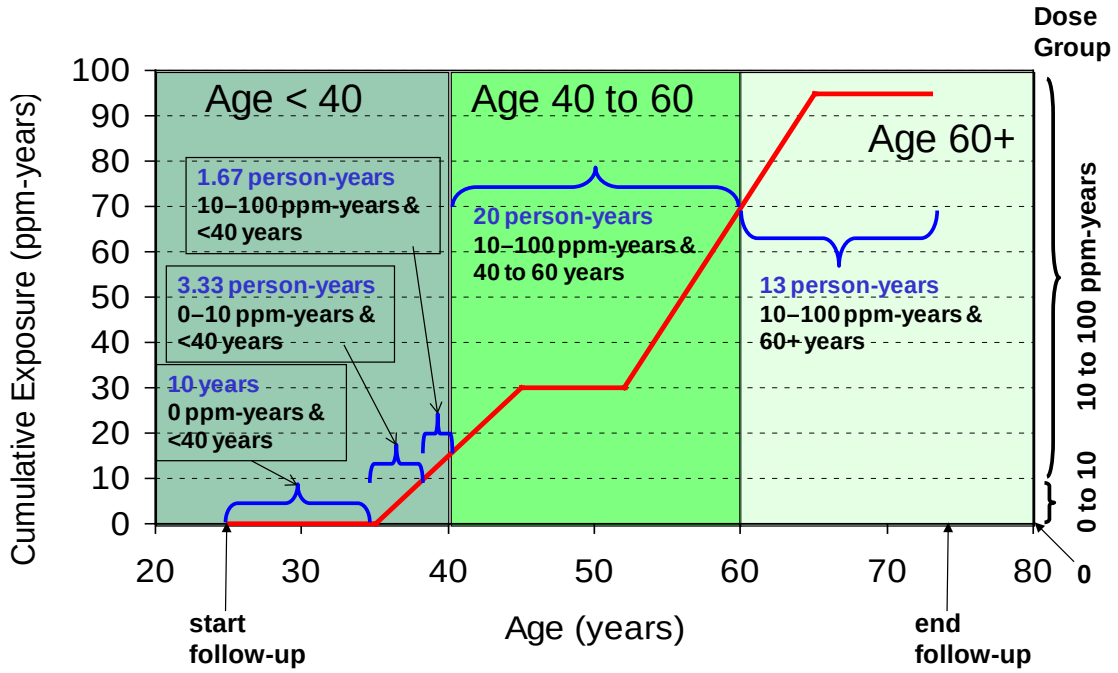
This worker's aged changed over the period he was observed (i.e., he belonged for different periods of time to different age groups in the table above). The cumulative exposure also changed over the period the worker was observed (i.e., his cumulative exposure was in different exposure intervals at different times in the table above). The following graph shows the exposure profile for this worker.



The corresponding cumulative exposure (ppm-years) is in the following graph.



It is clear that all the time this worker is at risk (i.e., being observed) he is male (i.e., all his person-years belong to the male group). It is also clear how many years of this worker's time at risk he was younger than 40, between 40 and 60 and over 60 years (i.e., 15, 20 and 13 years, respectively). However, it is not so clear the length of time this worker spent in each of the age groups and each of the cumulative exposure intervals. The following graph overlaps the age groups and cumulative exposure intervals for this particular worker.



The above graph shows length of time in years that this worker belonged to each combination of age group and exposure interval. During the first 15 years of followup (age 25 to age 40) he was exposed to 0 ppm-years for 10 year, 3.33 years he was exposed to a cumulative exposure less than 10 ppm-years but more than 0 ppm-years, and for 1.67 years his exposure ranged between

10 and 100 ppm-years. The next 20 years of his life were spent between the ages of 40 and 60 years and his cumulative exposure was between 10 and 100 ppm-years. He was older than 60 during his last 13 years of observation and his cumulative exposure was between 10 and 100 ppm-years. The numbers of person-years for this worker in each group of sex, age and cumulative exposure are summarized in the following table

Covariates		Cumulative Exposure (ppm-years)			
Sex	Age	0	0 to 10	10 to 100	100+
Male	< 40	10	3.33	1.67	
	40 to 60			20	
	60 +			13	
Female	< 40				
	40 to 60				
	60 +				

A table similar to the above table tallies the number of events in each group. For example, if the worker of the example had died with the response (e.g., cancer) then the worker would contribute with one event to the cell of his last person-year of follow-up.

After following a large number of workers in an epidemiological study, the table may look like

Covariates			Cumulative Exposure (ppm-years)			
Sex	Age		0 (e_0)	0 to 10 (e_1)	10 to 100 (e_2)	100+ (e_3)
Male (s_0)	< 40	(a_0)	0 / 1233.0	0 / 755.3	1 / 693.2	0 / 121.5
	40 to 60	(a_1)	1 / 1433.0	1 / 957.8	1 / 893.7	2 / 357.9
	60 +	(a_2)	1 / 893.1	2 / 1055.0	1 / 752.1	0 / 523.8
Female (s_1)	< 40	(a_0)	1 / 1793.0	0 / 867.2	1 / 323.6	
	40 to 60	(a_1)	0 / 899.0	0 / 739.5	1 / 651.8	0 / 225.3
	60 +	(a_2)	2 / 795.3	1 / 516.7	1 / 419.6	1 / 337.2

In the table above, the #1 of the entry #1 / #2 refers to the number of events (e.g., cancer deaths) and #2 refers to the number of person-years at risk. The labels s_0, s_1, a_0, a_1, a_2, e_0, e_1, e_2, and e_3 indicate different levels of the factors or covariates. A Poisson regression model for the above summary data would account for each of the factors (covariates). That is, the probability of observing the number of deaths in each cell follows a Poisson distribution given by

$$p(R=r_{ijk}) = (\lambda_{ijk} n_{ijk})^{r_{ijk}} \times e^{-\lambda_{ijk} n_{ijk}} / r_{ijk}!$$

where the subscript refers to the i-th sex (s_0 or s_1), j-the age group (a_0, a_1, or a_2) and k-th cumulative exposure interval (e_0, e_1, e_2, or e_3). The unknown hazard rate λ_{ijk} for the ijk-th cell in the table can be modeled using a multiplicative background model as follows,

$$\lambda_{ijk} = \lambda \times s_i \times a_j \times e_k$$

where λ is the overall background hazard rate for the cohort, s_i is the effect of sex, a_j is the effect of age and e_k is the effect of cumulative exposure. In this model the values of s_0, a_0 and e_0 are defined as 1 and λ and the values of the variables with other subscripts are unknown parameters.

The purpose of dose-response modeling is to determine the relationship between a dose (e.g., cumulative exposure ppm-years) and the observed response (e.g., cancer deaths). In other words, the effect e_k in the above equation is modeled as a function of dose rather than as a categorical effect. For example, the effect of cumulative exposure on the rate can be modeled as a linear function of the cumulative exposure as in the following

$$e_k = 1 + \beta \times d_k,$$

where d_k is the cumulative exposure of the k-th dose interval (the average dose for the person years in the interval is the most appropriate value) and β is an unknown parameter that needs to be estimated. After replacing the effect of cumulative exposure by the linear function, the multiplicative model is as follows:

$$\lambda_{ijk} = \lambda \times s_i \times a_j \times (1 + \beta \times d_k)$$

The parameters λ, s_1, a_1, a_2 and β can then be estimated using maximum likelihood. The likelihood function is given by

$$\text{Likelihood} = \prod_{ijk} p(R=r_{ijk})$$

$$= \prod_{ijk} \exp\{-\lambda_{ijk} n_{ijk}\} \times (\lambda_{ijk} n_{ijk})^{r}_{ijk} \; / \; r_{ijk}!$$

where the product is over all cells in the table and $\lambda_{ijk} = \lambda \times s_i \times a_j \times (1 + \beta \times d_k)$. The values of λ, s_1, a_1, a_2 and β that maximize the likelihood are the maximum likelihood estimates of those parameters. The product $\lambda \times s_i \times a_j$ in the Poisson regression model is the estimate of background hazard rate for the ijk-th cell in the table when the cumulative exposure is equal to zero. The relation $(1 + \beta \times d_k)$ in the model is the relative risk and describes the effect of cumulative exposure on the background hazard rate. In the estimation of risks for a different population the effect of the cumulative exposure is used but the background hazard rate ($\lambda \times s_i \times a_j$) is replaced by the sex- and age-dependent target population's background rate.

A.2 Summary Estimates of Standardized Mortality/Incidence Rates

Summary data presented in most published epidemiological studies are in the form of SMRs or SIRs. The SMRs are the ratio of the observed number of events in the cohort of the epidemiology study (e.g., number of cancer deaths) to the expected number of events in a reference population. The summary data can be used to fit a dose-response model using Poisson regression if the observed and expected numbers of deaths are given for different dose intervals. For example, the following table was reported by Enterline et al. 1995 for workers exposed to arsenic.

Cumulative Exposure (mg/m3-year)	Mean Cumulative Exposure (mg/m3-year)	Number of Deaths with Respiratory Cancer Observed in the Cohort	Number of Deaths with Respiratory Cancer Expected in the Reference Population*	SMR
[0, 0.75)	0.405	22	14.29	154.0
[0.75, 2)	1.305	30	17.10	175.5
[2, 4)	2.925	36	17.17	209.7
[4, 8)	5.708	36	17.00	211.7
[8, 20)	12.334	39	15.48	252.0
[20, 45)	28.336	20	7.04	284.0
45+	58.957	5	1.58	315.7

*White men in the State of Washington were used as the reference population because the plant of the study is located in that state and because all 2,802 workers in the study were men and most of them where white.

The data in the summary table can be fit using Poisson regression and the multiplicative background hazards model. The model for the linear dose-response model was specified in the previous section as

$$\lambda_{ijk} = \lambda \times s_i \times a_j \times (1 + \beta \times d_k)$$

where the s_i and a_j reflected the effects of sex and age, respectively. However, if only the effect of cumulative exposure is to be modeled (because that is the only information available in the Enterline et al. 1995 paper), the dose-response model reduces to

$$\lambda_k = \lambda \times (1 + \beta \times d_k).$$

If both sides of the equation are multiplied by the number of person-years in the k-th cell, then the expression is as follows,

$$\lambda_k \times n_k = \lambda \times n_k \times (1 + \beta \times d_k).$$

This is equivalent to

$$Observed_k = \text{Expected}_k \times (1 + \beta \times d_k)$$

Where

$Observed_k$ is the number of deaths in the k-th exposure interval predicted by the model,

Expected_k is the expected number of deaths in the reference population corresponding to the person-years in the k-th exposure interval.

The parameter β can then be estimated using maximum likelihood. The likelihood function is given by

$$\text{Likelihood} = \prod_k p(R=\text{Observed}_k)$$

$$= \prod_k \exp\{-\text{Expected}_k \times (1 + \beta \times d_k)\} \times$$

$$[\text{Expected}_k \times (1 + \beta \times d_k)]^{\text{Observed}_k} \; / \; \text{Observed}_k!$$

where the product is over all exposure intervals in the table and Observed_k is the actual number of respiratory cancer deaths observed in the k-th dose interval. (Note that Observed_k is the actual number while $Observed_k$ is the number predicted by the model). The value of β that maximizes the likelihood is the maximum likelihood estimate of that parameter.

The relation $(1 + \beta \times d_k)$ in the model describes the effect of cumulative exposure on the background rate. In the estimation of risks for a different population, the effect of the cumulative exposure is used along with the background hazard rate of that population.

A.3 Adjustments for Possible Differences Between the Population Background Cancer Rate and the Cohort's Cancer Rate in the Relative Risk Model

A multiplicative relative risk model that uses reference population background cancer rates to fit the cohort's observed cancer rates should adjust for possible discrepancies between the background cancer rates in the reference population and the background cancer rates in the cohort.

In the example given in Section A.2, the multiplicative background dose-response model relates the number of observed respiratory cancer deaths to the product of the number of respiratory cancers expected in a reference population and the effect of the dose. The underlying background respiratory cancer hazard rates of the workers in the cohort may be (and usually are) different than the underlying background respiratory hazard rates in the reference population. If no adjustment for this difference is made, the dose-dependent function (i.e., the term $1+\beta \times d_k$ in the model) is forced to explain not only the effect of the dose in the observed respiratory cancer mortalities but also any discrepancies between the study and reference population background rates. In other words, ignoring discrepancies between the cohort's background hazards rates and the reference population hazard rates may result in distorted (i.e., biased) dose-response relationships.

Crump and Allen (1985) discuss the relative risk model with a factor that accounts for the possibility of different background rates in an epidemiological cohort and its reference population. This factor may adjust for issues like the healthy worker effect, the difference between internally and externally derived background cancer rates, covariate effects not explicitly incorporated in the summary epidemiological data, etc. For example, the multiplicative background relative risk model with no adjustment for differences in background rates can be extended from

$$Observed_k = \text{Expected}_k \times (1 + \beta \times d_k)$$

to

$$Observed_k = \alpha \times Expected_k \times (1 + \beta \times d_k)$$

where the α term adjusts for any possible difference between the population's background cancer rates and the cohort's observed cancer rates in unexposed workers.

