



Abt Associates Inc.

**Alternatives to
Pollutant-by-Pollutant
Dose-Response
Estimation for Air
Toxics**

Presented at
HAP Workshop
April 30 - May 1, 2009

EP-W-05-022 WA 3-80
Task 2 White Paper

March 2010

Submitted to:
Ronn Dexter,
Work Assignment Manager
U.S. Environmental Protection
Agency
National Center for Environmental
Economics
1200 Pennsylvania Ave., NW
Washington, D.C. 20460

Submitted by:
Dale Hattis, Ph.D.
Meghan Keaney Lynch, Sc.D.

Abt Associates Inc.
55 Wheeler Ave
Cambridge, MA 02138
(617) 520-2442

Contents

1	Introduction.....	1
2	Assessing Risks for Traditional Toxic Effects for Chemicals With Incomplete Data—A Generic Approach.....	4
2.1	Assessing Uncertainties by Making Analogies with Previously Studied Cases	5
2.2	Developing Distributions to Represent Each of the Concerns Reflected in Traditional Uncertainty Factors	6
2.3	Incorporating an Assumption of Interacting Background Processes.....	12
2.4	Concerns from the Use of Empirically Derived Data Distributions and Other Aspects of the Estimation of Low-Dose Risks	15
2.5	Important Details in Projecting Low-Dose Risks and Benefits of Exposure Reductions	18
3	The Promise of Analyses of Specific Resources Available for Developing Potency Distributions for Air Toxics	28
3.1	RD ₅₀ Determinations by Alarie.....	28
3.2	Short-Term LC ₅₀ Concentration/Time Tradeoff Data Assembled by ten Berge	33
3.3	Prospects for a Compilation of LC ₅₀ s and Analyses of LC ₅₀ /LD ₅₀ Correlations	35
3.4	Potencies and Interindividual Variability/Susceptibility Distributions for Bronchoconstricting Agents—Guidance for a Common Mode of Action for Several Air Toxics—Acting in Asthmatics.....	35
3.5	Actions of and Human Susceptibility Distributions for Ozone—Another Basis for Analogies with Oxidant Chemicals?.....	39
3.6	Preliminary Categorization and Grouping of the Officially Designated Set of Air Toxics by a Priori Recognition of Basic Chemical Structural Alerts or Prominent Concerns for Carcinogenesis.....	40
4	The Promise of Early-Effect Biomarkers for Two-Step Assessments of Risks.....	46
4.1	Caveats to the Biomarker Approach	47
4.2	Application of the Biomarker Approach to HAPs	49
5	Conclusions.....	51
6	References.....	52
Appendix A – Birth Weight in Humans and Fetal Growth Reduction in Animals—Model Biomarkers for Assessing Generic Reproductive Toxicity in Animals and People..... A-1		
	Birth Weight and Infant Mortality	A-1
	The Barker (2001) Observations and a Hypothesis—Latent Effects Appearing in Later Life From Early-Life Reductions in the Reserve Capacity To Accomplish Homeostatic Control	A-11
	References - Appendix A.....	A-12
Appendix B – Case Study: Competitive Inhibition of Iodide Uptake by Perchlorate B-1		
	References - Appendix B	B-3

Tables

Table 2-1: Distributions ^a of Human Toxic Potency Relative to Animal Toxic Potency per Unit (Body Weight) ^{3/4}	8
Table 2-2: Distribution of Dose Multipliers Required to Estimate Uncertainty Associated with Database Deficiency	11
Table 2-3: Results of “Straw Man” Monte Carlo Simulations on Estimates of <u>Arithmetic Mean “Expected Value”</u> Risks of Mild Adverse Effects from 2,4,6-Trinitrotoluene for Chronic Human Dosage in the Neighborhood of the Current IRIS RfD	14
Table 2-4: Results of “Straw Man” Monte Carlo Simulations on Estimates of <u>95th Percentile</u> Risks of Mild Adverse Effects from 2,4,6-Trinitrotoluene for Chronic Human Dosage in the Neighborhood of the Current IRIS RfD	14
Table 2-5: Results of “Straw Man” Monte Carlo Simulations on Estimates of <u>Median (50th Percentile)</u> Risks of Mild Adverse Effects from 2,4,6-Trinitrotoluene for Chronic Human Dosage in the Neighborhood of the Current IRIS RfD	14
Table 2-6: Summary Statistics and Selected Percentiles of the Distributions of NOELs and LOELs (mg/kg-day)	21
Table 2-7: Summary Statistics and Selected Percentiles of the Distributions of Geometric Mean LD ₅₀ s for Chemicals Studied in More Than One Species.....	23
Table 2-8: Summary Statistics for Distributions of LOELs and NOELs for Chemicals Sorted into Cramer Classes	25
Table 3-1: Variance of RD ₅₀ s Explained by the Physical versus Chemical Reaction Categories of Alarie et al. (1998)	31
Table 3-2: Chemical Structural Categories Used By Alarie et al. (1998)	33
Table 3-3: Exponents of C ^h T Dose Metrics from log Probit Fits to Inhalation Lethality Data for a Variety of Chemicals	34
Table 3-4: Observations of Human Interindividual Variability in External Concentrations of Methacholine Needed to Produce Defined Short-Term Changes in Respiratory Parameters	36
Table 3-5: Famous Carcinogens	41
Table 3-6: Chemicals That Need Evaluation for Possible Carcinogenic Effects	42
Table 3-7: Chemicals Associated with Bladder Carcinogenesis	42
Table 3-8: Other Generic Large Categories	43
Table 3-9: Smaller Categories of Organic Compounds	44
Table 3-10: Inorganic Categories	45
Table 4-1: Basic Birth Weight Reduction Results Based on County-Average Air Pollutant Exposures During Gestation*	50
Table 4-2: Implications for Population Aggregate Birth Weight Changes*	50

Figures

Figure 1-1: Interpretation of Dose-Response Information for Quantal Effects in Terms of a Lognormal Distribution of Individual Threshold Doses.....	2
Figure 2-1: Lognormal Plot of the Ratio of the Reduced NOAEL Expected from the Addition of a Rat Reproductive/Developmental Study to a Base of a Rat Chronic Toxicity Study, Based on Evans and Baird (1998) Data for 35 Pesticides.....	11
Figure 2-2: Lognormal Plots of the Distributions of LOELs and NOELs Compiled by Munro et al. (1996) for All Chemical Structural Classes Combined	21
Figure 2-3: Lognormal Plot of the Distribution of Mean Log(LD ₅₀) for all 5,128 Chemicals with Data for Two or More Species in the Database of Rhomberg and Wolff (1998)	23
Figure 2-4: Lognormal Plots of LOEL Distributions Compiled and Sorted by Munro et al. (1996) Sorted into Cramer Classes	26
Figure 2-5: Lognormal Plots of NOEL Distributions Compiled and Sorted by Munro et al. (1996) into Cramer Classes	27
Figure 3-1: Lognormal Plot of the Distribution of Mean 10-Minute log(RD _{50s})	29
Figure 3-2: Lognormal Plot of 145 RD ₅₀ Values Compiled by Alarie et al. (1998).....	30
Figure 3-3: Comparison of Lognormal Plots of RD ₅₀ Distributions for Chemicals Classified by Alarie et al. (1998) as Likely to Have Physical versus Chemical Reaction Mechanisms Determine Their RD _{50s}	31
Figure 3-4: Lognormal Probability Plots of the Distribution of Methacholine Concentrations Needed To Cause a 20 Percent Reduction in FEV1 in Different Populations	37
Figure 3-5: Lognormal Probability Plots of the Distributions of SO ₂ Concentrations Needed to Cause a 15 or 20 Percent Reduction in FEV1 in Exercising Asthmatics	38
Figure 3-6: Lognormal Probability Plots of the Distributions of SO ₂ Concentrations Needed to Cause a 100 or 200 Percent Increase in Specific Airway Resistance in Exercising Asthmatics	39
Figure 3-7: Composite log Probit Plot of the Ozone Dose-Time-Severity-of-Response Relationship	40
Figure 4-1: Two-Step Risk Assessment Process Utilizing Biomarkers	46
Figure 4-2: Relationships Between Reported Cigarettes Smoked per Day, Average Birth Weight, and Infant Mortality.....	49
Figure A-1: Loss in Fetal Weight in CD-1 Mice Exposed to EGEE via Oral Gavage (Data of Wier et al. 1987).....	A-2
Figure A-2: Loss in Fetal Weight in New England White Rabbits Exposed to EGME via Inhalation	A-2
Figure A-3: Fetal Weight Response of Rats to Trichloroacetic Acid.....	A-4
Figure A-4: Fetal Weight Response of Rats to Dichloroacetic Acid.....	A-5
Figure A-5: Fetal Weight Response of Rats to Monobromoacetic Acid.....	A-5
Figure A-6: Fetal Weight Response of Rats to Trichloroacetonitrile.....	A-6
Figure A-7: Fetal Weight Response of Rats to Dichloroacetylnitrile.....	A-6
Figure A-8: Fetal Weight Response of Rats to Monobromoacetic Acid.....	A-7
Figure A-9: Results of Regression Analysis of the Fraction of Control Fetal Weight Response in Grouped Categories of TCA Equivalents—Interpretation with a Linear Model	A-8
Figure A-10: Results of Regression Analysis of the Fraction of Control Fetal Weight Response in Grouped Categories of TCA Equivalents—Interpretation with a Quadratic Model	A-9
Figure A-11: Results of Regression Analysis of the Fraction of Control Fetal Weight Response in Grouped Categories of TCA Equivalents—Interpretation with a Third-Degree Polynomial ...	A-10
Figure A-12: Relationship Between Weight at Birth (in 500 g Increments) and Infant Mortality ..	A-10

Figure A-13: Plot of the Incidence of Type 2 Diabetes in Relation to log(Mean Birth Weight)..... A-12
Figure B-1: Inferences of the Population Change in Median IQ in Relation to Spot Urinary Iodide
Measurements at Age 9..... B-2

Acronyms

DART	Developmental and Reproductive Toxicology
ED ₅₀	A dose that causes an effect in 50 percent of the test subjects
EGEE	Ethylene glycol ethyl ether
FAO	Food and Agriculture Organization
FEV1	Forced expiratory volume in 1 second
GSD	Geometric standard deviation
HAP	Hazardous air pollutant
IRIS	Integrated Risk Information System
JECFA	Joint FAO/WHO Expert Committee on Food Additives
LC	Lethal concentration
LC ₅₀	The concentration of a material in air that will kill 50 percent of the test subjects
LD	Lethal dose
LOAEL	Lowest observed adverse effect level
LOEL	Lowest observed effect level
MGEE	Methylene glycol ethyl ether
MTD	Maximum Tolerated Dose
NOAEL	No observed adverse effect level
NOEL	No observed effect level
NTP	National Toxicology Program
NRC	National Research Council
PBPK	Physiologically based pharmacokinetic
POD	Point of Departure
POM	Polycyclic Organic Matter
RfC	Reference concentration
RfD	Reference dose
RD ₅₀	Exposure concentration producing a 50 percent respiratory rate decrease
REL	Reference exposure level
RTECS	Registry of Toxic Effects of Chemical Substances
SO ₂	Sulfur dioxide
TCA	Trichloroacetic acid
WHO	World Health Organization

Executive Summary

Perceived needs for extensive chemical-specific toxicological information have impeded efforts to assess risks and evaluate likely public health protection benefits of possible standards for hazardous air pollutants (HAPs). This paper explores two approaches that regulatory toxicologists can use to analyze risks and associated uncertainties for noncancer effects of HAPs with limited toxicological databases. Both approaches show new ways to leverage human experimental and epidemiological data to draw inferences about human risks.

The first approach supplements available chemical-specific information by viewing specific HAPs as random draws from reference sets of putatively analogous chemicals that have been studied in the past, the “Straw Man” model. The variability among previously studied chemicals is used to create distributions that represent each of the concerns that are addressed by traditional “uncertainty factors” in the derivation of reference concentrations (RfCs) (Hattis and Lynch 2007; Hattis et al. 2002). The present paper adds to this previous work the capability to incorporate the expected implications for risks of a user-specified incidence of effects from interacting background processes, as recommended in a recent report by the National Research Council (2008). We also review several sources of data that can provide estimates of basic toxic potency (and associated uncertainties) in the light of chemical structural considerations and other information for pollutants where results of even the most basic toxicological tests are not available. The resulting uncertainty distributions for expected effects as a function of exposure can be used together with HAP-specific exposure information to judge the likely value of additional chemical-specific information for future regulatory decision-making.

Another path toward risk assessments for HAPs is a growing set of human biomarkers of early effect that allow assessors to “move upstream” from ultimate endpoints of concern (Woodruff et al. 2008). Using these biomarkers, projections of potential health risks are made in two steps: (1) assess relationships between exposure to the chemical of interest and the intermediate biomarker of effect using limited chemical-specific information, and (2) assess relationships between the biomarker of effect and ultimate endpoints of concern from general epidemiological data. Several candidate biomarkers are available for this approach, but on the basis of recent epidemiological findings, one that appears particularly promising for ambient air pollutants is fetal growth restriction, which produces changes in distributions of birth weights. There is strong evidence that incremental changes in birth weights, even within the normal range seen in the general population, can indicate developmental impairments with implications for infant mortality and a number of other long-term effects in later life. Estimates of relative potency for fetal growth restriction for some HAPs may be possible on the basis of animal experiments. Projected human effects might then be derived via comparisons with potencies of some criteria pollutants that have been the subjects of relatively extensive human epidemiological observations.

1 Introduction

There are a substantial number of chemicals in commerce for which no traditional toxicological studies are available. In the standard system of chemical risk assessment, these chemicals are unfortunately treated as having no toxic potency at all, and it is assumed that no benefits can be estimated for reducing human exposures to them.

We believe that regulatory toxicologists can draw upon a much wider range of information resources than has been customary in assessing public health hazards. Chemical-specific data can be usefully supplemented both in assessments of individual chemical exposures, and also in assessments of complex and variable mixtures that, on their surface, present mind-numbing possibilities for complex combinations of individual toxicants (e.g., “disinfection byproducts,” “diesel particulates,” “coke oven emissions,” “dioxins and other agonists of the Ah receptor,” “inhibitors of thyroid uptake of iodide”). This supplementation is especially important in the case of the approximately 188 HAPs,¹ which require evaluation but often lack sufficient epidemiology or toxicology data.

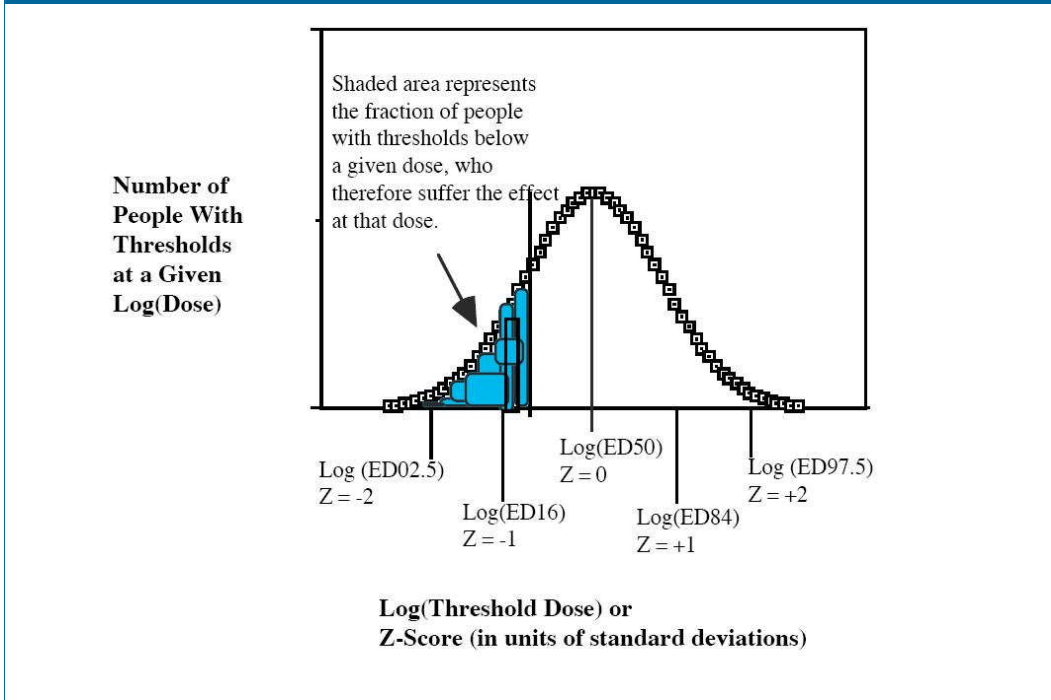
As noted in a recent report by the National Research Council (NRC 2008), the traditional analytical approach of RfCs and reference doses (RfDs) does not provide quantitative estimates of risk that can be used in comparative analyses of risks and benefits for different policy options. Based on some of our earlier work (Hattis et al. 2002; Hattis and Lynch 2007), we believe it is possible to produce such estimates without significantly more chemical-specific information than is now typically available for chemicals evaluated in the Integrated Risk Information System (IRIS) system. In this “Straw Man” approach, each of the traditional “uncertainty factors” is replaced with a distribution based on empirical observations of how large each of the factors proved to be for a putatively analogous reference set of chemicals. Finite risks are projected from an assumption that human thresholds for effects tend to be lognormally distributed² in the population (see Figure 1-1), although other or modified lognormal distributions could just as easily be used where indicated. **Section 2** provides an overview of this approach, associated uncertainties, and prospects for improvement via incorporation of a user-specified background interaction option as called for by NRC (2008). Incorporation of this adaptation essentially merges the NRC’s conceptual models 1 and 2—individual threshold dose-response models with and without an appreciable interaction with background pathological processes. The discussion in Section 2 draws from our previous work and includes additional information from an extensive database of potencies for traditional toxicological endpoints that was compiled by Munro et al. (1996). This database focuses primarily on oral exposures and thus the RfD.

¹ There are said to have been 188 original HAPs, published in 1996 (“This administrative regulation provides the list of hazardous air pollutants pursuant to 42 U.S.C. 7412(b) as amended in the *Federal Register*, 61 FR 30823, June 18, 1996 and the list of source categories and subcategories, as published in the *Federal Register*, 57 FR 31591, July 16, 1992.”) from which three were delisted: hydrogen sulfide, methyl ethyl ketone, and caprolactam were removed from Section 112(b) in 1991, 2005, and 1996 respectively. A recent “official list” of the HAPs retrieved from <http://www.epa.gov/ttn/atw/188polls.htm> has 190 by our count. The discrepancy may be partly related to the listing of three separate xylene isomers in addition to “xylenes (isomers and mixture) in the latter source.

² That is, the logarithms of individual human thresholds for effect are expected to have a normal Gaussian distribution. Such a distribution would be expected if there are many factors that cause humans to vary in their individual thresholds, and these factors tend to exert their influences multiplicatively. Although this assumption has considerable uncertainty as applied to individual cases, available data do not indicate that this assumption is appreciably biased in general.

Section 3 then reviews specific information resources that are available to help comparably evaluate potencies and risks for some endpoints for inhaled toxicants in the list of “hazardous air pollutants” defined in the current Clean Air Act. This discussion includes available databases for RD₅₀ (exposure concentration producing a 50 percent respiratory rate decrease) and LC₅₀ (the concentration of a material in air that will kill 50 percent of the test subjects) endpoints more relevant to the inhalation pathway. Preliminary databases are constructed and analyzed to demonstrate the feasibility of using the information to evaluate HAPs in conjunction with the Straw Man or alternative models.

Figure 1-1: Interpretation of Dose-Response Information for Quantal Effects in Terms of a Lognormal Distribution of Individual Threshold Doses



In addition to these “uncertainty factor” (or, better, “adjustment factor”) distributions, there is another important source of information that has become increasingly available in recent years from ongoing work by epidemiologists, but has not yet been tapped for use in EPA risk analyses. These human biomarkers of early effect, described in **Section 4**, allow assessors to “move upstream” from ultimate endpoints of concern (Woodruff et al. 2008). These biomarkers can allow assessors to project potential health risks in two steps:

- (1) Assess relationships between exposure to the chemical of interest and the intermediate biomarker of effect using limited chemical-specific information
- (2) Assess relationships between the biomarker of effect and ultimate endpoints of concern from general epidemiological data.

Generally, the biomarkers that are potentially useful for this two-step approach are continuous parameters (e.g., birth weight, sperm counts, lung-function measurements). The effects of ultimate

concern are often quantal parameters that are more difficult to measure directly in relation to toxicant exposures in epidemiological studies. The continuous nature of the biomarkers facilitates detection and quantification of the effects of different exposures. On the other hand, the continuous, “upstream,” early-effect biomarkers often have reasonably strong predictive relationships with the harder-to-measure quantal health effects of ultimate concern, which are easier for economists to work with in valuation studies.

A key observation that arises from the biomarker/health effects literature is that biomarkers typically do not need to exceed critical “thresholds” in order to have appreciable significance for predictions of public health risks. Most often, as in the case of birth weights or blood pressures,³ risks rise in a more or less continuous fashion in relation to biomarker levels. Thus, the health significance of biomarker changes is not usually confined to a small minority of people who are pushed from one side to another of quantitative criteria levels that physicians have defined in order to identify clinical disease states. The public health impacts generally extend through a large fraction of the population of ordinary people who would not normally be categorized as candidates for medical intervention, at least based on a single risk factor value by itself.

When epidemiological studies of early-effect biomarkers reach a critical mass so that comparative potency evaluations are possible for many exposures, as now appears to be the case for birth weight/fetal growth restriction effects of different air pollutants (Šrám et al. 2005; Bell et al. 2007), it is possible that wide-ranging comparisons of cost effectiveness will be possible for policy options that could reduce population exposures to different toxicants. Some of these possibilities are explored in **Section 4**.

Finally, **Section 5** offers a few conclusions about near-term prospects for helpful risk evaluations using the tools and resources described in the previous sections. Briefly, we believe it is possible to produce initial quantitative estimates of exposure-response relationships for major types of noncancer effects for the majority of HAPs with a reasonable level of effort over a few years. These estimates will typically have uncertainties that span several orders of magnitude, and will depend on numerous assumptions and analogies. However, as a group they will provide a consistent set of comparative evaluations that, together with exposure information, can be used to set priorities for both initial efforts at exposure reductions and efforts to obtain improved toxicity information for HAP control based on the greatest potential for both economic and health impacts.

³ See the birth weight example in Section 4; and also Kannel et al. (2003) for the graded increase in risks of cardiovascular disease with increasing blood pressures.

2 Assessing Risks for Traditional Toxic Effects for Chemicals With Incomplete Data—A Generic Approach

The focus in this paper is on noncancer effects.⁴ The traditional system for assessing risks from noncancer effects dates back to the foundational work of Lehman and Fitzhugh (1954). Originally, these Food and Drug Administration scientists recommended a rule of thumb to compensate for two sources of uncertainty in projecting “safe” levels for human daily intakes from animal toxicological observations: Take the “no effect level” (NOEL) indicated in studies of limited numbers of animals and divide it by 100 to arrive at an estimate of a dose that could be expected to be safe for people. Later, the 100-fold factor was decomposed into two factors of 10—one to account for animal-to-human extrapolation, and one to account for human interindividual variability.

Over time, three additional sources of uncertainty were added to these original two, each of which was represented by a single-point factor of 10 (or sometimes 3) when the toxicological database was deemed insufficient to resolve the issues in going from the Point of Departure (POD) such as a NOEL without further adjustment:

- Projection from an observed “low effect level” (LOEL) to a NOEL (or, more recently, a “no adverse effect level” (NOAEL)) when the principal toxicological study available did not provide a NOEL (or a NOAEL).
- Projection from an available subchronic to a chronic study.
- Projection from an “incomplete database” (lacking the full complement of studies such as chronic toxicity, reproductive, and developmental studies) to a database having all these component investigations.

One basic concept that lies at the heart of this analysis has not changed from the time of Lehman and Fitzhugh. This is the idea that many toxic effects result from placing a chemically induced stress on an organism that exceeds some homeostatic buffering capacity and/or exhausts available functional reserve capacity.⁵ Where it is applicable, the basic homeostatic-system-overwhelming model properly leads to an expectation that there should be individual thresholds for such effects. An individual person will show a particular response (or a response at a specific level of severity) only when his or her individual threshold exposure level for the chemical in question has been exceeded. However, this expectation of individual thresholds for response does not mean that we can necessarily specify a

⁴ A national analysis of exposures to HAPs in relation to cancer risks is currently in press (McCarthy et al. 2009).