In the equations above the variables are:

$Observed_k$ =	number of lung cancer deaths for exposure group k predicted by the model;
$Expected_k$ =	expected number of background lung cancer deaths for exposure group k based on the reference population background cancer rates;
β =	multiplicative factor by which background risk increases with cumulative exposure;
d_k =	cumulative exposure for exposure group k;
α =	multiplicative factor that accounts for differences in cancer mortality background rates between the study cohort and the reference population.

A.4 Estimating the Slope Parameter, β, in the Relative Risk Model Adjusting for Differences in Background Rates

As discussed in Section A.1, Poisson regression is a standard modeling technique in epidemiological studies. Poisson regression relies on the assumption that the number of cancer deaths in a dose group follows a Poisson distribution with mean equal to the expected number of cancer deaths and uses the maximum likelihood estimation procedure for the estimation of the parameters α and β in the model.

The Poisson distribution that describes probabilistically the number of cancers observed in a group is given by:

$$p(x) = \lambda^x \times e^{-\lambda} / x!$$

where p(x) is the probability of observing x cancers, x is the number of cancer deaths actually observed, x! = x (x-1) (x-2) … 1, and λ is the number of cancers in the group predicted by the model. Thus, for dose group k, x_k=$Observed_k$ and λ_k= $Observed_k = \alpha \times Expected_k \times (1 + \beta \times d_k)$. That is, for each group k of person-years with average dose d_k, the observed number of cancer deaths in the dose interval ($Observed_k$) follows a Poisson distribution with parameter λ_k= $Observed_k = \alpha \times Expected_k \times (1 + \beta \times d_k)$ and the likelihood of observing $Observed_k$ cancer deaths is given by,

$$p(Observed_k) = \exp\{-\alpha \times Expected_k \times (1 + \beta \times d_k)\} \times [\alpha \times Expected_k \times (1 + \beta \times d_k)]^{Observed_k} / Observed_k!.$$

The likelihood (L) is given by the product of the likelihoods of observing the number of cancer deaths in each dose group. That is,

$$L = p(\text{Observed}_1) \times p(\text{Observed}_2) \times \ldots.$$

or, equivalently,

$$L = \exp\{-\alpha\times\text{Expected}_1\times (1+\beta\times d_1)\}\times[\ \alpha\times\text{Expected}_1\times(1+\beta\times d_1)]^{\text{Observed}_1}/\text{Observed}_1!\times$$

$$\exp\{-\ \alpha\times\text{Expected}_2\times (1+\beta\times d_2)\}\times[\ \alpha\times\text{Expected}_2\times(1+\beta\times d_2)]^{\text{Observed}_2}/\text{Observed}_2!\times$$

$$\ldots..$$

where exp{.} is the base of the natural logarithm (e) raised to the power in the braces and Observed_k is the number of cancer cases observed for the person-years with cumulative exposures equal to d_k. The likelihood equation can be written using mathematical notation as follows:

$$L = \prod \exp\{-\alpha\times\text{Expected}_k\times (1+\beta\times d_k)\}\times [\alpha\times\text{Expected}_k\times(1+\beta\times d_k)]^{\text{Observed}_k}/\text{Observed}_k!$$

where the symbol $\prod$ indicates that it is the product over all dose groups k=1,2,…

The maximum likelihood estimates of the unknown parameters α and β can then be obtained by selecting the values of α and β that maximize the value of L. Finding the values of α and β that maximize the value of the likelihood L cannot be determined using a close-form solution because there are two variables. However, any routine that can maximize nonlinear functions of more than one variable can be used to calculate the maximum likelihood estimates of α and β.

The parameters α and β that maximize the likelihood function given above also maximize the logarithm of the likelihood because the logarithm is a monotone function. The logarithm of the likelihood function (LL) for the model given above is

$$LL = \sum \{\ -\alpha\times\text{Expected}_k\times (1+\beta\times d_k) + \text{Observed}_k \times \ln[\alpha\times\text{Expected}_k\times(1+\beta\times d_k)] - \ln(\text{Observed}_k!)\ \}$$

where the symbol $\sum$ indicates that it is the sum over all dose groups k=1,2,… and ln(x) is the natural logarithm of x. The LL function can also be written as

$$LL = \sum\{-\alpha\times\text{Expected}_k\times (1+\beta\times d_k) + \text{Observed}_k\times\ln(\alpha) + \text{Observed}_k \times \ln(\text{Expected}_k) + \text{Observed}_k \times \ln(1+\beta\times d_k)] - \ln(\text{Observed}_k!)\ \}$$

Note that the terms $\text{Observed}_k\times\ln(\text{Expected}_k)$ and $\ln(\text{Observed}_k!)$ in the LL equation above do not depend on the values of α and β, and hence, the values of α and β that maximize the LL also maximize the following simplified LL function:

$$LL = \sum\{-\alpha\times\text{Expected}_k\times (1+\beta\times d_k) + \text{Observed}_k\times\ln(\alpha) + \text{Observed}_k \times \ln(1+\beta\times d_k)]\ \}$$

Finally, the maximum likelihood estimates of α and β can also be obtained by solving for α and β in the following system of equations:

$$\frac{\partial LL}{\partial \alpha} = \sum \{ -Expected_k \times (1+\beta \times d_k) + Observed_k/\alpha \} = 0$$

$$\frac{\partial LL}{\partial \beta} = \sum \{ -\alpha \times Expected_k \times d_k + (Observed_k \times d_k) / (1 + \beta \times d_k) \} = 0$$

where $\partial LL/\partial \alpha$ and $\partial LL/\partial \beta$ are the partial derivatives of the logarithm of the likelihood with respect to α and β, respectively.

A.5 Estimating the Asymptotic Variance for the Slope Parameter in the Relative Risk Model

The system of equations of the partial derivatives of the logarithm of the likelihood given in the previous section can be used to estimate the asymptotic variance of the maximum likelihood estimates of α and β. The variance-covariance matrix of the parameters α and β is approximated by

$$Cov(\alpha,\beta) = - \begin{pmatrix} \partial^2 LL/\partial \alpha^2 & \partial^2 LL/\partial \alpha \partial \beta \\ \partial^2 LL/\partial \alpha \partial \beta & \partial^2 LL/\partial \beta^2 \end{pmatrix}^{-1}$$

where $[.]^{-1}$ is the inverse of the matrix, $\partial^2 LL/\partial \alpha^2$ is the second partial derivative of the logarithm of the likelihood with respect to α, $\partial^2 LL/\partial \beta^2$ is the second partial derivative of the logarithm of the likelihood with respect to β, and $\partial^2 LL/\partial \alpha \partial \beta$ is the partial derivative of the logarithm of the likelihood with respect to α and β. The approximation of the covariance is then given by

$$Cov(\alpha,\beta) = - \begin{pmatrix} \partial^2 LL/\partial \beta^2 & -\partial^2 LL/\partial \alpha \partial \beta \\ -\partial^2 LL/\partial \alpha \partial \beta & \partial^2 LL/\partial \alpha^2 \end{pmatrix} / \text{Determinant}$$

where

$$Determinant = 1 / [\partial^2 LL/\partial \alpha^2 \times \partial^2 LL/\partial \beta^2 - (\partial^2 LL/\partial \alpha \partial \beta)^2]$$

The second-order derivatives used for the estimation of the variance-covariance matrix are:

$$\frac{\partial^2 LL}{\partial \alpha^2} = \sum -Observed_k/\alpha^2$$

$$\frac{\partial^2 LL}{\partial \beta^2} = \sum -(Observed_k \times d_k^2) / (1 + \beta \times d_k)^2$$

$$\frac{\partial^2 LL}{\partial \alpha \partial \beta} = \sum -Expected_k \times d_k$$

A better asymptotic variance calls for substituting the variance-covariance matrix of α and β by the expected value of the above matrix. That is, by replacing the observed number of cancer deaths in a dose group k ($Observed_k$) by its expected value (i.e., $E(Observed_k) = Observed_k = \alpha \times Expected_k \times (1 + \beta \times d_k)$). After substituting $Observed_k$ by $\alpha \times Expected_k \times (1 + \beta \times d_k)$ in the second-order derivatives and the variance-covariance matrix given above, and some simplification, the better approximation of $Cov(\alpha,\beta)$ is given by:

$$Cov(\alpha,\beta) = \begin{pmatrix} \sum Expected_k \times (1 + \beta \times d_k)/\alpha & \sum Expected_k \times d_k \\ \sum Expected_k \times d_k & \alpha \times \sum (Expected_k \times d_k^2) / (1 + \beta \times d_k) \end{pmatrix}^{-1}$$

The determinant for the matrix is

$$Determinant = [\sum Expected_k \times (1 + \beta \times d_k)] \times [\sum (Expected_k \times d_k^2) / (1 + \beta \times d_k)] - (\sum Expected_k \times d_k)^2$$

and the variance of the maximum likelihood estimate of α is

$$var(\alpha) = [\alpha \times \sum (Expected_k \times d_k^2) / (1 + \beta \times d_k)] / Determinant,$$

while the variance of the maximum likelihood estimate of β is

$$var(\beta) = [\sum Expected_k \times (1 + \beta \times d_k)/\alpha] / Determinant,$$

and the standard errors (SE) of the estimated parameters are the square root of their respective variances.

Appendix B (Chapter 7): Example of a Meta-Analysis of Arsenic Cancer Dose Response Models Based on Published Summary Data

B.1 Introduction

The Toxicology Division of the Texas Commission on Environmental Quality (TCEQ) published a draft Development Support Document (DSD) for Arsenic and Inorganic Arsenic Compounds. Section 4.2 of the draft report presented the lung carcinogenic potential of arsenic. The carcinogenic potential was based on dose-response models fitted to summary data of four different cohorts widely published and discussed in the scientific literature.

The draft DSD presented the results of a combined estimate of the potential risk of arsenic on lung cancer in addition to the results derived from each of the four individual cohorts. Section 4.2 of the draft DSD was subjected to a comprehensive peer-review by a panel of scientific experts in the area of cancer risk assessment and dose-response modeling. These experts suggested that one of the cohorts be excluded from the risk assessment of arsenic and that a meta-analysis be performed on the remaining three cohorts. **THE FOLLOWING ANALYSIS IS PRESENTED FOR EXAMPLE PURPOSES ONLY AND DOES NOT REFLECT THE FINAL TOXICITY VALUES FOR ARSENIC ADOPTED BY THE TCEQ.**

B.2 Epidemiological Studies of Arsenic and Cancer

The USEPA developed a cancer risk assessment for arsenic in 1984 (USEPA 1984) and revised the risk estimated in IRIS in 2007. The risk estimates published by EPA are based on excess lung cancer mortality in workers at two smelters: the early data of the Asarco smelter in Tacoma, Washington (Enterline and Marsh 1982; more recently updated by Enterline *et al.* 1987 and 1995) and the Anaconda smelter in Montana (Brown and Chu 1982, 1983a, 1983b; Lee-Feldstein 1983; more recently updated by Lubin *et al.* 2000 and 2008).

In addition to the more recent updates of the Tacoma and Montana studies, there are now two more study cohorts that became available after USEPA's risk assessments and can be used for dose response modeling and risk assessment of arsenic. The Ronnskar Copper Smelter cohort study in Sweden (Järup et al. 1989; Viren and Silvers 1994), and the UK tin smelter cohort study in Humberside, UK (Binks et al. 2005; Jones et al. 2007) were not available in 1984 and also include adequate dose-response data that can be used for dose-response modeling and estimation of potential risks. Summary information about these four studies is given in Table B.1.

Table B.1. Summary information of the four epidemiological studies of arsenic with adequate dose-response data for cancer risk assessment

Study location and exposure period	Most recent dose-response data	Last year of cohort follow-up	Number of workers Person-years (PY)	Cancer site SMR[a] (p-value)	Range of cumulative arsenic exposure (mg/m^3- yr)[b]
Tacoma, WA Asarco copper smelter (1940-64)	Enterline *et al.* (1995)	1986	2,802 84,916	Respiratory 209.7 (p<0.01)	<0.750 to 45+
Montana copper smelter (1938-1958)	Lubin *et al.* 2000; Lubin *et al.* 2008	1989	8,014 144,851[c] (restricted cohort) 256,850 (full cohort)	Respiratory 187 (P<0.001) (restricted cohort) 156 (p<0.001) (full cohort)	1 to 26.2+
Ronnskar, Sweden copper smelter (1928-1967)	Järup *et al.* (1989); Viren and Silvers (1994)	1981	3,916 127,189	Lung 372 (p<0.001)	<0.25 to 100+
United Kingdom tin smelter (1937-1991)	Jones *et al.* 2007	2001	1,462 35,942	Lung 161 (p<0.001)	<0.002 to 4.5+[d]

[a] SMR, standardized mortality ratio; reported results from most recent study

[b] milligrams of arsenic per cubic meter per year (mg/m3-yr)

[c] PY data were obtained from Table 1 of Lubin et al. 2008 for the Montana copper smelter because it is the most recent analysis of the data

[d] the distribution of the cumulative arsenic exposures are given in Table 2 of Jones et al. (2008) for the UK smelter

B.2.1 Tacoma, Washington (Asarco Smelter, Enterline et al. 1982, 1987a, 1987b, and 1995)

Enterline *et al.* (1995) investigated the risks of cumulative arsenic exposure on updated worker respiratory cancer mortality information. An earlier study of 2,802 men who worked at a copper smelter at Tacoma, WA for a year or more during the period of 1940-1964 and were followed up for deaths during the period 1941-1976 was updated with follow-up until 1986. Estimates of exposure for the period 1977-1984 were added. Exposure to arsenic air concentrations were estimated from departmental air arsenic and workers urinary arsenic measurements. Information on smoking was not available. The SMR for respiratory cancer was 209.7 for the total cohort ($p < 0.01$). There were 1,583 deaths observed in the updated study versus 1,061 deaths (Enterline *et al.* 1982; 1987) and 188 versus 104 respiratory cancers (Enterline *et al.* 1982; 1987). Arsenic concentrations in air were estimated for each department starting from 1938 and urinary arsenic

measurements were estimated from each department and worker starting from 1948. While the measurements of arsenic in air were confined mostly to the departments in which arsenic was thought to be a problem, arsenic in urine was measured for all workers. Enterline *et al.* (1995) reported that the conversion of data of urinary arsenic to air arsenic was made by the identification of departments and years for which data from both air and urinary arsenic concentrations were available and by the determination of the mathematical relation between the two.