⁵ Other types of mechanisms do exist, however, such as an irreversible accumulating damage model (e.g., for chronic neurological degenerative conditions) or a “risk factor” model (e.g., for cardiovascular diseases). In this context a “risk factor” is a parameter that is found in prospective epidemiological studies to be directly predictive for the later risk of adverse cardiovascular events. Many of these are continuous parameters, such as blood pressure or cholesterol levels, which can take on a very large (theoretically infinite) gradation of values. This is in contrast to a “quantal parameter,” which can only take on values of yes or no (e.g., whether or not a particular patient has had a heart attack or whether or not an infant has died before its first birthday), or a “discrete parameter” that can take on only a finite number of specific values (e.g., days of the week). As mentioned earlier (Kannel et al. 2003), these continuous risk factors often have continuously graded influences on individual risks of adverse effects. For the greater statistical power of continuous variables for prediction of risks, see Hattis (1998) and Section 4 for further discussion of the relationship between birth weight and infant mortality.

level of exposure that poses zero risk for a diverse population. In a large group of exposed people with differing homeostatic buffering capacities and different pre-existing pathologies, there may be people for whom a marginal perturbation of a key physiological process is sufficient to make the difference between barely adequate and inadequate function to avoid an adverse response, or even to sustain life. Such situations include:

- Some individuals in the diverse population are already suffering from various kinds of pathological dysfunction in key parameters that may be marginally affected by different toxicants (e.g., a person undergoing a myocardial infarction may have a marginally expanded area of heart muscle death if the oxygen-carrying capacity of his or her blood is reduced by a marginal exposure to additional carbon monoxide).⁶
- Some individuals are presently engaged in a task (e.g., running a 100-yard dash or learning to read) that taxes some physiological capabilities to their limit, and marginal exposures to a toxicant marginally reduce those physiological capabilities.

In **Section 2.1** and **Section 2.2**, we provide background on our previous efforts in constructing the Straw Man model. In light of the considerations highlighted above, and the emphasis of the recent NRC report on interactions with background processes, **Section 2.3** describes an approach for building in an assumption of a finite interacting background incidence of effects that would otherwise have a highly nonlinear dose-response relationship at very low doses. The background incidence used for this assumption converts the population dose-response relationship to one that is linear at low doses. The low-dose slope of this relationship is greatly influenced by the magnitude of the interacting background. **Section 2.4** discusses some uncertainties and sources of bias in the model, and suggests ways in which these could be improved upon in future work. **Section 2.5** provides a discussion of important details necessary for applying the models to specific cases, including cases where there are known genetic polymorphisms likely to have significant effects and cases where there is not a chemical-specific primary chronic toxicity study that can be used to assess potency.

2.1 Assessing Uncertainties by Making Analogies with Previously Studied Cases

The basic approach for deriving distributions to replace the traditional uncertainty factors is to view each chemical as a random draw from the universe of other chemicals that might share the same or similar types of toxic properties. We of course do not have information for the universe of all chemicals. However, as our knowledge grows we can hope to accumulate more and more substantial numbers of observations from that universe.

If we can make the provisional assumption that our accumulated data from previously studied chemicals reasonably represent the universe of all chemicals that might be considered for similar risk evaluation, then we can represent our *uncertainty* about the properties of a particular chemical as approximately the same as the observable *variability* among previously studied chemicals for each of the concerns represented by traditional safety/uncertainty factors.

As the database grows, it is possible to do better than a simple random draw from all the chemicals with information relevant to a particular adjustment factor. We can perform regression analyses and

⁶ Recent epidemiological studies have detected a short-term elevation of general and cardiovascular mortality rates in relation to marginal increases in carbon monoxide exposure (Samoli et al. 2007).

mechanistic studies to identify specific properties of the chemicals for which we have data that are strongly related to the size of the factor being modeled. For example, in our analyses of human interindividual variability we have found that variability tends to be greater for immune system endpoints than for other types of systemic toxicity; and smaller for the most severe types of endpoints (e.g., death or serious irreversible injury) than for less severe endpoints (e.g., headache, mild irritation) (Hattis and Lynch 2007).⁷ Using this information, the “random draws” in our system for uncertainty in the extent of interindividual variability are made from different subsets of the chemicals and toxicological responses whose measurements of variability we have compiled, depending on the nature and severity of the response being assessed.

2.2 Developing Distributions to Represent Each of the Concerns Reflected in Traditional Uncertainty Factors

As described previously (Hattis et al. 2002; Hattis and Lynch 2007), the goal of the Straw Man model is to predict risks or doses associated with a defined upper confidence limit on risk, rather than levels of exposure postulated to pose either no risk or some level of “acceptable” risk. Because of this, when original experimental data are available, the Straw Man model uses as its point of departure ED₅₀s (doses that cause an effect in 50 percent of the test subjects) estimated from those data. Uncertainty distributions are then used in the risk projections rather than the traditional uncertainty factors.

In the Monte Carlo simulations used to derive estimated risks (and uncertainties in those estimates), human interindividual variability plays a key role. It provides the means for projection from estimated human chronic ED₅₀ values for adverse effects to much lower incidences of effects expected at lower doses. As illustrated in Figure 1-1 above, the basic calculations assume that individual thresholds for effect are lognormally distributed in the human population. Data from pharmacokinetic and pharmacodynamic variability distributions are combined to arrive at an overall standard deviation for human variability in log space. Uncertainties arising from this assumption are discussed further below.

Animal ED₅₀s are the preferred point of departure for projection of human ED₅₀s because with any other choice the risk projections are complicated by the differences between human and animal variability distributions. The interindividual variability in susceptibility (and hence the breadth of the distribution of individual thresholds) in experimental test animals is likely to be much smaller than would be seen in human populations. This is because test animals are often relatively uniform genetically and are deliberately raised and exposed to toxicants under conditions that are as uniform

⁷ The Straw Man analysis sorted endpoints for human effects into three crude levels of severity (Hattis et al. 2002): *severe effects* (death or permanent damage, including the severe ataxia produced by methyl mercury poisoning in an early 1970s incident in Iraq); *moderate reversible or irreversible effects* including elevated beta-2 macroglobulin urinary excretion (an indicator of kidney toxicity) seen in some populations occupationally and environmentally exposed to cadmium; dose-limiting gastrointestinal toxicity from an anti-cancer agent; and *mild reversible effects* including eye and throat irritation, sedation/drowsiness from anesthetics, and various levels of olfactory perception. In the latest analysis, for median chemicals the pharmacodynamic component of variability was associated with a lognormal distribution with geometric standard deviation of about 2.1 for mild and moderate effects, in comparison with about 1.7 for severe/irreversible effects. It should be stressed that in each case there was substantial variation among chemicals in the extent of human interindividual variability attributable to the pharmacodynamic components of susceptibility.

as possible with respect to nutrition, exposures to other compounds, and age at exposure, among other factors.

For illustrative purposes, the initial Straw Man publications have used a risk-specific dose that would generate a 1/100,000 excess risk of mildly adverse outcome in the general population or 1/1,000 in sensitive subpopulations and a confidence level of 95 percent for comparisons with existing RfDs. Choosing a set of incidence and confidence levels for replacing existing RfDs is a risk management decision to be made by an appropriate governing body.

The Straw Man model estimates risk-specific doses by drawing on information from other chemicals that is relevant to each of the traditional uncertainty factors used in traditional noncancer risk assessments. As with traditional single-point uncertainty factors, the uncertainty distributions depend on the nature of the chemical-specific toxicity information available for the chemical under study and therefore the data deficiencies that need to be offset by supplemental information from other sources—e.g., subchronic/chronic, animal/human, incomplete/complete data set for risk projection, and human interindividual variability for pharmacokinetic and pharmacodynamic steps in the causal pathway to production of adverse effects.

The following subsections describe our previous efforts to assemble several databases that can be used to replace the traditional point estimates applied as uncertainty factors. The focus during the construction of these databases was on the RfD and oral exposures. In **Section 3** we describe additional data sources that could be incorporated into the Straw Man model to reflect the inhalation exposures more common to HAPs.

2.2.1 Interspecies

Our previous work (Hattis et al. 2002) used a database generated by Price et al. (2008) to conduct an empirical distributional analysis. This database consisted of observations of human Maximum Tolerated Doses (MTDs) of anti-cancer agents that were compared to putatively equivalent LD₁₀s in mice, rats, and hamsters, and minimal values at which toxicity was observed in dogs and monkeys. The Price et al. database includes entries for 64 compounds, which is more than twice as many as those of earlier efforts by Watanabe et al. (1992) and Travis and White (1988). We evaluated these data for central tendency and a spread of the animal/human dose conversion factors for varying types of interspecies information. Specifically, we evaluated how much equivalent dose conversions might be different (a) when the “critical study” for calculating the RfD comes from different specific animal species (e.g., rats versus dogs), and (b) when data are available for multiple species (Hattis et al. 2002). This second aspect of the analysis allowed us to correct for the extra “conservatism” in calculations based on the choice of the most sensitive of the tested species as the basis for projection, when data for more than one species were available.

The distributional animal-to-human equivalent dose projections were made by Hattis et al. (2002) in three steps:

1. The doses for the critical effect were adjusted from mg/kg body weight to mg/(kg body weight)^{3/4} since this measurement tends to scale with (body weight)^{3/4} so as to reflect the pharmacokinetic elimination dependence on the metabolic rate, which tends to scale with (body weight)^{3/4} (Mordenti et al. 1986; Travis et al. 1990; Ad Hoc Working Group on Risk Assessment et al. 1992). The standardized body weights used in this scaling were from the Registry of Toxic Effects of Chemical Substances (NIOSH, 1982).

2. An additional adjustment factor was applied to reflect the median human MTD expected based on the identity and number of species that provided data that potentially could have been used as the basis for the RfD (Hattis et al. 2002).
3. The uncertainty in each type of animal-to-human potency projection was identified from the variability, as lognormal distributions, in the ratios of the observed human potencies to the animal-projected potencies for different chemicals.

Table 2-1 provides the geometric mean and geometric standard deviation (GSD) of the human/animal potency adjustment factors indicated for cases where the “critical” RfD-determining study comes from specific species and where particular other species also contributed potentially usable data. A value of 1 in the third column of the table means that for the typical chemical the center of the uncertainty distribution for the human potency is the same as that observed for the most sensitive animal species when potency is expressed as mg per (kg body weight)^{3/4}. Proceeding down the rows, it can be seen that the central values in the third column tend to decline below 1, as the most sensitive of the two or more species with available data has an increasing tendency to be more sensitive than humans. So the human potency, on average, will be less than the sensitivity of the most sensitive animal. This tends to offset the distortion of estimates of human potency caused by the presence of more species with relevant data for some chemicals compared to others. At the same time, for cases where data are available for more species, the spread of the uncertainty in the human potency is less, as indicated by smaller log(GSD)⁸ estimates in the final column of Table 2-1.

Species LD ₁₀ or MTD Information Used	Number of Chemicals	Geom. Mean	Arith. Mean	95th %tile	log(GSD)
Mouse (single species) ^a	54	1.222	2.71	7.07	0.464
Rat (single species)	18	0.888	1.45	4.29	0.416
Hamster (single species)	15	1.722	3.37	12.61	0.526
Monkey (single species)	34	1.139	1.87	7.51	0.498
Dog (single species)	56	0.609	2.14	5.45	0.579
Mouse or Rat (more sensitive)	18	0.734	1.03	2.74	0.347
Mouse or Hamster (more sensitive)	15	0.805	1.15	3.43	0.383
Mouse or Monkey (more sensitive)	31	0.819	1.13	3.53	0.386
Mouse or Dog (more sensitive)	49	0.529	1.78	3.76	0.518
Rat or Hamster (more sensitive)	12	0.687	1.15	3.52	0.431
Rat or Monkey (more sensitive)	15	0.715	0.89	2.22	0.299
Rat or Dog (more sensitive)	17	0.598	0.92	3.10	0.434
Mouse Rat Hamster (most sensitive)	12	0.625	0.98	2.81	0.397
Mouse Rat Monkey (most sensitive)	15	0.689	0.86	2.10	0.295
Mouse Rat Dog (most sensitive)	17	0.551	0.85	2.72	0.421
Mouse Rat Monkey Dog (most sensitive)	14	0.610	0.80	2.09	0.325
All Five Species (most sensitive)	10	0.470	0.67	1.79	0.353

^a Inferred from the Data of Price et al. (2008) for Cancer Chemotherapy Drugs—Human Projections Based on the Most Sensitive of the Species Listed for Each Chemical

^b The “single species” analyses are based on all chemicals where data were available for that species and humans, regardless of whether data were also available for other species. Similarly, the analyses for two

⁸ The log(GSD) is the logarithm of the geometric standard deviation of the parameter being considered. All the logarithms used in the model are common base 10 logarithms.

species are based on all chemicals where data were available for the two named species plus humans even though in some cases data might also have been available for other species.

2.2.2 Interindividual Variability

The variability in response expected for different individuals is treated by combining estimates of interindividual variability in pharmacokinetic and pharmacodynamic parameters from a database of human susceptibility to a variety of biological responses. This human interindividual variability database was updated from our previous work (Hattis et al. 2002). Briefly, the data sets selected for analysis are primarily from the pharmaceutical literature, and provide individual data on measurements for at least five reasonably healthy people for pharmacokinetic parameters, or at least histogram-type dose-response data (e.g., the numbers of people who respond at two or more dose or exposure levels) for pharmacodynamic observations. The parameters measured cover the human interindividual variability for various portions of the pathway from external exposure to effect. The database is classified and analyzed to break down the component variability into the following steps:

- Contact Rate (breathing rates/body weight; fish consumption/body weight)
- Uptake or Absorption (mg/kg)/Intake or Contact Rate
- General Systemic Availability Net of First-Pass Elimination
- Dilution via Distribution Volume
- Systemic Elimination/Clearance or Half Life
- Active Site Availability/General Systemic Availability
- Physiological Parameter Change/Active Site Availability
- Functional Reserve Capacity—Change in Baseline Physiological Parameter Needed to Pass a Criterion of Abnormal Function

The current database (summarized in Excel spreadsheets, available at <http://www2.clarku.edu/faculty/dhattis>) has a total of 447 data groups, each of which has been analyzed to yield an estimate of interindividual variability expressed in lognormal form as a log(GSD). Broadly, the parameters covered include:

- 11 contact rate (2 for children)
- 343 pharmacokinetic (71 include children)
- 93 with pharmacodynamic (and often also pharmacokinetic) information (6 include children)

Interindividual variability in pharmacokinetics (absorption, distribution, metabolism, and excretion) is represented as a central estimate log(GSD) of 0.202 with a lognormal uncertainty represented as a log[log(GSD)] of 0.092.

Three components make up the pharmacodynamic variability: active site availability, functional parameter change in relation to active site availability, and functional reserve capacity in relation to functional parameter changes. In previous work the variability in functional reserve capacity was found to be smaller for more severe effect endpoints (Hattis and Lynch 2007). Summing the variances for each of the pharmacodynamic steps yields:

$$\begin{aligned} & \text{Overall central estimate of pharmacodynamic log(GSD)} \\ & = [(\log(\text{GSD}) \text{ for active site availability})^2 + (\log(\text{GSD} \text{ for functional parameter change})^2 + \log(\text{GSD}) \\ & \quad \text{for functional reserve capacity})^2]^{0.5} \\ & = [(0.0917)^2 + (0.229)^2 + (0)^2]^{0.5} = 0.246 \end{aligned}$$

Similar calculations of the log(GSD)s for overall pharmacodynamic interindividual variability for moderate and mild effects result in larger central estimates of log(GSD)s of 0.267 and 0.330, respectively. In all cases the lognormal *uncertainty* in the estimates of the pharmacodynamic interindividual variability is taken as a log[log(GSD)] of 0.161 (Hattis and Lynch 2007).

This database could be enlarged with further work. The literature on the topic is rapidly expanding. Recently, a corrected estimator was developed (Lynch et al. 2007), which corrects for the bias in including measures from studies that publish only summary statistics (our previously constructed database was restricted to data sources that provided data for individual subjects). Use of this estimator would allow several hundred additional pharmacokinetic measures to be included, and this expansion could be focused on including additional data for potentially susceptible populations such as children or genetically susceptible subgroups.

2.2.3 Subchronic/Chronic

Data from Baird et al. (1996), Weil and McCollister (1963), and Nessel et al. (1995) are applicable for this uncertainty factor and have been incorporated into our previous efforts. The 61 chemicals for which comparative data are available span a wide range of industrial chemicals. The data include 50 assays in rats, and also include 11 sets of assays by inhalation (with the remainder via oral dosing). We found the distribution of the subchronic/chronic ratios to be approximately lognormal with a geometric mean of 2.01 and a geometric standard deviation of 2.17 (Hattis et al. 2002).

2.2.4 Database Deficiency (Missing Either Full Chronic Study or a Reproductive/Developmental Study)

NOAEL data for 35 pesticides with complete toxicological data were gathered and analyzed by Evans and Baird (1998). The toxicological database for each of the 35 pesticides included chronic toxicity studies in rats, dogs, and mice as well as reproductive and developmental studies in rats. To reduce the complications that are associated with multiple species, only rat data are taken into consideration in developing the uncertainty factor distribution used in our previous work (Hattis et al. 2002).

Analyzing these databases, we found that for about three-quarters of the pesticides (26 out of 35), a missing reproductive/developmental study would have made no difference in the assessed NOAEL because the chronic toxicity study yielded a lower estimate of the NOAEL. For the remaining quarter (9 out of 35) of the cases, the needed dose multiplier to correspond to the pesticide data was well described by the lowest quarter of a lognormal distribution with approximately the following the distribution shown graphically in Figure 2-1 and numerically in Table 2-2. In other words, in the Monte Carlo simulations, this distribution was used for 9/35 = 26 percent of cases; for the other 74 percent of cases, a value of 1 was used.

Figure 2-1: Lognormal Plot of the Ratio of the Reduced NOAEL Expected from the Addition of a Rat Reproductive/Developmental Study to a Base of a Rat Chronic Toxicity Study, Based on Evans and Baird (1998) Data for 35 Pesticides

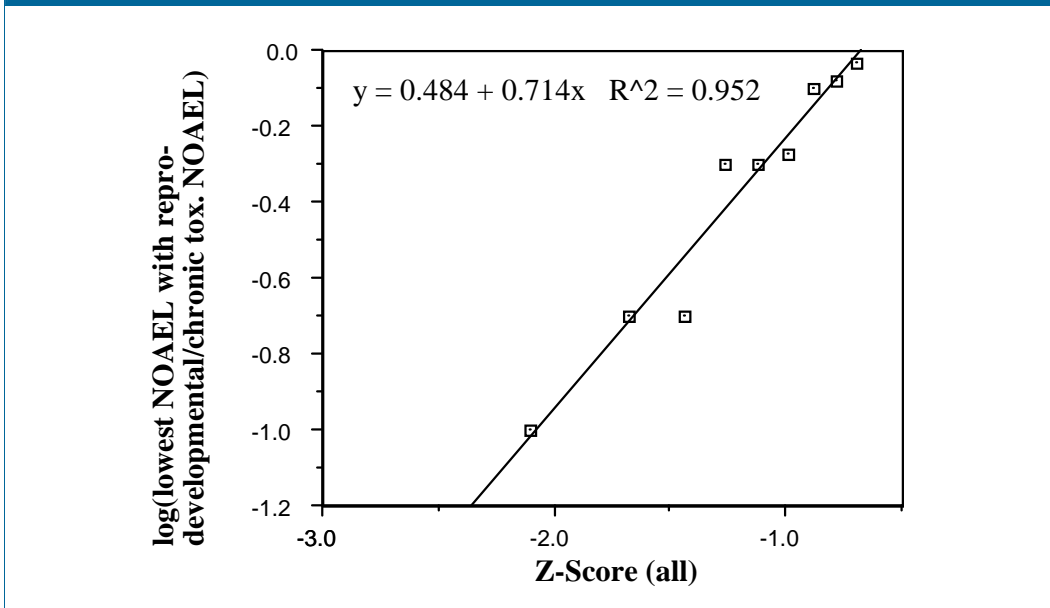


Table 2-2: Distribution of Dose Multipliers Required to Estimate Uncertainty Associated with Database Deficiency

%tile	Dose Multiplier for Database Deficiency
1	0.062
5	0.201
10	0.353
15	0.544
20	0.770
25+	1

In other words, a dose reduction factor of 3 (multiplier of dose of 0.333) would correspond to approximately the 9th percentile of the pesticide data—in the remaining 91 percent of cases, a smaller reduction in dose would be indicated to compensate for the missing data. A similar distributional correction factor was derived for cases with a missing chronic toxicity study. The Evans and Baird (1998) database indicates that addition of a satisfactory chronic toxicity study to pre-existing reproductive/developmental data would not have led to a lowering of the overall lowest NOAEL in 20 of 35 cases (57 percent). Similar to the previous database deficiency analysis, the remaining 15 cases are reasonably described by a lognormal distribution with a geometric mean of 0.98 and a geometric standard deviation of 4.47.

2.2.5 Lowest Observed Adverse Effect Level (LOAEL)/NOAEL

This conversion is not generally needed in the Straw Man model. The goal of the Straw Man model is to predict risks or doses associated with a defined upper confidence limit on risk, rather than levels of exposure postulated to pose either no risk or some level of “acceptable” risk. Because of this, when the original experimental data are available, the Straw Man model uses as its point of departure ED₅₀s estimated from those data. When the original animal data are not available, animal ED₅₀s and associated uncertainties are estimated from the lowest observed adverse effect levels (LOAELs), or if necessary NOAELs, using assumptions about what incidences of effect are usually associated with these traditional endpoints. LOAELs are preferred to NOAELs for this purpose because they require smaller projections in dose to reach ED₅₀ levels, likely with less uncertainty.

2.2.6 Monte Carlo Simulations That Combine the Distributions

A final step in the “Straw Man” approach is the combination of information about the uncertainty in all of the factors used for particular RfDs in Monte Carlo simulations of the uncertainty in risk as a function of dose. The Monte Carlo simulations are run in Excel spreadsheets. Each individual “trial” is on a separate line of the spreadsheet, and consists of a selection of random values from the distributions defined for each uncertainty factor contributing to the RfD being reevaluated in our historical case studies. The outputs from each trial include risks expected at the original EPA RfD and doses tenfold lower and tenfold higher than this; and the dose expected to be associated with a 1/100,000 risk. Arithmetic mean expected values and uncertainty percentiles for these outputs are then calculated from the results of each run of 5,000 trials taken as a whole. These results are then summarized as the averages of distributional values calculated from three parallel runs of 5,000 trials each.

2.3 Incorporating an Assumption of Interacting Background Processes

Chapter 5 of the recent NRC report (2008) recommends a unified framework for dose-response analysis that incorporates the possibility that marginal increments to toxic processes from chemicals will interact with background processes to contribute to ongoing pathologies in the human population. This will lead to a linearization of the incremental population incidence of adverse effects that is expected to occur with incremental dose of the toxicant at low doses.

It is fairly straightforward to add a user-specified amount of background interaction to the spreadsheets used for the Straw Man modeling. The key is to build in a calculation that translates the population incidence of the interacting background that the user believes likely to be present into an equivalent background dose of the toxicant being modeled. In the Straw Man model, this background dose equivalent varies from trial to trial depending on all the values of the uncertain parameters selected for that trial (e.g., chronic/subchronic factor, interspecies projection factor, amounts of pharmacokinetic and pharmacokinetic variability). Risks are then calculated as before for the combined “dose” of the chemical plus the equivalent dose from the background interaction. Finally, incremental risks attributable to the chemical under study over the interacting background are calculated by subtracting out whatever background incidence of effect was incorporated earlier. As mentioned above, this adaptation essentially merges the NRC’s Conceptual Models 1 and 2. These models, as described in NRC (2008) are detailed below.

Conceptual Model 1: *Nonlinear individual response, low-dose linear population response with background dependence.*

Low-dose linearity can arise when the dose-response curves for individuals in the population are nonlinear or even have thresholds but the exposure to the chemical in question adds to prevalent background exposures that are contributing to current disease. The dose-response relationship would be determined to a great extent by human variability and background exposure.