In order to estimate actual worker exposure concentrations, the average concentrations of arsenic in air were weighted by hours per shift at the sample location, number of men at the location per shift, and frequency of operation for samples. While prior to 1971 arsenic data were obtained by tape and spot samples, starting in 1971, personal air measurements were available. The authors then constructed an exposure matrix for arsenic in air by department and year from 1938 up to the time the smelter closed in1984. The exposure estimates from 1938 were used for years before 1938. Finally, the cumulative exposure was calculated and reported as $\mu g/m^3$-yr. Enterline *et al.* (1995) reported significant excesses for all malignant neoplasms taken together (i.e., cancers of the large intestine, respiratory system, and bone). For < 20 years since first exposure, only respiratory cancer was significant. However, for ≥ 20 years since first exposure SMRs were generally higher but those that were significant were the same as for the total cohort. Enterline *et al.* (1995) reported the relation between cumulative exposure to arsenic in air expressed as $\mu g/m^3$-yr and cancer of the respiratory system for the entire cohort, for the cohort hired before 1940, and for the cohort hired in 1940 or later. This stratification helped separate workers hired before 1940, who were estimated to have had relatively high exposure concentrations coupled with poor respiratory protection, from workers with relatively lower exposure concentrations coupled with better respiratory protection. USEPA developed a URF from the Enterline and Marsh (1982) study which was based on the results from the Pinto et al. (1976) study. The Pinto et al. (1976) study reported an association between airborne arsenic concentrations and urinary arsenic concentrations. The basis of this relationship was that the urinary arsenic concentration could be used as a biomarker for airborne exposure, and the dose response for arsenic-related lung cancer mortality could be expressed in terms of cumulative urinary arsenic exposure (μg/As/liter urine years). This relationship was expressed with the following formula:

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

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

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

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

where As_{air} is measured as $\mu g/m^3$ and As_{urine} is measured as μg/liter. Enterline *et al.* (1987) reported that Pinto et al. (1976) did not take into account prior arsenic exposure through diet. This resulted in high baseline levels of urinary arsenic (about 150 μg/liter). As such, Enterline *et*

al. (1987) indicated that the Pinto *et al.* (1976) study did not depict the true relationship between urinary arsenic measurements and airborne arsenic levels.

Table B.2 presents the summary information of the most recent update of the Tacoma cohort used in dose-response modeling.

Table B.2. Observed (O), Expected (E) and Standard Mortality Rates (SMRs) of respiratory cancer deaths from Enterline *et al.* (1995)

Cumulative exposure interval ($\mu g/m^3$-yr)	Mean exposure (total cohort)	Total cohort			Hired < 1940			Hired ≥ 1940		
		O	E	SMR	O	E	SMR	O	E	SMR
[0, 750)	405	22	14.29	154.0	2	3.08	65.0	20	11.21	178.4*
[750, 2,000)	1,305	30	17.10	175.5**	5	7.33	68.2	25	9.77	256.0**
[2,000, 4,000)	2,925	36	17.17	209.7**	22	8.93	246.2**	14	8.23	170.1
[4,000, 8,000)	5,708	36	17.00	211.7**	15	10.01	149.9	21	7.00	300.2**
[8,000, 20,000)	12,334	39	15.48	252.0**	28	10.98	255.1**	11	4.50	244.4*
[20,000, 45,000)	28,336	20	7.04	284.0**	14	5.56	251.7**	6	1.48	405.5**
≥45,000	58,957	5	1.58	315.7*	5	1.48	338.7*	0	0.11	---

*P < 0.05 ; **P < 0.01

B.2.2 Montana (Anaconda copper smelter, Lubin et al. 2000, 2008)

Lubin *et al.* (2000) investigated the shape of the dose-response curve in an updated study in workers at the Anaconda copper smelter in Montana, initially investigated by Brown and Chu (1982, 1983a, 1983b) and Lee-Feldstein (1986). More recently, Lubin *et al.* (2008) evaluated the shape of the dose-response relationship between respiratory cancer mortality and cumulative exposure to arsenic and the modification of this relationship by the average exposure concentration. The Lubin *et al.* (2000, 2008) analyses are based on an earlier study of 8,047 men who worked at the Montana smelter for a year or more from 1938 to 1957. These workers were followed for deaths during the period 1938-1977 and the study was updated until the end of 1989. Significantly increased SMRs of 156 and 187 were found for respiratory cancer in a restricted sub-cohort and in the full cohort, respectively.

Cumulative exposure estimates in mg/m^3-yr were based on employment records of workers in jobs with light (L), medium (M), and heavy (H) arsenic exposure and measurements of airborne arsenic concentration between 1943 and 1958. The cumulative exposure to arsenic was calculated as

$$\text{Cumulative Exposure} = 0.29\times L + 0.58\times M + (\gamma\times 11.3)\times H$$

where L, M and H are the lengths of time in years that the worker was in jobs with low, medium and high exposure concentrations, respectively, and 0.29, 0.58 and 11.3 are the arsenic air concentrations in jobs with low, medium and high exposures. The weighting factor, γ, is a fraction that measures the proportion of exposure reduction in high-concentration jobs due to the use of protective equipment.

The cumulative exposure estimates calculated by Lubin *et al.* (2000) based on duration in jobs with low and medium exposure concentrations and time of exposure in areas of heavy exposure

were re-calculated using a weighting factor γ of 0.1 to take into account the reduction in exposure due to the use of air filtration masks in heavy-exposure jobs. This resulted in more representative arsenic cumulative exposure estimates that were lower than the estimates using a weighting factor γ of 1.0, particularly at the highest cumulative exposures. Furthermore, using the weight of 0.1 on high-exposure jobs resulted in: 1) rate ratios that conformed to a linear dose-response relationship with cumulative exposure to arsenic and 2) steeper estimates of the slopes, which imply more health-protective excess risks of respiratory cancer deaths.

Approximately 40% of the 8,014 workers in the Montana cohort quit work at the smelter at an early age: 1,616 (20%) were under 30 years of age, and 1,565 (20%) were between 30 and 39 years of age. In order to minimize the impact of unmeasured exposures on workers that stop working at the smelter, Lubin *et al.* (2000) performed analyses both on the full cohort and on data restricted to current workers and to former workers who stopped working at the smelter at age 50 years or older.

Table 2 in Lubin *et al.* (2008) lists the mean cumulative exposure to arsenic (mg/m^3-yr), the number of respiratory cancers and the standardized mortality ratios (SMRs) for six cumulative exposure intervals for the full cohort, unadjusted and adjusted for calendar period and country of birth. The SMRs for respiratory cancers adjusted for calendar period and country of birth are more appropriate that the unadjusted SMRs that are also listed in Table 2. The adjusted SMRs include the effects of possible fluctuations of background respiratory cancer mortality rates in different calendar years and different countries of birth. The relevant data extracted from Table 2 of the Lubin *et al.* (2008) paper are shown in Table B.3.

Table B.3. Observed, Expected and Standard Mortality Rates (SMRs) of respiratory cancer deaths from Table 2 in Lubin *et al.* (2008)

Cumulative exposure interval ($\mu g/m^3$-yr)	Mean exposure ($\mu g/m^3$-yr)	Observed number of respiratory cancer deaths	Expected* number of respiratory cancer deaths	SMR
[0, 750)	470	62	73.81	84
[750, 2,000)	1,240	96	75.00	128
[2,000, 5,000)	3,430	74	68.52	108
[5,000, 10,000)	7,270	83	74.77	111
[10,000, 15,000)	11,900	84	50.00	168
≥15,000	21,900	47	20.00	235

*Expected = Observed / SMR

B.2.3 Ronnskar, Sweden (copper smelter, Järup et al. 1989 and Viren and Silvers 1994)

Järup *et al.* (1989) investigated lung cancer mortality in a cohort of 3,916 male Swedish smelter workers employed for at least three months from 1928 to 1967, and followed through 1981. Lung cancer mortality was related to the estimated average intensity of exposure to arsenic but not to duration. There was also no evident dose-response relationship between estimated exposure to sulfur dioxide and lung cancer.

Järup *et al.* (1989) indicate that "data suggest that arsenic concentration is more important than duration of exposure for the risk of developing lung cancer." In addition, Järup *et al.* (1989) "did not find a clear dose-response relationship in the low exposure categories." These two statements in Järup *et al.* (1989) are consistent with Lubin *et al.* (2008) conclusions that their results suggested a "direct concentration effect on the exposure-response relationship, indicating that for a fixed level of cumulative arsenic exposure, inhalation of higher concentrations of arsenic over shorter durations was more deleterious than inhalation of lower concentrations over longer durations."

Viren and Silvers (1994) used summary data from Järup *et al.* (1989) to fit a dose-response model but did not provide variance estimates of the parameters in the model.

Table B.4 contains the summary data from Järup *et al.* (1989) and Viren and Silvers (1994). Ranges of dose categories in (mg/m^3-yr) were provided by Järup *et al.* (1989) and Viren and Silvers (1994) used the midpoint of each cumulative exposure level as a measure of dose. For cumulative exposure exceeding 100 mg/m^3-yr, an interval that was open-ended, Viren and Silvers (1994) assumed that the median exposure in this group was 25% greater than the lower bound of the given interval.

Table B.4. Observed (O), Expected (E) and Standard Mortality Rates (SMRs) of lung cancer deaths from Järup *et al.* (1989) and Viren and Silvers (1994)

Cumulative exposure interval ($\mu g/m^3$-yr)	Range midpoint[a] ($\mu g/m^3$-yr)	Total cohort[b]			Hired < 1940[b]			Hired ≥1940[a]		
		O	SMR[c]	E[d]	O	SMR[c]	E[d]	O	SMR[c]	E[d]
[0, 250)	125	14	271	5.17	3	284	1.06	11	267	4.12
[250, 1,000)	625	13	360	3.61	3	603	0.50	10	319	3.13
[1,000, 5,000)	3,000	17	238	7.14	6	223	2.69	11	247	4.45
[5,000, 15,000)	10,000	15	338	4.44	10	285	3.51	5	537	0.93
[15,000, 50,000)	32,500	29	461	6.29	27	448	6.03	2	757	0.26
[50,000, 100,000)	75,000	6	728	0.82	6	728	0.82	---	---	---
≥100,000	125,000	12	1,137	1.06	12	1,137	1.06	---	---	---

[a] Data from Viren and Silvers (1994)

[b] Data from Table IV and V (Järup et al. 1989)

[c] SMR, no latency period

[d] E was not provided, but was calculated based on E = O / SMR

B.2.4 Humberside, United Kingdom (tin smelter, Jones et al. 2007)

Jones *et al.* (2007) investigated the relationships between excess lung cancer mortality at a UK tin smelter in Humberside, UK and inhalation exposure to lead, antimony, arsenic, cadmium and radioactivity, with the aim of identifying the cause or causes of the excess lung cancer and quantitative measures of exposures. The cohort was composed of male former employees at the tin smelter initially investigated by Binks *et al.* (2005).

Jones *et al.* (2007) results indicated there were no significant associations found between lung cancer mortality and simple cumulative exposure to any of the substances studied. However, when cumulative exposures were weighted according to time since exposure and attained age, significant associations were found between lung cancer mortality and exposures to arsenic, lead and antimony. Jones *et al.* (2007) concluded that the excess of lung cancer mortality in the cohort can most plausibly be explained if arsenic is the principal occupational carcinogen (for which the ERR diminishes with time since exposure and attained age) and if there is a contribution to excess mortality from an enhanced prevalence of smoking within the cohort. The implications of the dose-response for arsenic exposure for risk estimation merit further consideration.

Jones *et al.* (2007) established exposure matrices for arsenic, cadmium, lead, antimony, and polonium-210 (210Po). They used numerous air sample measurements to estimate the concentrations of the different agents at the smelter. These measurements were recorded for the period 1972 to 1991. Jones *et al.* (2007) commented that the use of 'area' or 'static' air samples may not be representative of the air breathed by workers and are likely to underestimate true personal exposures. In addition, Jones *et al.* used work history for each cohort member to

calculate the exposure profile of each worker. The air concentration measurements for jobs that started before calendar years 1972 were not available (there were work histories starting in 1937). Jones *et al.* extrapolated exposures concentrations to years prior to 1972 using three alternative extrapolation assumptions. The alternative extrapolation assumption that back-extrapolates each process area on a linear increasing trend from the baseline value (average of the three earliest years for which data were available) to values 2-fold higher in the early 1940s is based on a weak trend seen in per-caput average exposure levels over the period 1972-1991. This extrapolation assumption is consistent with exposure estimation model of most epidemiological studies. That is, in most epidemiological studies, exposure concentrations tend to be higher in the early years and decrease to lower exposure concentrations in later years. In addition, Dr. Jones, in a personal communication, also recommended this extrapolation assumption as being more realistic.

Jones *et al.* (2007) provided summary data only for weighted cumulative exposure to arsenic but did not provide summary data for un-weighted cumulative exposure to arsenic. The weighted cumulative exposure to arsenic is an exposure that is modified by other factors that weight the effect that the concentration might have on lung cancer. Arsenic is rapidly cleared from the body after intake, so weighting factors were used to investigate how arsenic concentration and the time since exposure exert a modifying effect on the carcinogenic process during the period of exposure, similar to radon daughters. Jones *et al.* suggest using a weighted dose metric that diminishes the risk of lung cancer with the time since exposure and the age of the worker. They indicate that Binks *et al.* (2005) "found evidence of diminution of lung cancer risk with time since exposure." The weights used by Jones *et al.* to calculate the weighted cumulative exposure were taken from the "exposure-age-concentration model" in BEIR VI (Tables 3-3 and A-4 in NRC 1999). These weights were initially derived from dose-response models for exposures to radon progeny. The weighted cumulative exposure used by Jones *et al.* is as follows:

$$\text{Weighted Cumulative Exposure at age n} = \varphi_n \times \sum_{i=1 \text{ to } n} C_i \times \theta_{n-i}$$

where C_i is the exposure concentration at age i,

$$\varphi_{age} \begin{cases} = 1 & \text{if age} < 50 \text{ years} \\ = 4.8 - 0.105 \times \text{age} + 0.000575 \times \text{age}^2 & \text{if } 50 \text{ years} \le \text{age} < 80 \text{ years} \\ = 0.09 & \text{if age} \ge 80 \text{ years} \end{cases}$$

and, defining time since exposure (tse) as age n minus i in the above equation,

$$\theta_{tse} \begin{cases} = 0 & \text{if tse} < 5 \text{ years} \\ = 1 & \text{if } 5 \text{ years} \le \text{tse} < 10 \text{ years} \\ = 1.17 - 0.0145 \times \text{tse} - 0.00025 \times \text{tse}^2 & \text{if } 10 \text{ years} \le \text{tse} < 30 \text{ years} \\ = 0.51 & \text{if tse} \ge 30 \text{ years} \end{cases}$$

Jones *et al.* smoothed the step function for φ_{age} and θ_{tse} specified in Tables 3-3 and A-4 in BEIR VI (NRC 1999). Jones *et al.* indicate that fitting the models using the smoothed function and the step-function version of the weights result in approximately the same estimates. For a fixed

concentration or exposure rate, the weighted cumulative exposure used by Jones *et al.* (and also proposed in the "exposure-age-concentration model" in BEIR VI) is zero for the first 5 years of exposure, then increases for the next 45 years followed by a decrease for the next 30 years, to slowly increase following 80 years after the first exposure. Figure B.1 shows the weighted cumulative exposure as a function of age using the smoothed weights derived by Jones *et al.*

Figure B.1. Weighted Cumulative Exposure using the Smoothed Weights for a Concentration of 1 (Jones *et al.* 2007)

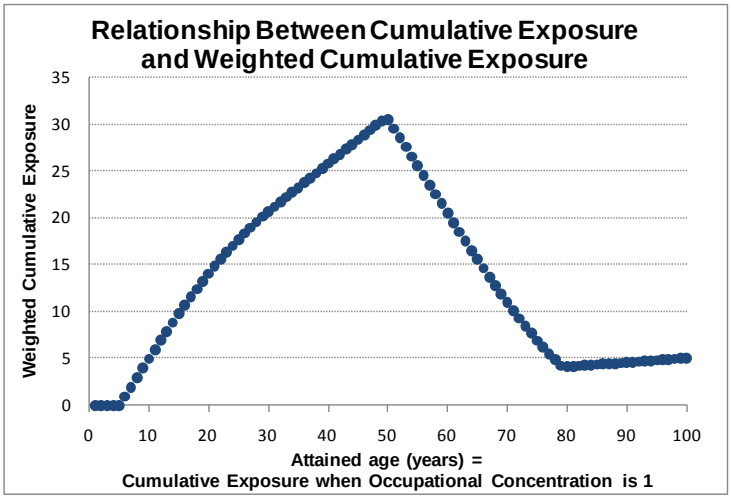


Table B.5 contains summary the data from Jones *et al.* (2007). Ranges and averages of weighted cumulative exposure to arsenic (µg/m^3-yr) were provided.

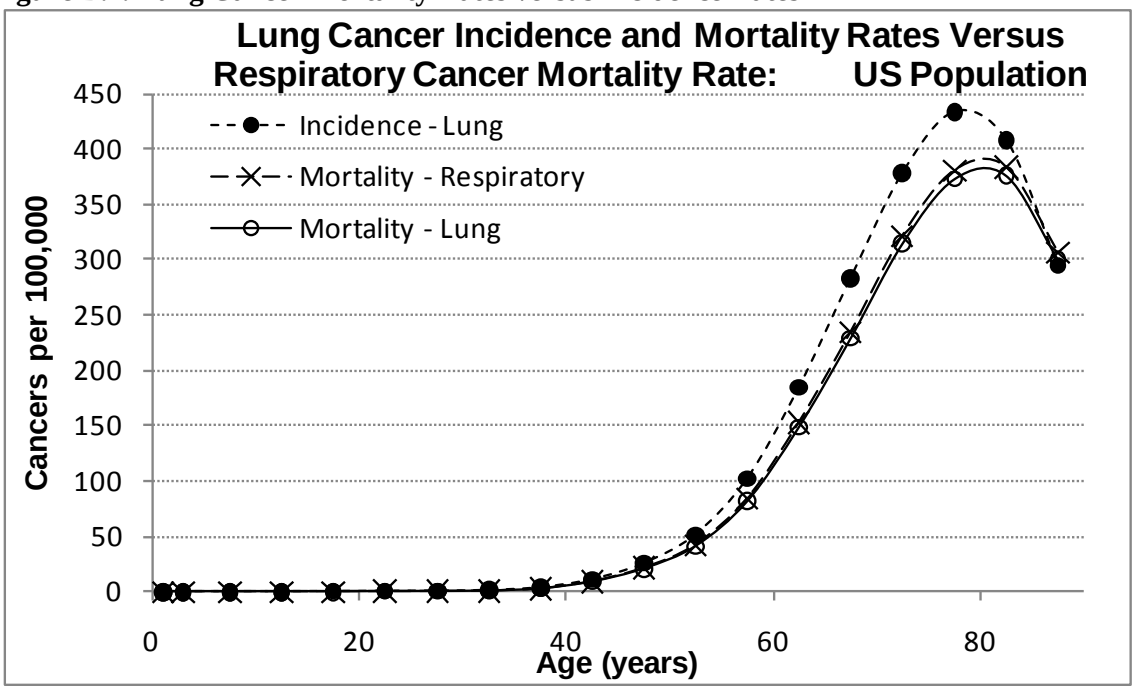
Table B.5. Observed, Expected, and SMR of lung cancer deaths from Jones *et al.* (2007)

Arsenic weighted cumulative exposure range (µg/m^3-yr)	Arsenic Mean of weighted cumulative exposure (µg/m^3-yr)	Observed	Expected	SMR
[0.0, 45]	9.7	13	10.4	125
(45, 120]	81	12	6.7	178
(120, 320]	210	12	8.8	137
(320, 710]	480	12	7.5	160
> 710	1400	13	4.9	267

B.3 Dose Metric and Dose-Response Assessment

Enterline *et al.* (1995) and Lubin *et al.* (2000; 2008) examined *respiratory cancer mortality* (i.e., larynx, bronchus, trachea, lung, and other residual) by cumulative arsenic exposure level whereas the other two epidemiology studies investigated lung cancer mortality (Järup *et al.* 1989; Jones *et al.* 2007). Lung cancer mortality is the cancer endpoint of interest for all four studies for risk estimation purposes. The respiratory cancer mortality data from Enterline *et al.* (1995) and Lubin *et al.* (2000; 2008) are a reasonable surrogate for lung cancer as most (96 %) of the observed deaths (i.e., 182 out of 188 and 428 out of 446, respectively) were due to lung cancer. Also, as lung cancer mortality is reasonably predictive of lung cancer incidence (i.e., five-year survival is only about 15% according to the American Cancer Society 2005), the cancer potency estimates based on the two studies and the resulting calculations are comparable (i.e., lung cancer incidence and mortality are sufficiently similar as to be comparable for purposes of this assessment; see Figure B.2).

Figure B.2. Lung Cancer Mortality Rates versus Incidence Rates[a]



[a]Based on US lung cancer mortality and incidence rates

The dose metric used for the dose-response assessments is cumulative arsenic exposure ($\mu g/m^3$-yr) because it is the only measure available for all cohort studies and because there are no biological/mechanistic data or statistical evidence which indicates that another dose metric is more appropriate, except for Lubin *et al.* (2008). Lubin *et al.* (2008) investigated a dose metric of cumulative arsenic exposure modified by the concentration. Similarly, Jones *et al.* (2007) used a novel weighted cumulative exposure to arsenic where the weights were based on age attained and the time since exposures occurred. Unfortunately, Jones *et al.* (2007) did provide summary statistics corresponding to standard, un-weighted, cumulative exposure to arsenic similar to the other three studies.

B.4 Dose-Response Models

Standard regression analysis approaches for survival data (Poisson regression and Cox regression) are considered more reliable and less restricted to calculate the maximum likelihood estimates of the slope (β) and corresponding variance when the necessary detailed data are available (e.g., can adjust for covariate effects and use internally-derived background hazard rates). While results of the standard Poisson regression analysis were available for Lubin *et al.* (2000; 2008) and Jones *et al.* (2007), only summary data (i.e., observed and expected deaths versus cumulative arsenic exposure levels) were available for Enterline *et al.* (1995) and Järup *et al.* (1989). For these last two studies with the summary data, the linear multiplicative relative risk model and Poisson regression modeling (Crump and Allen 1985) can be used to obtain maximum likelihood estimates of β and the asymptotic variance for β (Appendix A). In addition to the more plausible assumptions regarding the amount of increase in risk with age, the linear multiplicative relative risk model naturally results from the Poisson regression and Cox proportional hazards models when only summary data are available.

The following standard parameterization of the multiplicative relative risk linear model with intercept (Crump and Allen 1985), is used more often and readily usable for excess risk estimation:

$$Observed_k = \alpha \times Expected_k \times (1 + \beta \times d_k)$$

where:

$Observed_k$ = estimated number of cancer cases for exposure group k
$Expected_k$ = expected number of background cancer cases for exposure group k
α = accounts for differences in cancer background rates between the study population and the reference population
β = multiplicative factor by which the background risk increases with cumulative exposure
d_k = cumulative exposure for exposure group k

The model can be modified to adjust for other covariate effects. For the Tacoma and Sweden studies, that divide the workers into two sub-cohorts of workers hired before 1940 and after 1939, an additional parameter can be added to the model to account for potential differences between background mortality rates of the two sub-cohorts. That is, the modified model that accounts for the effect of year of hire is as follows:

$$Observed_k = h \times \alpha \times Expected_k \times (1 + \beta \times d_k)$$

where all the parameters are as before and the parameter h is the effect of being hired after 1939 as opposed to being hired before 1940.

B.5 Model Fitting to Individual Study Cohorts

The **respiratory** cancer mortality data in Tables B.2 and B.3 and the lung cancer mortality data in Tables B.4 and B.5 were fit using the multiplicative relative risk models described above. The data were fit using Poisson regression, external background lung cancer rates, and assuming a linear dose-response model. The maximum likelihood estimates (MLEs) of the parameters (intercept α and slope β, and, when appropriate, the effect of year of hire h). Poisson regression with externally derived background cancer rates implicitly adjusts for age when the reference population background rates are calculated using age-dependent background cancer rates.

The Tacoma and Sweden study cohorts made a distinction between workers hired 1940 and those hired after 1939. Estimates for all workers adjusting for the year of hire include all the information in the cohort and, therefore, are emphasized here because they are more reliable. However, models were also fit to i) all workers with no adjustments, ii) workers hired <1940, and iii) workers hired ≥1940 for comparison purposes. The MLE, standard error (SE), and 95% upper confidence limit on the β (95%UCL) were also calculated and are presented.

Table B.6 shows the MLEs for the effect of year of hire (when estimated), intercept (α) and slope (β) along with the standard error of the estimated slope and the corresponding asymptotic 95% UCL on the slope.