Conceptual Model 2: *Low-dose nonlinear individual and population response, low-dose response independent of background.*

This is the dose-response conceptual model currently in use for noncancer endpoints. For these dose-response relationships, the fraction of the human population responding drops to inconsequential levels at low doses. At very low doses, the threshold dose for toxicity is not exceeded in individuals, or the risk is infinitesimal. The same is true for the population, with the shape of its dose-response relationship determined by the variability in individuals' thresholds.

Specifically, our analysis takes into account:

- Individual threshold dose-response models with and without an appreciable interaction with background pathological processes.
- If essentially no background interaction is considered likely, the user can specify a vanishingly small interacting background incidence of the effect, such as 10^{-8} used in the numerical example below.

Table 2-3 through Table 2-5 show the results of implementing this approach for one of the 18 randomly selected RfD chemicals previously studied in Straw Man analyses—2,3,6-Trinitrotoluene.⁹ The top line of these tables show results assuming absolutely zero incidence of background causes of liver toxicity. Subsequent lines of the tables show the effects of progressively increasing the incidence of interacting background. Columns 2–4 of the tables show the incremental risks over background expected as one proceeds from daily dose rates at 0.1 to 1 to 10 times the IRIS RfD. The last two columns show the change in incremental risk per unit of chemical dose expressed in RfD units (that is, of 5×10^{-4} mg/kg-day). Each line of each table represents the arithmetic mean of three Monte Carlo simulation runs of 5,000 trials each.

⁹ The IRIS RfD for this chemical of 5×10^{-4} mg/kg-day is based on a LOAEL of mild liver toxicity (hepatic swelling and hepatocytomegaly) at 0.5 mg/kg-day in a dog study and a 1,000-fold combined uncertainty factor from LOAEL/NOAEL, interspecies, and human interindividual variability factors.

Table 2-3: Results of “Straw Man” Monte Carlo Simulations on Estimates of Arithmetic Mean “Expected Value” Risks of Mild Adverse Effects from 2,4,6-Trinitrotoluene for Chronic Human Dosage in the Neighborhood of the Current IRIS RfD

Background Risk	Risk Over Bkgd at .1*RfD	Risk Over Bkgd at RfD	Risk over Bkgd at 10*RfD	Slope—Change in Risk Per RfD Increment from .1 to 1 RfD	Slope—Change in Risk Per RfD Increment from 1 to 10 RfD
0	4.9E-06	1.4E-04	6.3E-03	1.6E-04	6.8E-04
1E-05	5.5E-06	1.7E-04	7.5E-03	1.9E-04	8.1E-04
1E-04	1.3E-05	2.4E-04	8.5E-03	2.5E-04	9.2E-04
1E-03	4.0E-05	5.3E-04	1.1E-02	5.4E-04	1.2E-03

Table 2-4: Results of “Straw Man” Monte Carlo Simulations on Estimates of 95th Percentile Risks of Mild Adverse Effects from 2,4,6-Trinitrotoluene for Chronic Human Dosage in the Neighborhood of the Current IRIS RfD

Background Risk	Risk Over Bkgd at .1*RfD	Risk Over Bkgd at RfD	Risk over Bkgd at 10*RfD	Slope—Change in Risk Per RfD Increment from .1 to 1 RfD	Slope—Change in Risk Per RfD Increment from 1 to 10 RfD
0	2.0E-08	7.9E-05	2.8E-02	8.8E-05	3.0E-03
1E-05	8.3E-06	3.0E-04	3.3E-02	3.2E-04	3.7E-03
1E-04	3.4E-05	6.5E-04	3.9E-02	6.8E-04	4.3E-03
1E-03	1.4E-04	1.8E-03	5.2E-02	1.8E-03	5.6E-03

Table 2-5: Results of “Straw Man” Monte Carlo Simulations on Estimates of Median (50th Percentile) Risks of Mild Adverse Effects from 2,4,6-Trinitrotoluene for Chronic Human Dosage in the Neighborhood of the Current IRIS RfD

Background Risk	Risk Over Bkgd at .1*RfD	Risk Over Bkgd at RfD	Risk over Bkgd at 10*RfD	Slope—Change in Risk Per RfD Increment from .1 to 1 RfD	Slope—Change in Risk Per RfD Increment from 1 to 10 RfD
0	1.4E-21	2.9E-12	7.7E-06	3.2E-12	8.6E-07
1E-05	4.8E-07	5.6E-06	2.0E-04	5.7E-06	2.2E-05
1E-04	2.5E-06	2.8E-05	5.5E-04	2.8E-05	5.8E-05
1E-03	1.2E-05	1.3E-04	1.7E-03	1.3E-04	1.8E-04

Table 2-3 displays the results for the single measure of effect that is most frequently needed for economic benefit analyses—the arithmetic mean “expected value” for population risk. This is simply the arithmetic mean over all of the individual trials representing different random draws from the uncertainty distributions included in the model. It can be seen in the first line of the table that this measure of effect is moderately nonlinear—there is approximately a 4 fold difference ($6.8/1.6 = 4.25$) in the slope of the incremental harm/incremental dose ratio as we proceed from the region between 0.1 and 1 times the RfD to the region between the RfD and 10 times the RfD. Proceeding down the table it can be seen that the largest assumed amounts of interacting background appreciably reduce the difference in slope between these two intervals.

Table 2-4 shows similar results using 95th percentiles of the estimated risks. These are the 251st highest values in each 5,000-trial uncertainty simulation run. This measure shows greater dose-

response nonlinearity than was seen for the arithmetic mean with zero or low levels of interacting background, which provides scope for a greater reduction in nonlinearity at the highest levels of interacting background.

Finally, Table 2-5 shows the results for median (50th percentile) estimates of risk—corresponding to the 2,501st highest risk value in each set of 5,000 trials. This measure shows the most nonlinearity with dose in the base (no background interaction) case. The reason why there is more nonlinearity for this measure of the risk distribution with low levels of background interaction is that higher percentiles cannot spread out on the high risk side as much because risks are constrained to be no greater than 1. This tends to compress the high risk tail of the risk uncertainty distributions for the mean and 95th percentile estimates of risks shown in Table 2-3 and Table 2-4. As we proceed to larger amounts of interacting background toward the bottom of the table, the lower end of the risk distribution tends to be compressed upward. For all of the interacting background explored in the tables, the effect of the highest levels of interacting background is to increase the incremental effect of the toxicant. However, if the interacting background levels were raised high enough, it is likely that the upper-end constraint of a risk of 1 would start to compress the ratio of incremental risk to incremental dose of the toxicant.

2.4 Concerns from the Use of Empirically Derived Data Distributions and Other Aspects of the Estimation of Low-Dose Risks

In the following subsections, we discuss several concerns that should be considered when using empirically derived data distributions in risk estimates. These concerns are not meant to discourage the use of such distributions, but to ensure that any uncertainty and bias in the resulting estimates is understood, and to suggest some potential opportunities to estimate and reduce the associated uncertainties through new research. The main concerns include:

- Representativeness of “convenience” samples
- Implicit inclusion of measurement errors in adjustment factor distributions
- Uncertainty in our categorization of a specific toxicant into a putatively predictive group
- Uncertainties in the assumption that human variability distributions for quantal effects are lognormal at low doses.

2.4.1 Representativeness of “Convenience Samples”

One concern with this approach is that the samples of chemicals and people studied for different properties suitable to informing the different uncertainty factors are not stratified random samples. They are, in terminology sometimes used in the trade, “convenience” or “haphazard” rather than random samples. Such samples may well be biased in some known and some unknown ways relative to the universe(s) they are intended to represent. Samples of people studied for pharmacokinetic parameters in Phase I drug studies, for example, may tend to have fewer children, elderly, and very sick people relative to the general population. This is not usually accidental; including individuals who are thought likely to be particularly vulnerable to adverse effects involves ethical concerns and additionally may complicate both regulatory approval and eventual marketing of a drug if a member of a vulnerable group suffers an adverse effect.

Because of this, regulatory analysts who wish to assess potential risks for a general unrestricted population may want to place greater emphasis on a subset of variability data where the populations studied are less restricted (e.g., those that include children, as in our analysis of pharmacokinetic variability data—Hattis and Lynch 2007). Other things being equal, it is likely that this kind of unrepresentative sampling problem produces a tendency for our estimates to understate the real variability in human populations, but only experience and expanded data sets will allow investigators to assess how much underestimation of variability occurs in this way.

The “haphazard sample” problem, however, extends beyond the interindividual variability factor to the general run of toxicological data bearing on the full array of “uncertainty”/adjustment factors. For example, for the data bearing on interspecies uncertainties, it will not generally be known how the choices of contributing investigators of what to study (e.g., selected anti-cancer agents) may have distorted the available uncertainty factor information in general. The only general cure for this problem is to authorize a series of new toxicological experiments on stratified random samples of chemicals designed to gather representative data on adjustment factor distributions for deliberately identified categories of chemicals.

2.4.2 Implicit Inclusion of Measurement Errors in Adjustment Factor Distributions

There is of course no such thing as a perfectly accurate measurement. And unfortunately, our estimates of the spread of “uncertainty”/adjustment factor distributions are not unbiased in the presence of such measurement errors. Measurement errors undoubtedly have caused our estimates of human interindividual variability, interspecies differences, and other adjustment factors to be more spread out than the underlying causal risk-determining contributors. In contrast to the consideration described in Section 2.4.1, this will tend to bias our estimates of the variance of adjustment factor distributions to the high side relative to reality.

Imagine that there was a causal factor that contributed directly to some individual’s exposure and risk—say, an individual’s breathing rate on a particular day when a chlorine cloud passed by after being released from an overturned railroad tank car. Neglecting the effect of variations in the chlorine concentration in parts of the cloud as it passed by the exposed subject, his or her real chlorine intake can be thought of as a single true number determined by the inhalation rate times the time that he or she spends in the cloud. Now let us suppose that we were observing this individual during the passage of the cloud. A laser interferometer fairly precisely measured the expansion and contraction of his or her chest, from which it was possible to estimate the volume of chlorine-contaminated air breathed in. Precise as the measurements might be, there would be some measurement error. Our knowledge of the actual amount breathed can at best be expressed as a number plus or minus some measurement error. Thus, a precise value determining exposure and possible consequences is fuzzed into a distribution by this error. Still, whether the individual lives or dies, or suffers more or less damage from the chlorine, depends only on the fact of the actual exposure, not our uncertainty in estimating it. The fact would not change regardless of the measurement method or its accuracy—for example, whether the measured distribution with error was produced by a relatively accurate laser interferometer, or estimated by general background experiments with people walking or running at a similar velocity in a laboratory.

Extending this example, what if we were not considering just a single individual in a single environmental exposure situation, but a group of individuals, such as underground coal miners? As analyzed in some older work (Hattis and Silver 1993), in a favorable case we might have triplicate

measurements of individual miners that would allow us to estimate both the variability in breathing rates among different miners and the short-term variability plus measurement error seen in the variance among the three measurements for each miner. If each of these variances were normally distributed, the overall observed variance of the measurements would be the sum of the variances from within- and between-miner components. Therefore, if we wished to estimate the long-term interindividual variation among miners that might be predictive of their relative long-term exposures to coal dust, then we should really subtract the short-term variability plus measurement variance from the total observed variance.

Suffice it to say this type of subtraction of measurement error/short-term variability variance is rarely if ever done; and has not yet been attempted for the observations that have gone into our estimates of distributions of interindividual variability and other adjustment factors. However, practically any source of observed data will be subject to this distorting bias relative to the real risk-producing variability distribution for the factor in question. We have not yet attempted to assess whether the net effect of this bias, plus distortions described in earlier sections, is to over-or underestimate real risk-causal variation among people or among chemicals.

2.4.3 How Do We Assess the Uncertainty in Our Categorization of a Specific Toxicant Into a Putatively Predictive Group?

The uncertainty distribution for human interindividual variability assigned to a specific chemical and effect depends on classification of the effect as to the organ system involved, whether the response is due to local versus systemic action of the compound, whether it is immune-mediated or not, and its degree of severity within our crude three-tier severity schema.

However sophisticated we may grow to be in predicting likely sites and modes of action for untested and incompletely tested chemicals, there will still be some chance that we will make a basically incorrect or seriously incomplete categorization. This will cause us to

“HABEAS CORPUS.—A writ by which a man may be taken out of jail when confined for the wrong crime.”