Table B.6. Summary of all maximum likelihood estimates, standard errors and 95% upper confidence limits on slopes for the four different cohorts used in the arsenic DSD

Data	Intercept MLE (α)	Slope MLE β $(\mu g/m^3\text{-}yr)^{-1}$	Standard Error Slope (β)	95% UCL on Slope β $(\mu g/m^3\text{-}yr)^{-1}$
Tacoma, Washington (Asarco Smelter, Enterline et al. 1982, 1987a, 1987b, and 1995)				
All workers adjusting for year of hire (h=1.38)	*1.46*	*3.15E-05*	*1.48E-05*	*5.59E-05*
All workers with no adjustment for year of hire	1.81	2.13E-05	1.13E-05	3.99E-05
Workers hired < 1940	1.43	3.44E-05	1.89E-05	6.56E-05
Workers hired 1940+	2.05	2.67E-05	2.33E-05	6.51E-05
Montana (Anaconda copper smelter, Lubin *et al.* 2000, 2008)				
Full Cohort	0.94	5.75E-05	1.61E-05	8.40E-05
Ronnskar, Sweden (copper smelter, Järup et al. 1989 and Viren and Silvers 1994)				
All workers adjusting for year of hire (h=1.19)	2.37	2.92E-05	1.63E-05	5.61E-05
All workers with no adjustment for year of hire	2.67	2.38E-05	9.14E-06	3.89E-05
Workers hired < 1940	2.48	2.62E-05	1.35E-05	4.84E-05
Workers hired 1940+	2.60	6.17E-05	5.92E-05	1.59E-04
Humberside, United Kingdom (tin smelter, Jones et al. 2007) *** Based on WEIGHTED cumulative exposure ***				
Scenario B	1.33	6.49E-04	4.90E-04	1.45E-03

Note that because the dose metric of the UK study is weighted cumulative exposure to arsenic, the slope derived for the UK study is not comparable to the slopes derived for the other three studies. The most reliable MLEs of the slope for the three studies that use the same dose metric (cumulative exposure to arsenic) range from 2.92×10^{-5} to 5.75×10^{-5} per $\mu g/m^3$-yr. The 95% UCL on the slope based on the three most reliable studies with cumulative exposure to arsenic range from 5.59×10^{-5} to 8.40×10^{-5} per $\mu g/m^3$-yr.

Figures B.3 to B.6 show the SMRs, 95% confidence intervals on the SMRs and the model fit to the four individual cohorts. Figures B.3 and B.5 show the model fit to the two sub-cohorts with two different intercepts but the same slope.

Figure B.3. Model fit to the SMRs of the Tacoma cohort with different intercept for the two sub-cohorts with different hiring calendar years

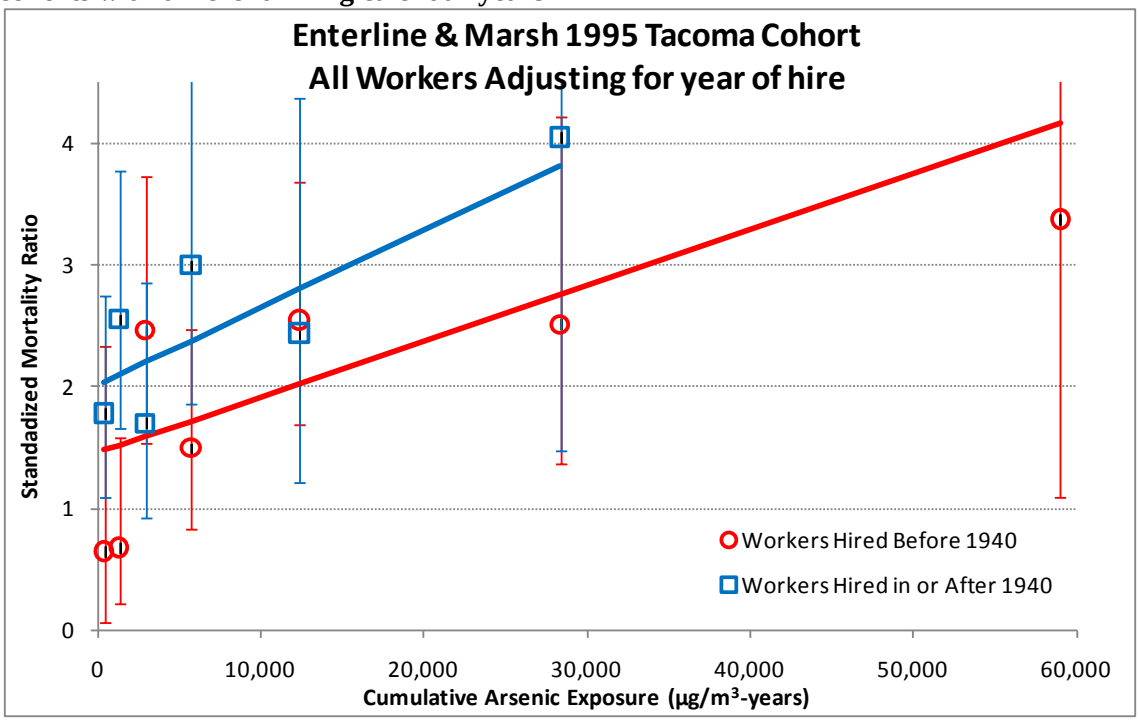


Figure B.4. Model fit to the SMRs of the Montana cohort

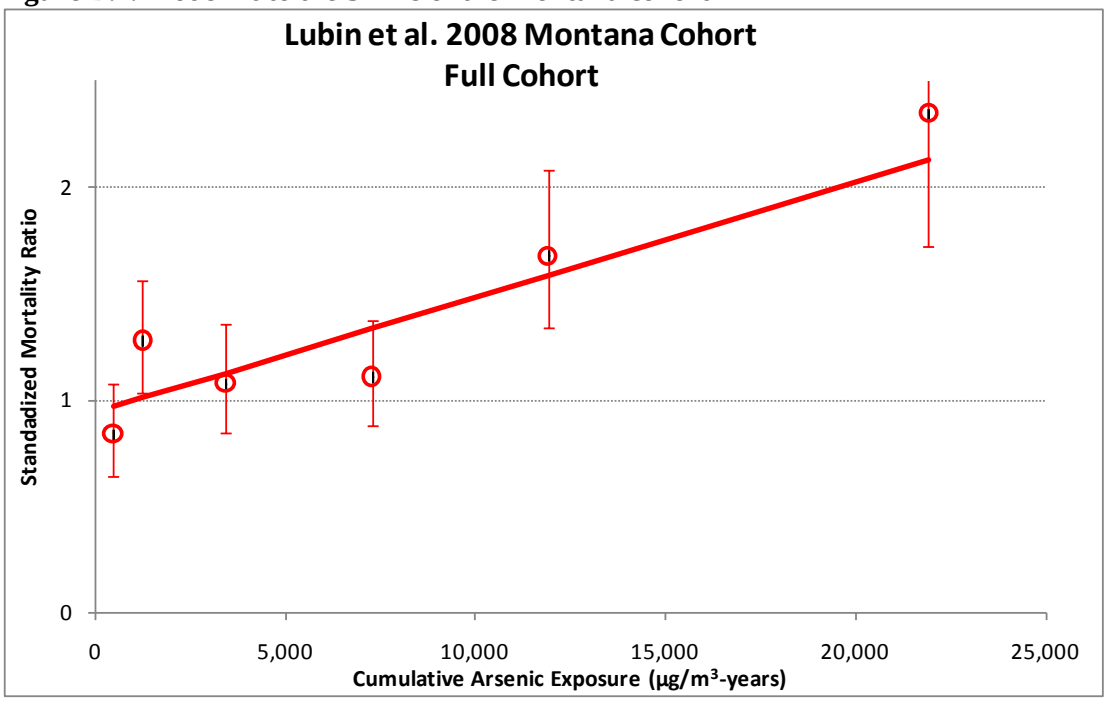


Figure B.5. Model fit to the SMRs of the Sweden cohort with different intercept for the two sub-cohorts with different hiring calendar years

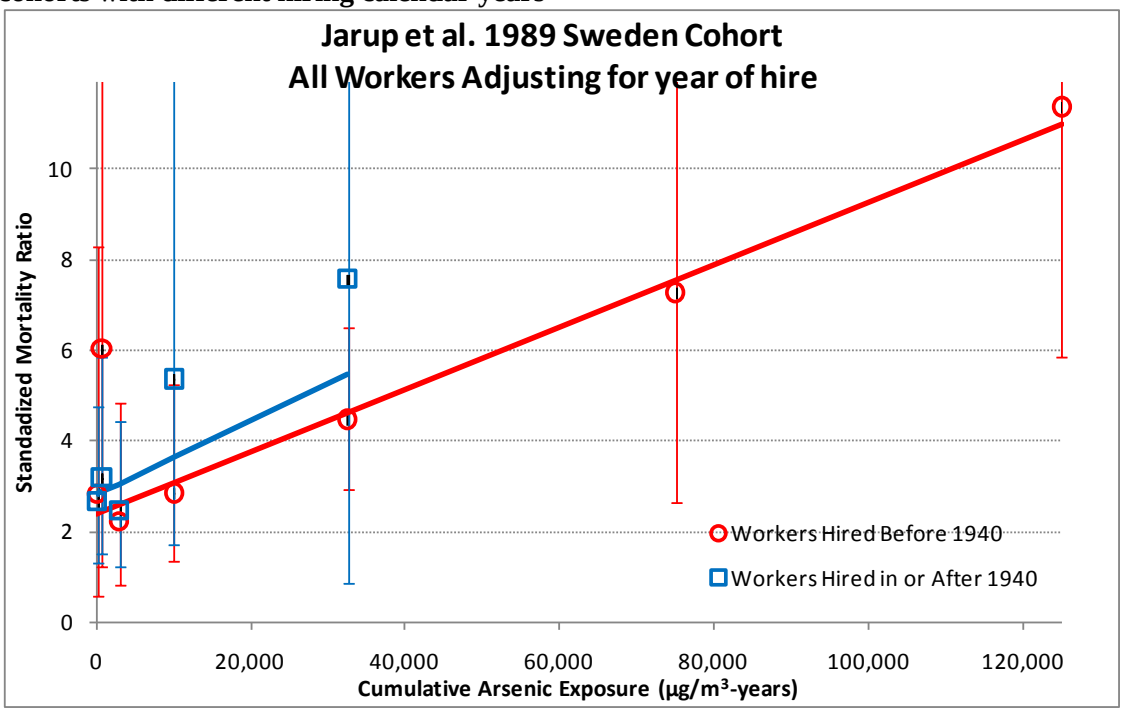
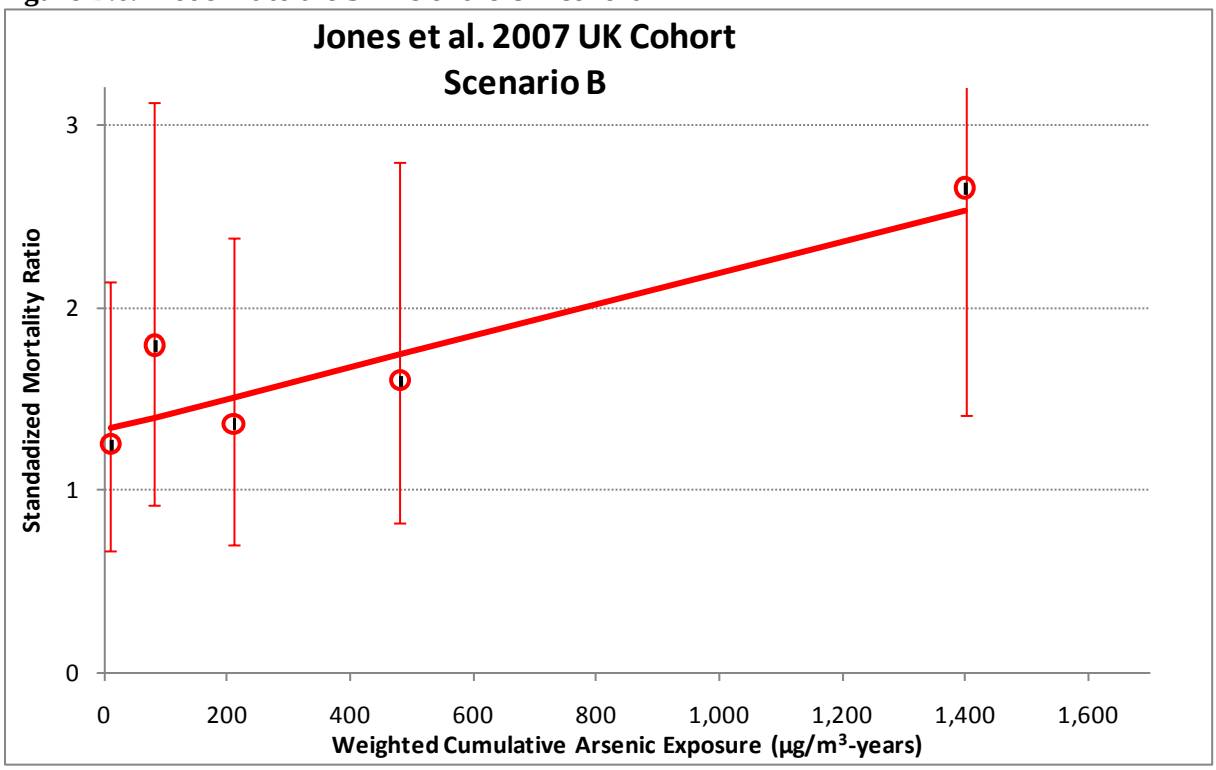


Figure B.6. Model fit to the SMRs of the UK cohort



B.6 Unit Risk Factors for Lung Cancer Mortality

Unit risk factors (URFs) express cancer potency in units of risk per air concentration (e.g., risk per µg/m^3) assuming continuous environmental lifetime exposure. They are calculated using linear low-dose extrapolation when the carcinogenic mode of action (MOA) is unknown, which is the case for arsenic. Where a dose-response curve is modeled for tumor or cancer mortality data, the URF is the slope of a straight line from a specified point of departure (POD) to the origin, with the POD being the lowest tumor response or cancer mortality response supported by the study data.