—from the *Devil's Dictionary*
by Ambrose Bierce

draw our random samples for adjustment factor distributions from the wrong base sets. One possible accommodation to this possibility is to make some fraction of the draws from a particular subset of the data using the best estimates of the relevant categorization of the chemical, but to draw the remaining fraction of the modeled cases from the rest of the available data for the various adjustment factor distributions. For example, 2,3,7,8-tetrachloro-dibenzo-dioxin (TCDD) causes cholangiocarcinoma (bile duct cancer) by an apparently non-mutagenic mechanism with a nonlinear dose response relationship in rats (NTP 2006). Therefore in making background interaction calculations similar to those in Table 2-3 through Table 2-5 the most natural assumption is that TCDD might interact with the “background” processes causing some or all the cases of human cholangiocarcinoma. However the anatomical sites at which excess cancers occur are notoriously variable across species. Therefore analysts might wish to include calculations based on the possibility of interaction with processes causing the background incidence of human liver or even lung cancers. This would allow the analyst to represent and assess the potential consequences of the degree of uncertainty in the accuracy of his or her qualitative categorization of the chemical as to mode of action and other factors. A similar type of approach in quantitative dose-response modeling sometimes goes by the name of “Bayesian model averaging” (Whitney and Ryan 2009).

2.4.4 Uncertainties in the Assumption That Human Variability Distributions for Quantal Effects Are Lognormal at Low Doses

In the system we have developed to date, the chief sources of uncertainty that we model are due to the parameters describing the extent of interindividual variability and the uncertainty inferred from data on other adjustment factors. We do not yet include uncertainties from potential departures from true lognormal distributions of individual thresholds for effect, although we have done some analysis of the effects of some possible departures from perfect lognormality in earlier papers (Hattis et al. 1999). Other than empirical data for the limited range over which we can test any distribution, the principal theoretical justification for lognormality is the asymptotic tendency of multiplicatively interacting variability factors to approach a lognormal distribution via the central limit theorem. While this consideration is valid, it is still possible that there could be a large contributor of major variability or several contributors that will provide departures from theoretical lognormal distributions in the cases of specific chemicals/effects. Developing appropriate modifications of our calculation methods to represent uncertainties arising from these sources of departure from lognormal distributions is challenging, but may be feasible with the aid of statistical/mathematical experts. Our expectation is that building in a user-selectable background interaction term will tend to reduce the contribution to overall uncertainty produced by possible departures from lognormality in the underlying distribution of individual thresholds for response.

2.5 Important Details in Projecting Low-Dose Risks and Benefits of Exposure Reductions

In the following subsections we discuss several details in estimating low-dose risks. These details often complicate the analysis, but are important to accurately estimate risks as well as the benefits of exposure reductions.

- How should we proceed if there are good reasons to suspect that there are discrete sensitive subpopulations?
- How can we proceed if we are missing even a basic study from which to infer potency?
- Do structure/activity categorization and potency estimation systems for different endpoints hold promise?
- What value is added by the incorporation of new research findings into the analyses?

2.5.1 Sensitive Subpopulations Because of Discrete Genetic Differences or Disease States

The most generally reasonable approach in these cases is to define separate distributions (each of which may be lognormal) to represent the different subgroups. This leads to an overall mixture distribution (e.g., two or three lognormal distributions in the case of a polymorphism for a gene leading to higher protein levels or higher activity for a specific enzyme.) The overall population may not be “multi-model” in the sense of producing multiple humps in either linear or log space, but it will usually be resolvable into separate subdistributions. It should be noted that postulating subdistributions for different subgroups will often greatly complicate the population model statistically. For example, if only a single lognormal mode is fit, it will be sufficient to derive only two parameters (a mean and a standard deviation in logarithmic space for a lognormal distribution). If there is only one other subdistribution, then at least five parameters are needed (two means, two standard deviations and a “mixing” parameter that specifies what fraction of the population is in each modal subpopulation). Moreover, the uncertainties in each of these parameters will generally have dependencies on the other parameters.

2.5.2 Inference from the Distribution of Potencies from Previously Studied Chemicals

There are a substantial number of chemicals in commerce for which no traditional toxicological studies are available. In the standard system of chemical risk assessment, these chemicals are unfortunately treated as having no toxic potency at all, and it is assumed that no benefits can be estimated for reducing human exposures to them.

As we have emphasized earlier, even if we have no case-specific information about an uncertainty, we can often find relevant generic information that we can use at least to characterize the uncertainty about what the missing data point would be if chemical-specific information were to appear from a previously unknown source or be generated from a newly commissioned experiment. This section describes broad empirical potency distributions that can serve as the starting points for understanding the “value” of chemical structural information and other chemical-specific data for narrowing uncertainties in risks posed by different substances.

The “relevant generic information” referred to in the previous paragraph will usually take the form of a database of prior observations for other compounds that may provide analogies to the untested compound being considered. One of the largest available databases of traditional toxicological information that we know of was compiled by Munro et al. (1996). They describe the compilation process as follows:

Four data sources were chosen to represent a variety of chemical substances (e.g. pesticides, food additives, industrial chemicals, etc.). These included the National Toxicology Program (NTP) technical reports (post-1984), the toxicological monographs prepared by the Joint FAO/WHO Expert Committee on Food Additives (JECFA), the Integrated Risk Information System (IRIS) database, and the Developmental and Reproductive Toxicology (DART) database. No studies on organometallics, inorganic substances, chemical mixtures, or nonstructurally defined substances (such as gums, resins, oils, etc.) were included in the reference database. Only studies with oral exposure were entered into the database with the dosing method specified (i.e. gavage, diet, drinking water or capsule) for each study. The study types included those typically conducted in toxicology, such as subchronic, chronic, reproduction and teratology studies. Short-term and acute studies were not included, since these were considered neither to be relevant to establishing chronic NOELs nor to be representative of other endpoints. Emphasis was placed on obtaining data from chronic studies. The database consisted mainly of studies in rodents and rabbits. Initially, studies conducted in other species, including humans, dogs and ferrets, also were included in the reference database; however, very few studies in these species were found that met the established criteria...

Although the intent was to develop a database consisting mainly of NOELs from chronic studies, in many cases, the lowest and thus most conservative NOEL for a substance came from a subchronic study. In order to group NOELs for substances with only subchronic studies with those with chronic studies to derive the cumulative distribution of NOELs, subchronic NOELs were divided by a factor of three to approximate the most likely NOEL that would be derived from a chronic study...

The data entered into the reference database included the name of the chemical, Chemical Abstracts Service Registry Numbers (CAS No.), structural classification as assessed using the decision tree of Cramer and Ford (1978) (as discussed below), species, sex, route of administration, dose levels tested, study type, duration, endpoints reported, lowest-observed-effect level (LOEL), NOEL and references. A further criterion for inclusion in the database was that a study had to have a demonstrated LOEL as well as a NOEL, thus ensuring that the study was rigorous enough to detect toxic effects. In some instances, however, NOELs were included for studies expected not to demonstrate a LOEL since these were substances, such as major food ingredients, that were without toxicity at the highest dose tested in well-conducted studies. In these cases, the NOEL was conservatively chosen as the highest dose tested. It should be noted that the inclusion of such substances in the database would not bias the database in

favor of higher NOELs since the true NOEL for such substances probably would exceed the NOEL established from the available studies.

In an effort to be conservative in the construction of the reference database, NOELs selected by the author(s) of each study were used, even though in some cases authors tended to be highly conservative in the interpretation of their data. In some instances, it was found that the stated NOEL may have been based on a misjudgment of an adverse effect by the author (e.g. physiological v. toxicological effects) or maternal toxicity). An example of this is isopropyl alcohol, which has been reported to produce teratogenic effects at very low doses (0.018 mg/kg body weight) in one study; however, its structure, known metabolism and other toxicological data provide no evidence for concluding teratogenicity. Even though, scientifically, some of these author-derived NOELs were not thoroughly substantiated, they were included in the reference database, thereby increasing the degree of its conservative nature. NOELs selected by the US Environmental Protection Agency for the IRIS database were entered without further review. For each of the 613 substances, the most conservative NOEL was selected, based on the most sensitive species, sex and endpoint.

The easiest way to use the information in databases such as those of Munro is to describe the data with distributions fit to the data as a whole, or subsets of the data that may form appropriate “priors” for chemicals with specific characteristics that are likely to be related to their toxic potency. Figure 2-2 shows lognormal probability plots¹⁰ of the distributions of the 553 LOELs and 294 NOELs that are helpfully provided in an appendix to the Munro et al. (1996) paper. It can be seen that the distribution of assessed NOELs is close to the regression line representing the fitted lognormal distribution. The LOEL distribution departs appreciably from the fitted lognormal line at the high (right) end—corresponding to the least toxic substances with the highest values. It is possible that many experimentalists do not bother to test substances at dose rates in the thousands of mg/kg-day that would correspond to the part of the plot where the curvature is seen. Table 2-6 further characterizes the two distributions numerically.

¹⁰ In this type of plot, the correspondence of the individual data points to the straight line (representing a normal distribution of the logs of the reported LOELs and NOELs) is a quick qualitative indicator of the fit of the data points to the assumed distribution. The Z-score is the number of standard deviations to the left or the right of the mean of a cumulative normal distribution calculated only from the order of each point in the data. The intercept and slope of each fitted regression line are estimates of the mean and standard deviation of the log-transformed values. (In all cases in this paper we use common base 10 logarithms.)

Figure 2-2: Lognormal Plots of the Distributions of LOELs and NOELs Compiled by Munro et al. (1996) for All Chemical Structural Classes Combined

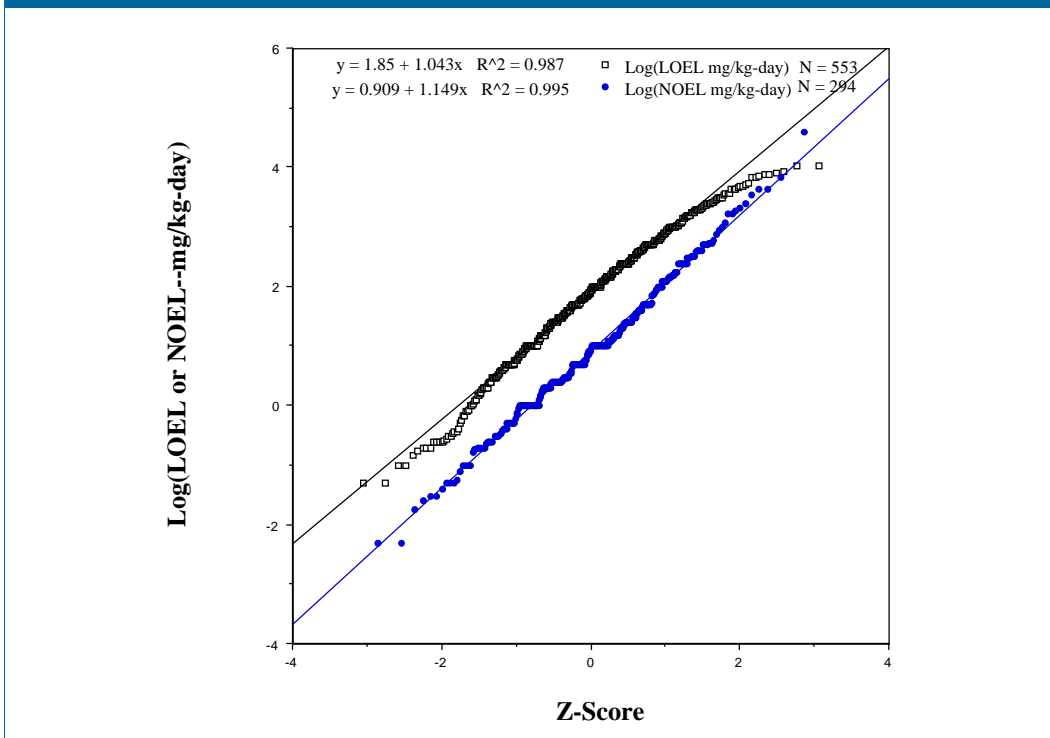


Table 2-6: Summary Statistics and Selected Percentiles of the Distributions of NOELs and LOELs (mg/kg-day)

	1st	5th	50th	95th	99th	Geometric Mean	Geometric Std. Dev. (GSD)
NOEL (mg/kg-day)	0.018	0.050	9.0	44	234	8.1	14.0
LOEL (mg/kg-day)	0.18	0.86	86	2600	7600	71	11.1

Source: Munro et al. (1996).

It can be seen that simply assuming that a random chemical is a candidate for toxicological testing and therefore likely to fall somewhere in these distributions of NOELs and LOELs represents some, but not a great deal of, information on its potency. The ratio of the NOEL for the 95th percentile chemical to the NOEL for the 5th is nearly 1,000 fold. The similar ratio for the LOELs is upwards of 3,000 fold. These ranges of observed variability can be considered 90% confidence limits on NOEL and LOEL measures of potency for a randomly selected untested chemical. Similarly, 98 percent confidence limits (ratios of the 99th to the 1st percentile values) span over 13,000 fold for NOELs and over 40,000 fold for LOELs.

Another endpoint for which a large amount of toxic potency information is available is the LD₅₀. LD₅₀s are likely to be more precisely measured than NOELs and LOELs, and have the advantage that they are not constrained to be one of the experimental doses selected by an investigator for study.

Rhomberg and Wolff (1998), as part of a study of interspecies differences in acute toxicity, compiled oral LD₅₀s for 5,128 chemicals that had been studied in two or more species drawn from the National Institute for Occupational Safety and Health's Registry of Toxic Effects of Chemical Substances (NIOSH 1982). Figure 2-3 shows the a lognormal plot of the distribution of the geometric means of LD₅₀s for all available species for each chemical, and Table 2-7 shows percentiles of the distribution in parallel with Table 2-6. The distribution shows a distinct departure from lognormality, and is generally narrower than was seen for the NOEL and LOEL values compiled by Munro et al. (1996). These data indicate that for this more severe endpoint of acute lethality, the 95th to 5th percentile potency ratio is about 280 fold; similarly, 98 percent confidence limits (the ratio of the 99th to 1st percentile values) span a range of nearly 9,000 fold.

The broad empirical potency distributions in this section should be seen as the starting points for understanding the "value" of chemical structural information and other chemical-specific data for narrowing uncertainties in risks posed by different substances. However, an important point is that even in the absence of such further information, it is possible to quantify uncertainties in toxic potencies for different toxicity endpoints. What is necessary for such an analysis is that either the whole body of previously tested chemicals or some identifiable subset of those chemicals can be fairly identified as a reference set from which the chemical of interest can be seen as a random draw.

Figure 2-3: Lognormal Plot of the Distribution of Mean Log(LD₅₀) for all 5,128 Chemicals with Data for Two or More Species in the Database of Rhomberg and Wolff (1998)

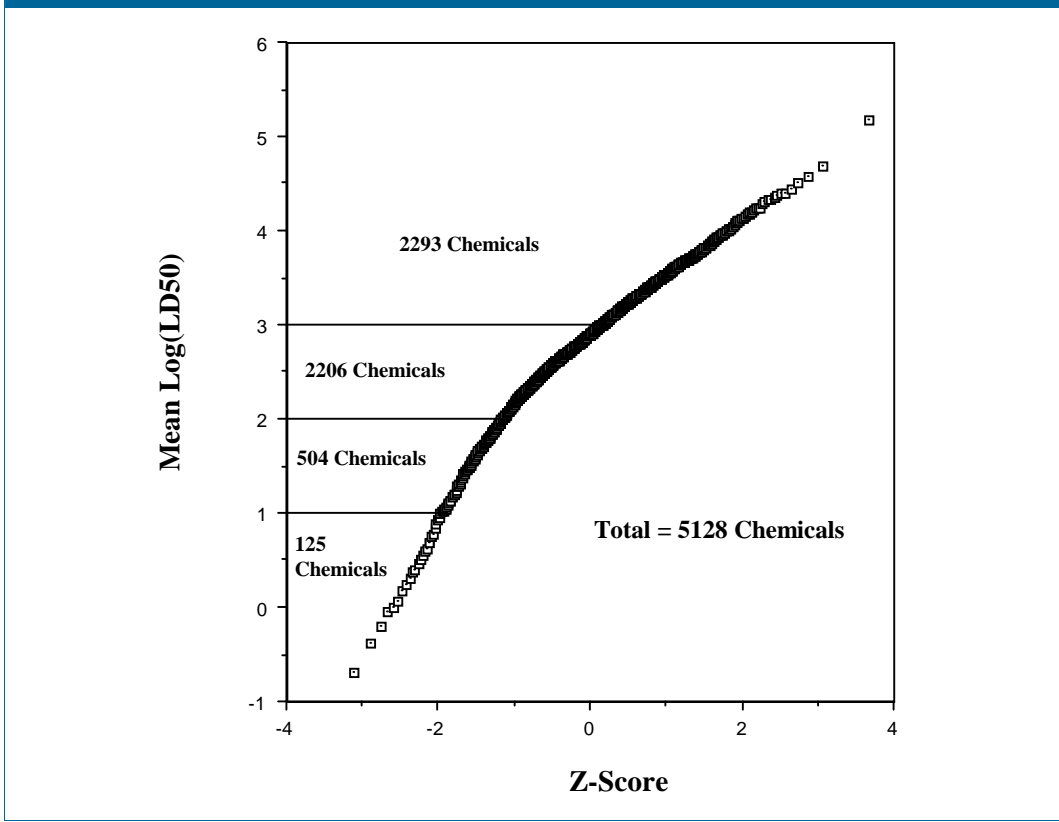


Table 2-7: Summary Statistics and Selected Percentiles of the Distributions of Geometric Mean LD₅₀s for Chemicals Studied in More Than One Species

	1st	5th	50th	95th	99th	Geometric Mean	Geometric Std. Dev. (GSD)
Acute Oral LD ₅₀ (mg/kg)	2.4	29	810	8,120	21,300	680	5.8

Source: Compiled by Rhomberg and Wolff (1998) from data recorded in NIOSH (1982).

2.5.3 Structure/Activity Categorization and Potency Estimation

One important contribution of the paper by Munro et al. (1996) was to test out the predictive value of a classification/ranking system for chemical potencies developed by Cramer and Ford (1978) of the Flavor and Extract Manufacturers Association.¹¹ The Cramer and Ford (1978) system is a fairly

¹¹ One of the coauthors of the Munro et al. paper lists his affiliation as “Research Institute for Fragrance Materials, Inc.” The complexity of the decision tree, the fact that there are only three aggregated classes at the end, and the fact that it was developed by an interested industry association leave open the concern that some of the categorization advice may have been influenced by the equivalent of gerrymandering—having the effect of grouping favored sets of compounds with chemicals known to have less toxicity in the hope of

elaborate decision tree based on answers to 33 questions (Q1 through Q33), many with multiple parts. Chemical groups or other features considered by Cramer and Ford to warrant classification into a category of higher toxicity potential include:

- Q2—Aliphatic secondary amines, cyano, N-nitroso, diazo, triazo and halogenated compounds
- Q3—Elements other than carbon, hydrogen, oxygen, nitrogen, or divalent sulphur; except for some common salts
- Q9—Lactones fused to another ring, or five- or six-membered unsaturated lactones
- Q33—Substances that bear on every major structural component at least one sodium, potassium, or calcium sulphonate or sulphamate for every 20 or fewer carbon atoms without any free primary amines except those adjacent to the sulphonate or sulphamate. (This feature indicates a need for a higher toxicity classification, probably because increased numbers of charged groups make the molecule more soluble and therefore more easily excreted.)

The output of the classification is to sort chemicals into one of only three classes, referred to as Cramer classes, with III being the highest and I being the lowest predicted toxicity grouping. Our analysis of the LOEL and NOEL data in the Munro et al. paper by these classes is shown in Table 2-8; lognormal plots of the distributions by class are shown in Figure 2-4 and Figure 2-5.

It can be seen in Table 2-8 that overall the system does appear to have predictive value for toxic potencies measured as both LOELs and NOELs. The mean logs of each endpoint within the classes differ by an order of magnitude, or a bit more in the case of the NOELs.

However, there is likely to be considerable room for improvement for purposes of quantitative risk and benefits analysis. For both LOEL and NOEL endpoints the Cramer classification captures less than 20 percent of the overall variance in chemical potencies. In our view it would be better to use a multiple regression analysis to sort out which types of chemical constituent groups (and/or which combinations of constituent groups) have predictive value and what their approximate individual contributions are to toxicity in a quantitative sense. (For example, it is doubtful to us that diverse constituents listed in the bullets describing questions 2 and 3 above have the same quantitative implication for potency.) The result of this type of analysis would be a model that would yield quantitative estimates of LOELs and NOELs (with an estimate of uncertainty in the form of the root mean square error of the model). The data are available in the Munro et al. paper to create such a quantitative predictive model, although it should of course be subject to further testing, ideally with independently gathered data.

gaining more favorable regulatory treatment. However, this kind of possibility could only be tested by considerably more analysis of the Straw Man model than we have been able to perform thus far.

Table 2-8: Summary Statistics for Distributions of LOELs and NOELs for Chemicals Sorted into Cramer Classes						
A. Analysis for LOELs						
Class	N	Mean log	log(GSD)	Log Variance	Gmean	GSD
III	442	1.68	1.026	1.05	48.0	10.62
II	21	1.91	0.662	0.44	80.7	4.59
I	90	2.66	0.825	0.68	454	6.69
	553			0.969 = weighted average within-group variance		
All Chemicals	553	1.85	1.047	1.10	70.5	11.14
				0.969/1.10 = 0.88 = fraction of total variance within Cramer classes		
B . Analysis for NOELs						
Class	N	Mean log	log(GSD)	Log Variance	Gmean	GSD
III	241	0.682	1.027	1.06	4.8	10.65
II	10	1.295	0.829	0.69	19.7	6.75
I	43	2.094	1.104	1.22	124.1	12.70
	294			1.067 = weighted average within-group variance		
All Chemicals	294	0.909	1.146	1.31	8.1	14.00
				1.067/1.31 = 0.81 = fraction of total variance within Cramer classes		

Figure 2-4: Lognormal Plots of LOEL Distributions Compiled and Sorted by Munro et al. (1996) Sorted into Cramer Classes

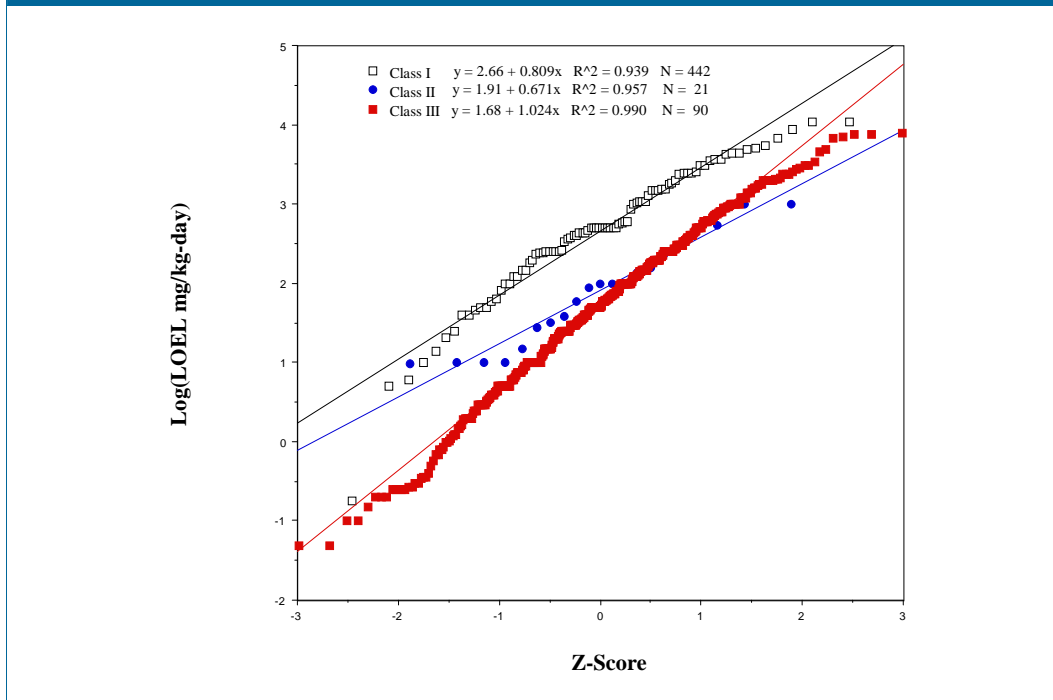
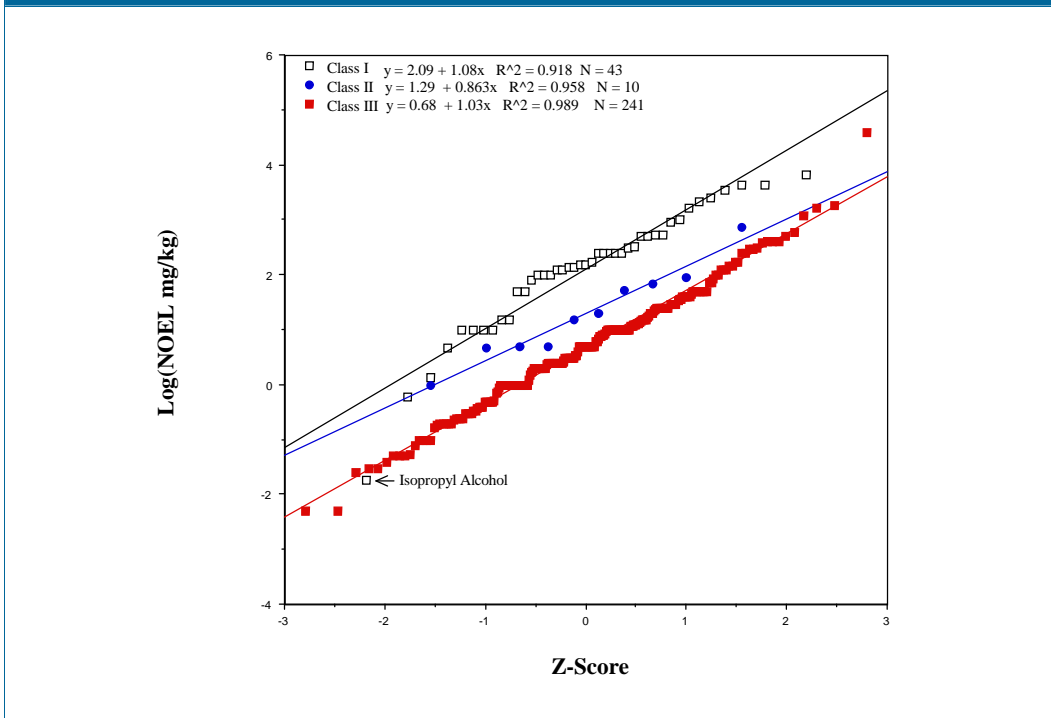