An excess risk of 1 additional lung cancer in a 100,000 is a reasonable POD for models based on human data. The models used here are based on human epidemiological studies and have been fit to a linear equation (linear multiplicative relative risk model) for use with the BEIR IV methodology (NRC 1988). Thus, extrapolated URFs using a high POD and URFs estimated using a small POD are approximately equal.

Air concentrations are solved iteratively with life-table analyses using the BEIR IV approach (NRC 1988). Air concentrations based on extra risk are calculated as opposed to added risk. Mortality and survival rates are used to calculate air concentrations based on a lifetime exposure of 70 years, the default used by TCEQ for exposure analysis (TCEQ 2006). Texas-specific mortality rates for 2001-2005 for lung cancer and Texas-specific survival rates for 2005 are used in the calculation of PODs and URFs.

Table B.7 shows estimates of URFs based on PODs of air concentrations at 1 in 100,000 excess lung cancer mortality risk (10^{-5}-risk air concentrations) for on β (MLE) and β (95% UCLs) from Table B.6 and using Texas mortality and survival rates. URFs and arsenic 10^{-5}-risk air concentrations were calculated using β values for all workers adjusting for year of hire, all workers with no adjustment for year of hire, and workers hired <1940. URFs were not calculated for workers hired 1940+ because Viren and Silvers (1999) found no association for those workers, (i.e., regression didn't achieve statistical significance at $P < 0.01$ based on the corresponding likelihood ratio statistic).

Table B.7. Preferred maximum likelihood estimates and 95% upper confidence limits on the on the URFs for the four individual studies

Study and number of person years	URF per $\mu g/m^3$	
	MLE	95% UCL
Tacoma, Washington (Asarco Smelter, Enterline et al. 1982, 1987a, 1987b, and 1995) All workers adjusting for year of hire 84,916 person-years	1.19E-04	2.12E-04
Montana (Anaconda copper smelter, Lubin et al. 2000, 2008) Full cohort 256,850 person-years	2.18E-04	3.19E-04
Ronnskar, Sweden (copper smelter, Järup *et al.* 1989 and Viren and Silvers 1994) All workers adjusting for year of hire 127,189 person-years	1.11E-04	2.13E-04
Humberside, United Kingdom (tin smelter, Jones et al. 2007) *** Based on WEIGHTED cumulative exposure *** 35,942 person-years	7.04E-04	1.57E-03

In Table B.7 the URFs are comparable because all are excess risks per unit of environmental concentration of arsenic. That is, the model for the U.K. study has been appropriately adjusted in the estimation of the POD to factor out the weights of the cumulative exposure and, consequently, the resulting POD and URF are in units of environmental concentrations of arsenic. The MLEs of URFs range from 1.11×10^{-4} to 7.04×10^{-4} while the 95% UCLs on the URFs range from 2.12×10^{-4} to 1.57×10^{-3}.

B.7 Meta-Analyses Derivation of URFs

The values of the URFs and the corresponding 95 UCLs on the URFs derived from the individual studies listed in Table B.7 have a range of values. Although the range of URFs is not very large (less than a factor of seven), it is advantageous for regulatory purposes to have a

single URF that best describes the potential excess of lung cancers per unit of environmental concentration.

There are several alternative methods that could lead to the establishment of a single URF. Regulatory agencies oftentimes choose the largest URF because it is the most protective URF for risk assessment purposes. Alternatively, other simple approaches take the arithmetic average or the geometric mean of the URFs. These simplistic approaches, however, do not make the best use of the available information in the data.

Meta-analyses are statistical methods that use objective criteria to combine results from the analyses of individual studies or to fit models to the combined epidemiological data. These meta-analyses usually result in more powerful models that take advantage of all the information in the epidemiological studies. At the same time, meta-analyses are more robust than the simplistic approach of estimating MLEs and should result in better upper bound estimates on the URF.

Meta-analyses methods can take on several different forms. The meta-analyses for the risk assessment of arsenic take on four different forms. The results of seven alternative meta-analyses performed using these four forms on the individual studies are presented here.

B.7.1 Meta-Analysis Using Person-Years to Weight Individual URFs

The preferred URFs based on Enterline *et al.* (1995), Lubin *et al.* (2000 ; 2008), Järup *et al.* (1989), and Jones *et al.* (2007) are considered appropriate estimates of the carcinogenic potency of arsenic based on their respective studies, and ranged from 1.11×10^{-4} per µg/m^3 to 7.04×10^{-4} per µg/m^3, a 6.4-fold difference (Table B.7). The Lubin *et al.* (2000; 2008) study with 256,850 person-years and the Järup *et al.* (1989) studies with 127,189 person-years had significantly more workers and person-years included in the study than the Enterline *et al.* (1995) study with 84,916 person-years and Jones *et al.* (2007) study with 35,942 person-years. Therefore, a weighted URF using person-years (py) was calculated to estimate a combined URF.

$$Combined\ URF\ (Risk\ per\ \mu g/m^3) = \frac{(URF_1 \times py_1) + (URF_2 \times py_2) + (URF_3 \times py_3) + (URF_4 \times py_4)}{py_1 + py_2 + py_3 + py_4}$$

$$Combined\ URF\ (Risk\ per\ \mu g/m^3) = \frac{(1.19\times10^{-4}\times84{,}916) + (2.18\times10^{-4}\times256{,}850) + (1.11\times10^{-4}\times127{,}189) + (7.04\times10^{-4}\times35{,}942)}{84{,}916 + 256{,}850 + 127{,}189 + 35{,}942}$$

$$Combined\ URF\ (Risk\ per\ \mu g/m^3) = 2.09\times10^{-4} per\ \mu g/m^3$$

The combined best estimate of the person-year-weighted URF based on Texas lung cancer mortality rates and survival probabilities is 2.09×10^{-4} per µg/m^3 which is similar to the geometric mean of the URFs from the four studies of 2.12×10^{-4} per µg/m^3.

The 95% UCL on the URF can be calculated in addition to the combined central estimate of the URF. The standard error of the URF derived from each individual study can be back-calculated from the central estimates of the URFs and the corresponding 95% UCLs on the URFs given in Table B.7. The standard error (S.E.) of the URF for each individual study can be calculated as follows:

$$S.E.(URF_i) = (95\%\ UCL\ on\ URF_i - MLE\ of\ URF_i)/1.645$$

where i=1, 2, 3 or 4 is an indicator for the study and 1.645 is the 95-th percentile of the standard normal distribution. The resulting standard errors for the Tacoma, Montana, Ronnskar and UK studies are equal to 5.65×10^{-5}, 6.14×10^{-5}, 6.20×10^{-5}, and 6.26×10^{-4}, respectively.

The standard error of the person-year-weighted combined URF is then given by,

$$S.E.(combined\ URF) = \sqrt{\frac{py_1^2 \times S.E.(URF_1)^2 + py_2^2 \times S.E.(URF_2)^2 + py_3^2 \times S.E.(URF_3)^2 + py_4^2 \times S.E.(URF_4)^2}{(py_1 + py_2 + py_3 + py_4)^2}}$$

$$S.E.(combined\ URF) = \sqrt{\frac{(84{,}916 \times 5.65 \times 10^{-5})^2 + (256{,}850 \times 6.14 \times 10^{-5})^2 + (127{,}189 \times 6.20 \times 10^{-5})^2 + (35{,}942 \times 6.26 \times 10^{-4})^2}{(84{,}916 + 256{,}850 + 127{,}189 + 35{,}942)^2}}$$

$$S.E.(combined\ URF) = 5.21 \times 10^{-5}$$

Thus, the 95% UCL on the combined URF is given by

$$95\%\ UCL\ on\ combined\ URF(Risk\ per\ \mu g/m^3) = combined\ URF\ (Risk\ per\ \mu g/m^3) + 1.645 \times S.E.(combined\ URF)$$

$$95\%\ UCL\ on\ combined\ URF(Risk\ per\ \mu g/m^3) = 2.09 \times 10^{-4} + 1.645 \times 5.21 \times 10^{-5} = 2.95 \times 10^{-4}\ per\ \mu g/m^3$$

Using the same procedure, but excluding the UK study, it can be shown that the combined URF and corresponding 95% UCL based on the three studies that use the same cumulative exposure to arsenic (Tacoma, Montana and Sweden) are 1.71×10^{-4} and 2.35×10^{-4} per $\mu g/m^3$, respectively.

B.7.2 Meta-Analysis Using Inverse Variance of the URFs to Weight Individual URFs

The inverse of the variance of the individual URFs can be used as weights instead of person-years in each study to combine the individual URFs. Simulation experiments have shown that the inverse-variance weighting results in minimum variance estimates while sample-size or person-years weighting results in less biased estimates (Sanchez-Meca and Marin-Martinez, 1998). Although, the number of person-years plays a role on the size of the estimated variance of the URF, the inverse-variance weighting is a more standard statistical procedure used in meta-analyses.

The inverse-variance weighting meta-analyses require the estimated variance of the individual URFs. The variance of the URFs are given as before using the following formula

$$S.E.(URF_i) = (95\%\ UCL\ on\ URF_i - MLE\ of\ URF_i)/1.645$$

where, as before, i=1, 2, 3 or 4 is an indicator for the study and 1.645 is the 95-th percentile of the standard normal distribution. The resulting standard errors for the Tacoma, Montana, Sweden and UK studies are equal to 5.65×10^{-5}, 6.14×10^{-5}, 6.20×10^{-5}, and 6.26×10^{-4}, respectively.

The MLE of the combined URF is now given by the weighted average of the MLEs of the individual URFs. The weights are the inverse of the squared S.E.'s of the individual URFs. That is,

$$Combined\ URF\ (Risk\ per\ \mu g/m^3) = \frac{(w_1 \times URF_1) + (w_2 \times URF_2) + (w_3 \times URF_3) + (w_4 \times URF_4)}{w_1 + w_2 + w_3 + w_4}$$

where $w_i=[1/S.E.(URF_i)]^2$ for i=1, 2, 3, 4. The combined best estimate of the inverse-variance-weighted URF based on Texas lung cancer mortality rates and survival probabilities is 1.50×10^{-4} per $\mu g/m^3$ which is approximately 1.4-fold smaller than the person-years-weighted combined URF of 2.09×10^{-4} per $\mu g/m^3$.

The 95% UCL on the combined URF can be calculated in addition to the combined central estimate of the URF. Here, the standard error of the inverse-variance-weighted URF is simply given by the square root of the inverse of the sum of the weights; that is,

$$S.E.(combined\ URF) = \sqrt{\frac{1}{w_1 + w_2 + w_3 + w_4}}$$

where, again, $w_i=[1/S.E.(URF_i)]^2$ for i=1, 2, 3, 4. The resulting standard error of the combined URF is then equal to 3.45×10^{-5}. Thus, the 95% UCL on the combined URF is given by

$$95\%\ UCL\ on\ combined\ URF(Risk\ per\ \mu g/m^3) = URF\ (Risk\ per\ \mu g/m^3) + 1.645 \times S.E.(combined\ URF)$$

$$95\%\ UCL\ on\ combined\ URF(Risk\ per\ \mu g/m^3) = 1.50 \times 10^{-4} + 1.645 \times 3.45 \times 10^{-5} = 2.07 \times 10^{-4}\ per\ \mu g/m^3$$

Using the same procedure, but excluding the UK study, it can be shown that the inverse-variance-weighted combined URF and corresponding 95% UCL based on the three studies that use the same cumulative exposure to arsenic (Tacoma, Montana and Sweden) are 1.48×10^{-4} and 2.05×10^{-4} per $\mu g/m^3$, respectively.

B.7.3 Meta-Analysis Using Inverse Variance of the Estimated Slopes to Weight Individual Slopes

The meta-analyses based on the weighted URFs can combine the four studies because the individual URFs are in units of environmental air concentration. Thus, the person-year and inverse-variance combined URFs had estimates that included the four studies in addition to estimates that included only the three studies with similar dose metric.

In contrast to the URFs that are based on the same units, the estimated parameters of the models fit to the data (specifically the slope β) are not based on the same dose metric of arsenic exposure. In fact, some experts recommended that the UK study be excluded from the meta-analyses because the dose metric used in unconventional and not well understood. The weighted cumulative exposure to arsenic used in the UK study differed from the standard cumulative exposure to arsenic used in the other three studies (Tacoma, Montana and Sweden).

Consequently, only the later three studies can be used if a meta-analysis based on the slopes of the individual studies is to be used to derive a single URF.

A meta-analysis based on the slopes to derive a single URF should result in more accurate estimates in that some assumptions made when combining individual URFs are no longer necessary. This alternative meta-analysis based on the slopes, however, does require other assumptions that are more commonly satisfied by the MLE of the slopes. The main assumption is the asymptotic normality of the maximum likelihood estimate of the parameter.

The meta-analysis based on the slopes estimated from the individual studies first determines a combined slope and then the combined slope is used calculate an URF. The combined slope results from a weighted combination of the individual slopes. The standard inverse-variance weighting can be used to obtain a combined slope. The inverse-variance weighting meta-analyses to estimate a combined slope β requires the estimated variance of the individual β estimates. The squared root of the variance (S.E.) of the individual slopes are given in Table B.6 and are equal to 1.48×10^{-5}, 1.61×10^{-5} and 1.63×10^{-5} for the Montana, Tacoma and Sweden studies, respectively.

The MLE of the combined slope is now given by the weighted average of the MLEs of the individual slopes. The weights are the inverse of the squared S.E.'s of the individual URFs. That is,

$$Combined\ \beta\ (Slope\ per\ \mu g/m^3 - yr) = \frac{(w_1 \times \beta_1) + (w_2 \times \beta_2) + (w_3 \times \beta_3)}{w_1 + w_2 + w_3}$$

where $w_i=[1/S.E.(\beta_i)]^2$ for i=1, 2, 3. The central estimate of the inverse-variance-weighted combined β is 3.90×10^{-5} per $\mu g/m^3$-yr.

In order to estimate a 95% UCL on the URF, the 95% UCL on the combined slope needs to be calculated first. This combined 95% UCL on the inverse-variance weighting combined slope can be calculated using the same approach used before. That is, the standard error of the inverse-variance weighting combined slope is calculated first and this standard error is used to calculate a 95% UCL on the combined slope.

The standard error of the inverse-variance weighting combined slope is given by

$$S.E.(combined\ \beta) = \sqrt{\frac{1}{w_1 + w_2 + w_3}}$$

where, again, $w_i=[1/S.E.(\beta_i)]^2$ for i=1, 2, 3. The resulting standard error of the combined β is then equal to 9.06×10^{-6}. Thus, the 95% UCL on the combined β is given by

$$95\%\ UCL\ on\ combined\ \beta(Slope\ per\ \mu g/m^3 - yr) = combined\ \beta\ (Slope\ per\ \mu g/m^3 - yr) + 1.645 \times S.E.(combined\ \beta)$$

$$95\%\ UCL\ on\ combined\ \beta(Slope\ per\ \mu g/m^3 - yr) = 3.90 \times 10^{-5} + 1.645 \times 9.06 \times 10^{-6} = 5.39 \times 10^{-5}\ per\ \mu g/m^3 - yr$$

The central estimate of the URF and the 95% UCL on the URF are calculated by applying the same methodology used in the calculations of the URFs and 95% UCLs on the URF for the individual studies. That is, air concentrations are solved iteratively with life-table analyses using the BEIR IV approach (NRC 1988). Mortality and survival rates are used to calculate air concentrations based on a lifetime exposure of 70 years, the default used by TCEQ for exposure analysis (TCEQ 2006). Texas-specific mortality rates for 2001-2005 for lung cancer and Texas-specific survival rates for 2005 are used in the calculation of PODs and URFs. The URF is then equal to 10^{-5} divided by the air concentration corresponding to a 1 in 100,000 excess lung cancer mortality risk.

The estimated URF based on the inverse-variance-weighted combined β and the corresponding 95% UCL are equal to 1.48×10^{-4} and 2.04×10^{-4} per $\mu g/m^3$, respectively. These values are essentially equal to the estimates calculated using inverse-variance weighting to combine individual URFs (1.48×10^{-4} and 2.05×10^{-4} per $\mu g/m^3$, respectively).

B.7.4 Meta-Analyses Using Dose-Response Models to Fit the Combined Data

Meta-analyses that combine URFs or slopes are useful when only that information is available. In cases where the data used to derive the individual URFs and slopes are available, more robust meta-analyses can be developed. The available data of the individual studies can be combined and models fit to the combined data.

The arsenic DSD has sufficient information to conduct meta-analyses on the combined data from the three studies with similar dose metric. The linear multiplicative rate ratio model was fit to the combined data using Poisson regression and maximum likelihood estimation. Two alternative model parameterizations were explored. The first model was identical to the models described before and assumed that the intercept α is identical for all the cohorts. The second model assumed that each cohort or sub-cohort could have a different intercept (i.e., each cohort or sub-cohort was allowed to have a different background respiratory or lung cancer mortality rate). The different intercepts in the second model estimated different background cancer mortality rates for the Tacoma workers hired before 1940, the Tacoma workers hired after 1939, the Montana workers, the Sweden workers hired before 1940 and the Sweden workers hired after 1939. Both, the first and second models, estimated a single slope β and its corresponding S.E. and 95% UCL on β.

Table B.8 shows the estimates of the intercepts, slopes, standard errors and 95% UCL on the slopes for the two models fit to the combined data. The table also shows, for comparison purposes, the intercept, slope, standard error and 95% UCL on the slope derived using inverse variance weighting discussed in the previous section.

Table B.8. Estimates of the intercept, slope, standard error and 95% UCL on the slopes resulting from meta-analyses that combine the Tacoma, Montana and Sweden cohorts

Cohort or sub-cohort included in the analyses	Intercept MLE (α)	Slope MLE β $(\mu g/m^3\text{-yr})^{-1}$	Standard error slope (β)	95% UCL on slope β $(\mu g/m^3\text{-yr})^{-1}$
Meta-analysis using inverse-variance weighting to estimate a common intercept α and a common slope β from the intercepts and slopes of the individual cohorts				
All Workers	1.14	3.90E-05	9.06E-06	5.39E-05
Meta-analysis estimating an intercept and a single slope of the multiplicative rate ratio model fit to the combined data from the three cohorts				
All Workers	1.14	5.46E-05	9.84E-06	7.08E-05
Meta-analysis estimating an intercept for each cohort or sub-cohort and a single slope of the multiplicative rate ratio model fit to the combined data from the three cohorts				
Tacoma: Workers hired < 1940	1.35	4.22E-05	9.42E-06	5.77E-05
Tacoma: Workers hired 1940+	1.94			
Montana: All Workers	1.00			
Sweden: Workers hired < 1940	1.97			
Sweden: Workers hired 1940+	2.72			

Figure B.7 shows the fit of the first model to the combined data of the three cohorts (five cohorts and sub-cohorts). The sizes of the symbols are proportional to the number of person-years in the respective cohorts. Cohorts with larger number of person-years carry more weight when the model is fit to the data using the maximum likelihood estimation procedure. Figure B.7, for example, shows that the model emphasizes the fit of the Montana cohort because it has the largest number of person-years.

Figure B.8 shows the fit of the second model to the combined data of the three cohorts (five cohorts and sub-cohorts). Here, too, the sizes of the symbols are proportional to the number of person-years in the respective cohorts. In Figure B.8, however, the cohorts with the larger number of person-years carry more weight in the estimation of the slope but not on the estimation of the intercepts. Different intercepts are estimated for different cohorts or sub-cohorts. Figure B.8, then, shows five lines corresponding to a model that has one single slope but five different intercepts.

Figure B.7. First model with one intercept and one slope fit to the SMRs of the Tacoma, Montana and Sweden cohorts combined

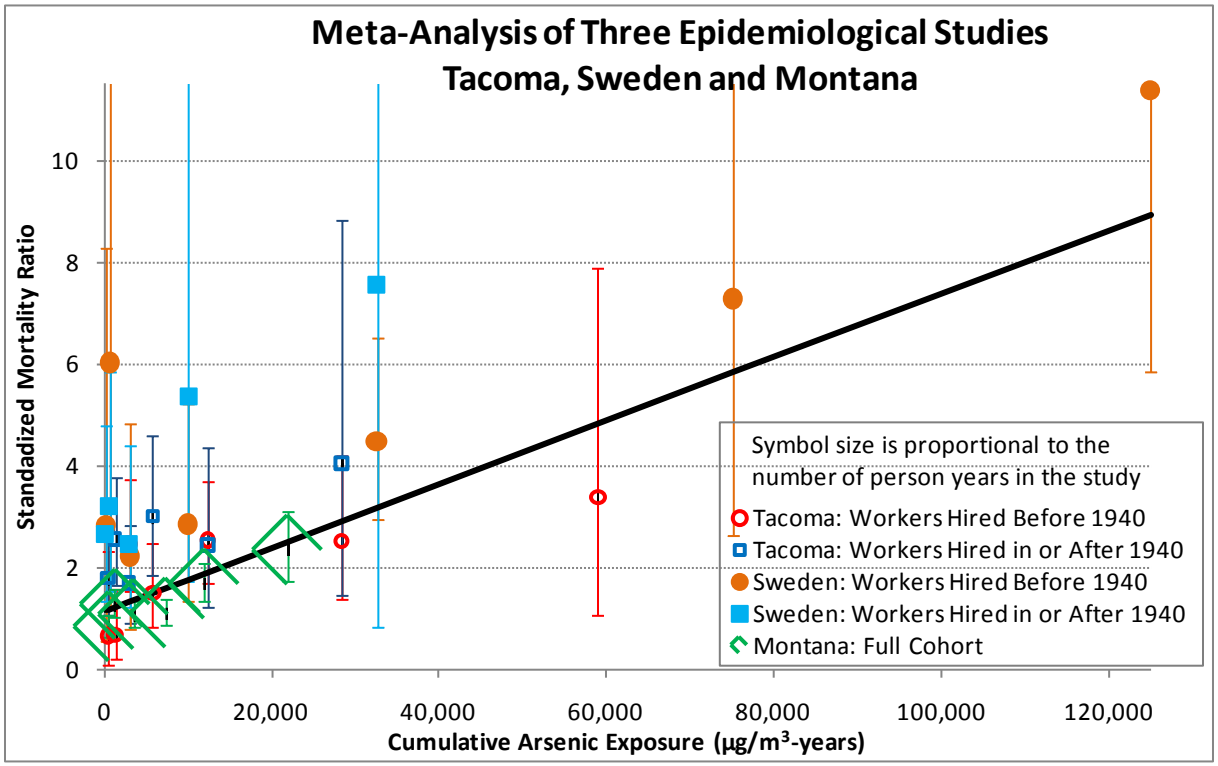
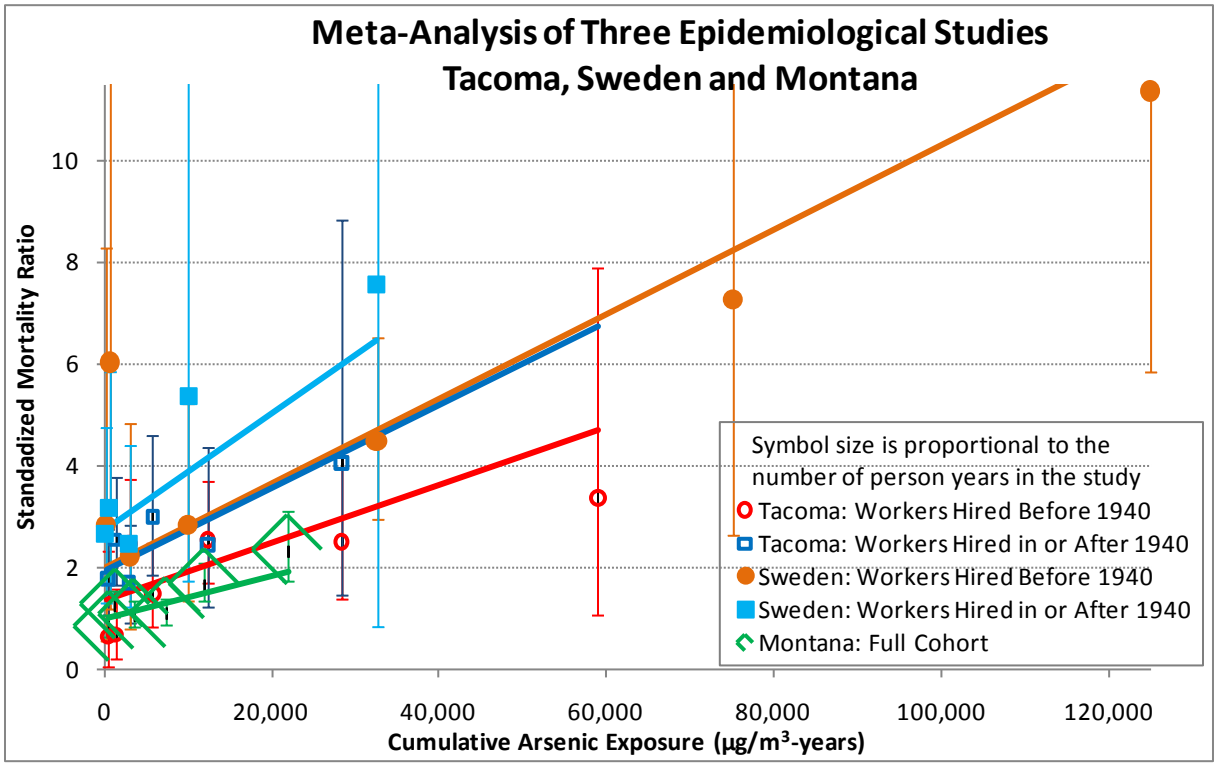


Figure B.8. Second model with five intercepts and one slope fit to the SMRs of the Tacoma, Montana and Sweden cohorts combined



The central estimate of the URF and the 95% UCL on the URF are calculated by applying the same methodology used in the calculations of the URFs and 95% UCLs on the URF for the individual studies and described before. Table B.9 shows the central estimate of the URF and the 95% UCL on the URF for the three meta-analyses shown in Table B.8.

Table B.9. Estimated URF and 95% UCL on the URF resulting from meta-analyses that combine the Tacoma, Montana and Sweden cohorts

Meta-analysis	URF per $\mu g/m^3$	
	MLE	95% UCL
using **inverse-variance weighting** to estimate a common intercept α and a common slope β from the intercepts and slopes of the individual cohorts	1.48E-04	2.04E-04
estimating **one intercept** and a single slope of the multiplicative rate ratio model fit to the combined data from the three cohorts	2.07E-04	2.69E-04
estimating **an intercept for each cohort** or sub-cohort and a single slope of the multiplicative rate ratio model fit to the combined data from the three cohorts	1.60E-04	2.19E-04

The range of the MLE URFs is 1.48×10^{-4} to 2.07×10^{-4} per $\mu g/m^3$ for the analyses shown in Table B.9. The URF and 95% UCL on the URF corresponding to the second model (last row in Table B.9) are more reliable than the estimates based on the first model because the second model fits the data statistically significantly better than the first model. In addition, the first model is rejected because of lack of fit whereas the second model is not rejected as a model fitting the data.

B.8 Summary

The original arsenic DSD report written by the TCQ included a meta-analysis that weighted individual URFs derived from each study separately using the number of person-years in the cohort. A panel of experts that reviewed the arsenic DSD suggested using inverse-variance weighting instead of person-years weighting. In addition, the panel suggested estimating the 95% UCL on the URF in addition to the central estimate of the URF.

The panel of experts additionally recommended excluding one study (UK study) because it used a weighted cumulative exposure to arsenic as the dose metric as opposed to the standard cumulative exposure. After excluding the UK study, the panel suggested performing a meta-analysis on the SMRs of the three studies (Tacoma, Montana and Sweden) that used the same cumulative exposure to arsenic dose metric. Though the panel proposed that a model with one intercept be fit to the combined data, a model with multiple intercepts was also fit to the combined data. The model with one intercept did not fit the data satisfactorily while the lack-of-fit test did not reject the model with multiple intercepts.

Based on the expert panel suggestions and the statistical analyses performed on the combined data, the model with multiple intercepts fit to the SMRs of the Tacoma, Montana and Sweden cohorts generate the most reliable and scientifically defensible estimates of the URFs.

References for Appendices A and B

Binks, K., R. Doll, M. Gillies, C. Holroyd, S. R. Jones, D. McGeoghegan, L. Scott, R. Wakeford, and P. Walker. 2005. Mortality experience of male workers at a UK tin smelter. *Occup Med (Lond)* 55 (3):215-26.

Brown, C. C., and K. C. Chu. 1983a. Implications of the multistage theory of carcinogenesis applied to occupational arsenic exposure. *J Natl Cancer Inst* 70 (3):455-63.

———. 1983b. A new method for the analysis of cohort studies: implications of the multistage theory of carcinogenesis applied to occupational arsenic exposure. *Environ Health Perspect* 50:293-308.

Crump and Allen (1985) Methods of Quantitative Risk Assessment Using Occupational Studies. The American Statistician, 39(4): 442-450.

Enterline, P. E., R. Day, and G. M. Marsh. 1995. Cancers related to exposure to arsenic at a copper smelter. *Occup Environ Med* 52 (1):28-32.

Enterline, P. E., V. L. Henderson, and G. M. Marsh. 1987. Exposure to arsenic and respiratory cancer. A reanalysis. *Am J Epidemiol* 125 (6):929-38.

Enterline, P. E., and G. M. Marsh. 1982. Cancer among workers exposed to arsenic and other substances in a copper smelter. *Am J Epidemiol* 116 (6):895-911.

Enterline, P. E., G. M. Marsh, N. A. Esmen, V. L. Henderson, C. M. Callahan, and M. Paik. 1987. Some effects of cigarette smoking, arsenic, and SO2 on mortality among US copper smelter workers. *J Occup Med* 29 (10):831-8.

Järup, L., and G. Pershagen. 1991. Arsenic exposure, smoking, and lung cancer in smelter workers--a case-control study. *Am J Epidemiol* 134 (6):545-51.

Järup, L., G. Pershagen, and S. Wall. 1989. Cumulative arsenic exposure and lung cancer in smelter workers: a dose-response study. *Am J Ind Med* 15 (1):31-41.

Jones, S. R., P. Atkin, C. Holroyd, E. Lutman, J. V. Batlle, R. Wakeford, and P. Walker. 2007. Lung cancer mortality at a UK tin smelter. *Occup Med (Lond)* 57 (4):238-45.

Integrated Risk Information System (IRIS). 2007. Arsenic. U.S. Environmental Protection Agency. http://www.epa.gov/iris/subst/index.html. August 28, 2007. Washington, D.C.

Lee-Feldstein, A. 1983. Arsenic and respiratory cancer in humans: follow-up of copper smelter employees in Montana. *J Natl Cancer Inst* 70 (4):601-10.

———. 1986. Cumulative exposure to arsenic and its relationship to respiratory cancer among copper smelter employees. *J Occup Med* 28 (4):296-302.

Lubin, J. H., and N. E. Caporaso. 2006. Cigarette smoking and lung cancer: modeling total exposure and intensity. *Cancer Epidemiol Biomarkers Prev* 15 (3):517-23.

Lubin, J. H., L. E. Moore, J. F. Fraumeni, Jr., and K. P. Cantor. 2008. Respiratory cancer and inhaled inorganic arsenic in copper smelters workers: a linear relationship with cumulative exposure that increases with concentration. *Environ Health Perspect* 116 (12):1661-5.

National Research Council (NRC). 1999. Committee on Health Risks of Exposure to Radon. Health Effects of Exposure to Radon: BEIR VI; Washington, D.C. National Academy Press.

Pinto, S. S., M. O. Varner, K. W. Nelson, A. L. Labbe, and L. D. White. 1976. Arsenic trioxide absorption and excretion in industry. *J Occup Med* 18 (10):677-80.

Sanchez-Meca, J. and F. Marin-Martinez. 1998. Weighting by inverse variance or by sample size in meta-analysis: a simulation study. *Educational and Psychological Measurement* 58: 211-20.

Texas Commission on Environmental Quality (TCEQ). 2006. Guidelines to develop effects screening levels, reference values, and unit risk factors. RG-442. Chief Engineer's Office. available from http://www.tceq.state.tx.us/comm_exec/forms_pubs/pubs/rg/rg-442.html

United States Environmental Protection Agency (USEPA). 1984. Health Assessment Document for Inorganic Arsenic Final Report. EPA-600/8-83-012F. Office of Health and Environmental Asssessment. Washington, D.C.

Viren, J., and A. Silvers. 1999. Nonlinearity in the lung cancer dose-response for airborne arsenic: apparent confounding by year of hire in evaluating lung cancer risks from arsenic exposure in Tacoma smelter workers. *Regul Toxicol Pharmacol* 30 (2 Pt 1):117-29.

Viren, J. R., and A. Silvers. 1994. Unit risk estimates for airborne arsenic exposure: an updated view based on recent data from two copper smelter cohorts. *Regul Toxicol Pharmacol* 20 (2):125-38.

Appendix C (Chapter 7): Glossary

Benchmark Dose (BMD) or Concentration (BMC): A dose or concentration that produces a predetermined change (called the benchmark response or BMR) in a specified response rate of an adverse effect compared to background.

Benchmark Response (BMR): A predetermined response rate change for an adverse effect, used to define a benchmark dose from which an RfD (or RfC) can be developed. For quantal responses (as opposed to continuous response) the change in response rate over background corresponding to the BMR is usually in the range of 5-10%, which is the limit of responses typically observed in well-conducted animal experiments.

Categorical variable: A variable that is restricted to a finite number of possible values. Categorical variables are usually names or labels such as gender, health status, and type of job. Categorical variables may also be labels for groups of discrete or continuous variable values. For example, body weight under 100 pounds might be labeled 1; body weight between 100 and 200 pounds might be labeled 2; and body weight over 200 lbs might be labeled 3; in which case this label would be a categorical variable.

Continuous variable: A variable that can take on any value between its minimum value and its maximum value. Continuous variables typically correspond to measurements. For example, body weight is a continuous variable if the weight can be measured to as many decimal points as desired and is not restricted to be a whole number of units.

Discrete variable: A variable that is restricted to a countable number of possible values. For example, if body weight is restricted to a whole number of pounds (or kilograms), then body weight is discrete.

Effective Dose (ED_{10}) or Effective Concentrations (EC_{10}): The dose or concentration corresponding to an increase of 0.10 in the probability of the occurrence of an adverse effect, an increase of 0.10 above the probability of the occurrence of an adverse effect among controls.

Exposure-Response Assessment: A determination of the relationship between exposure, which has elements of both magnitude and temporality, and a specific biological response. Response can be expressed as measured or observed incidence or change in level of response, percent response in groups of subjects (or populations), or the probability of occurrence or change in level of response within a population.

Exposure-Response Relationship: The relationship between a quantified exposure and the proportion of subjects demonstrating specific biologically significant changes in incidence and/or in degree of change (response).

Incidence: The number of new cases of a specified response that develop within a specified population over a specified period of time.

Linear Exposure-Response: A pattern of frequency or severity of biological response that varies directly with the amount of exposure of an agent. This linear relationship may hold (or be assumed to hold) only at low exposures in the range of extrapolation.

Model: A mathematical function with parameters that can be adjusted so the function closely describes a set of empirical data. A mechanistic model usually reflects observed or hypothesized biological or physical mechanisms, and has model parameters with real world interpretation. In contrast, statistical or empirical models selected for particular numerical properties are fitted to data, and model parameters may or may not have real world interpretation. When data quality is otherwise equivalent, extrapolation using mechanistic models (e.g., biologically based dose-response or exposure-response models) often carries higher confidence than extrapolation using empirical models (e.g., logistic model).

Multistage Model: A mathematical function used to extrapolate the probability of cancer from animal bioassay data, using the form…

$P(d) = 1 - \exp[-(q_0 + q_1 d + q_2 d^2 + ... + q_k d^k)$

where:

P(d) = the probability of a tumor (or other specified response) from lifetime continuous exposure at level d;

q_i = fitted model parameters, i=0, 1, ..., k;

k = usually restricted to be no greater than the number of dose (or exposure) levels -1.

Multistage-Weibull Model: A mathematical function used to extrapolate the probability of cancer from animal bioassay data, using the form

$$P(d,t) = 1 - e^{-(q_0 + q_1 d + q_2 d^2 + \ldots + q_k d^k)(t - t_0)^z}$$

Where: P(d,t) = the probability of a tumor (or other response) from lifetime, continuous exposure at dose d until age t (when tumor is fatal);

q_i = fitted dose parameters, i=0, 1, . . . , k;

k = no greater than the number of dose groups - 1;

t_0 = the time between when a potentially fatal tumor becomes observable and when it causes death; and

z = fitted time parameter (also called "Weibull" parameter).

Nonlinear Exposure-Response: A pattern of frequency or severity of biological response that does not vary directly with the amount of dose (exposure) of an agent. When MOA information indicates that responses may not follow a linear pattern below the dose (exposure) range of the observed data, nonlinear methods for determining risk at low dose (exposure) may be justified.

Odds: If the probability of a specified event is p, then the odds in favor of that specified event is p/(1-p).

Odds Ratio (OR): The odds of disease among exposed individuals divided by the odds of disease among unexposed.

Person-Years at Risk: The number of years that an individual is at risk of responding to exposure. In a cohort, usually the person-years at risk include the time since start of follow-up or start of employment in the plants being studied and the time that the individual was observed.

Point of Departure (POD): The dose (exposure) point that marks the beginning of a low-dose (or low-exposure) extrapolation. This point can be the lower bound on dose (exposure) for an estimated incidence or a change in response level from a dose-response model (BMD), or a NOAEL or LOAEL for an observed incidence, or change in level of response.

PPM-Years: Units of exposure (ppm) in the epidemiology study corresponding to inhaling 10 m^3 per day for 5 days a week. PPM-years of exposure in the inference situation might correspond to environmental exposure of 20 m^3 per day for 7 days per week. The conversion from an environmental concentration relevant to the general population of 1 ppm to an occupational exposure concentration used in the estimated dose-response would be:

(1 ppm) × (20 m^3/10 m^3) × (7 days/5 days)

Similarly, if the "slope" in an estimated occupational dose-response model is β per ppm-day, then that slope can be converted to units of ppm-years as follows:

β × 365 per ppm-year.

Random Variable: A function that associates a unique numerical value with every outcome of an experiment.

Relative Risk (RR): The relative risk ratio, or more simply the relative risk, is the probability of a specified event occurring in the exposed group divided by the probability of a specified event occurring in the non-exposed group.

Risk Assessment (in the context of human health): The evaluation of scientific information on the hazardous properties of environmental agents (hazard characterization), the exposure-response relationship (exposure-response assessment), and the extent of human exposure to those agents (exposure assessment). The product of the risk assessment is a statement regarding the probability that populations or individuals so exposed will be harmed and to what degree (risk characterization).

Risk Characterization (in the context of human health): The integration of information on hazard, exposure, and exposure-response to provide an estimate of the likelihood that any of the identified adverse effects will occur in exposed people.

Standardized Mortality Ratio (SMR): The ratio of observed deaths in a study population to the expected number of deaths calculated for a specified standard population usually comparable to the population being observed.