Figure 2-5: Lognormal Plots of NOEL Distributions Compiled and Sorted by Munro et al. (1996) into Cramer Classes



2.5.4 Prospectively Assessing the “Value of Information” for Adding Information from New Research

The variance-reduction analysis for Cramer classes in the previous section illustrates the basic technique for evaluating the contribution of new types of information. If a specific test or other data reduces uncertainty (summarized for convenience as a log variance in the previous section) then it will lead to improved choices among policy options. The “value” of this reduced uncertainty in the context of a specific type of regulatory choice can be evaluated if there is a defined decision rule that depends on the toxicity estimate and the residual uncertainty in that estimate, as in the case of the “risk-specific dose” proposed in the original Straw Man model.¹²

¹² The proposal in that case was that the RfD be redefined as the dose expected to produce less than a 1/100,000 incidence of mild adverse effects with 95 percent confidence, based on evaluation of a defined set of sources of uncertainty.

3 The Promise of Analyses of Specific Resources Available for Developing Potency Distributions for Air Toxics

The discussion in the previous section reflects the orientation of many toxicological data and risk analysis methods toward the oral route of exposure. This section surveys resources that are available for extending the same type of analysis to the inhalation route. A literature search was conducted in order to ascertain the availability of data that could be used for constructing reference distributions. These distributions could be used following the framework described in Section 2 to inform the potency of various HAPs in the absence of chemical-specific data.

3.1 RD₅₀ Determinations by Alarie

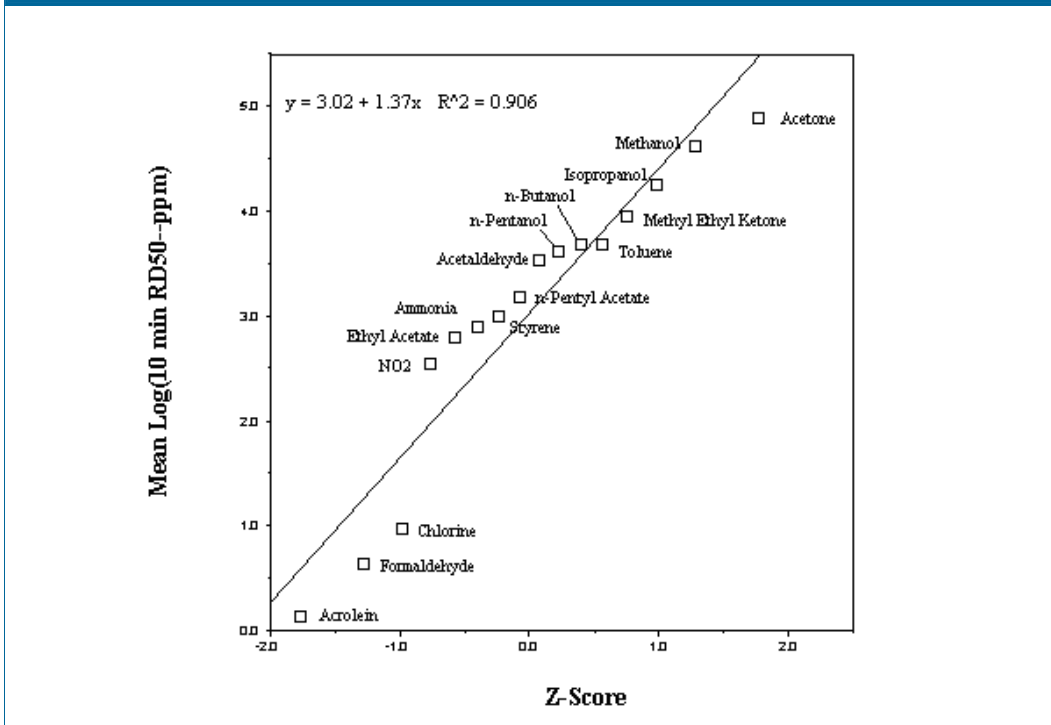
Measurements of RD₅₀s are the single most promising source of relative potency data for irritation responses via the inhalation route. The response is caused by activation of receptors triggering a reflex via the trigeminal nerve (Kasanen et al. 1998).¹³ Dose-response relationships appear to follow standard Michaelis-Menten kinetics (Nielsen and Vinggaard 1988), and structure-activity studies of the response have been productive (Alarie et al. 1998; Steinhagen and Barrow 1984; Kristiansen and Nielsen 1988). Based on an analysis of a database of 145 RD₅₀s, Alarie et al. (1998) report that it is possible to relate potencies to either the chemical's reactivity or physical properties.

The idea of relating these simple, short-term measurements in mice to standards for protection of humans from respiratory irritation from a wide variety of air pollutants originated with Yves Alarie decades ago (Alarie 1966, 1981). Most recently, Kuwabara et al. (2007) of the California Environmental Protection Agency have compared RD₅₀s with results of human irritation measurements and with California's "acute reference exposure levels" (RELs)—which are analogous to short-term RfCs.

Figure 3-1 shows a lognormal plot of the RD₅₀ data compiled by Kuwabara et al. (2007) for 10-minute measurements. Many other data are available in the Kuwabara et al. paper and elsewhere for other exposure durations, and a more extensive analysis of inhalation potency distributions for chemicals is therefore possible, analogous to Figure 2-2 through Figure 2-5 in the previous section.

¹³ On the other hand, it has also been reported that these reflex responses are not necessarily predictive of inflammatory and tissue damage responses seen over somewhat longer time scales (Bos et al. 2002).

Figure 3-1: Lognormal Plot of the Distribution of Mean 10-Minute log(RD₅₀s)



Source: Data Compilation of Kuwabara et al. 2007

Analysis of even this limited sample suggests that irritant potencies of different chemicals vary to a degree similar to that seen for LOELs and NOELs, and may have a pattern of departure from a lognormal distribution similar to that seen for LOELs and LD₅₀s.

Figure 3-2 extends this analysis to the much larger database (covering 145 chemicals) of Alarie et al. (1998). It can be seen that this set of data is relatively well described by the fitted lognormal distribution line. Moreover, the variability in potencies among chemicals seen in the slope of the lognormal line is similar to the variability seen for LOELs and NOELs in the completely independent oral potency distribution data collected by Munro et al. (1996) and plotted in Figure 3-2.

Figure 3-2: Lognormal Plot of 145 RD₅₀ Values Compiled by Alarie et al. (1998)

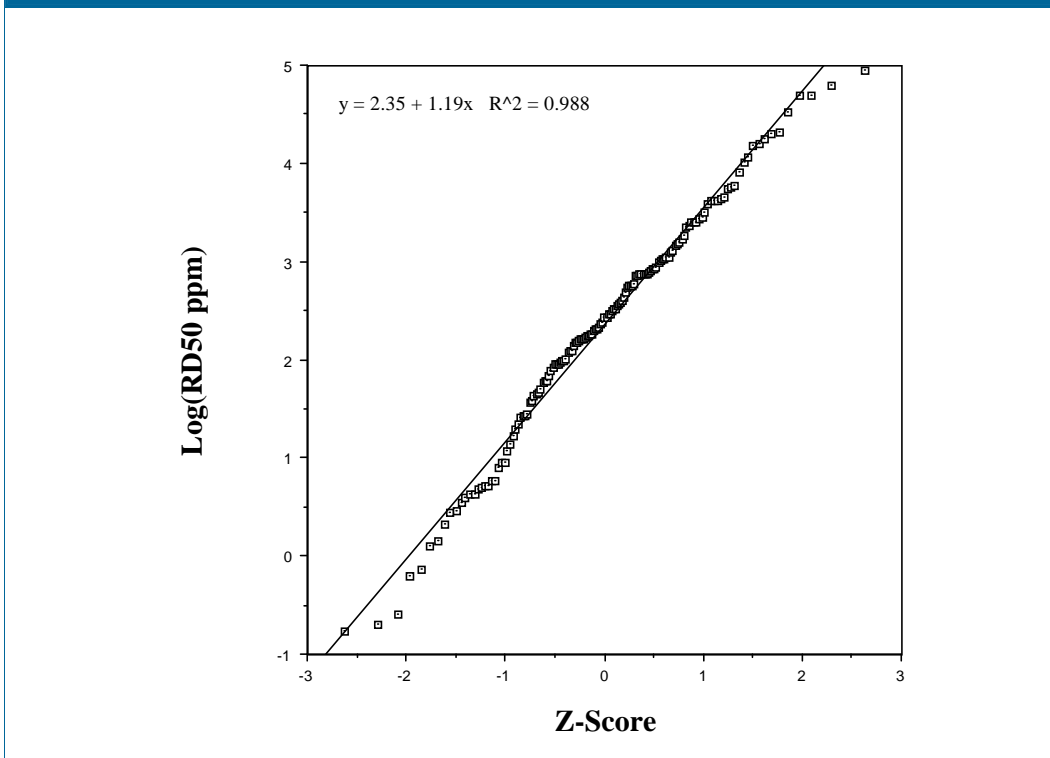


Figure 3-3 shows similar plots of the Alarie et al. RD₅₀ data broken down into groups of chemicals where the authors believe the RD₅₀s are largely determined by either (a) physical properties (called *p* chemicals) or (b) chemical reactivity (called *r* chemicals). Unfortunately, these categorizations were made on the basis of whether the RD₅₀ itself differed from the vapor pressure by tenfold or more. In our view it would be better not to incorporate a relationship involving the primary dependent variable used for evaluation to be part of the putatively independent variable used for classification. A better approach would be to use first principles of chemical structure or reaction rates with potential biologically relevant macromolecules. However, this subcategorization does capture an appreciable amount of the variability in RD₅₀ potencies present in the database,¹⁴ as indicated in Table 3-1.

¹⁴ About 35 percent = (1 - 0.65, where .65 is the fraction of within-group versus total variance).

Figure 3-3: Comparison of Lognormal Plots of RD₅₀ Distributions for Chemicals Classified by Alarie et al. (1998) as Likely to Have Physical versus Chemical Reaction Mechanisms Determine Their RD₅₀s

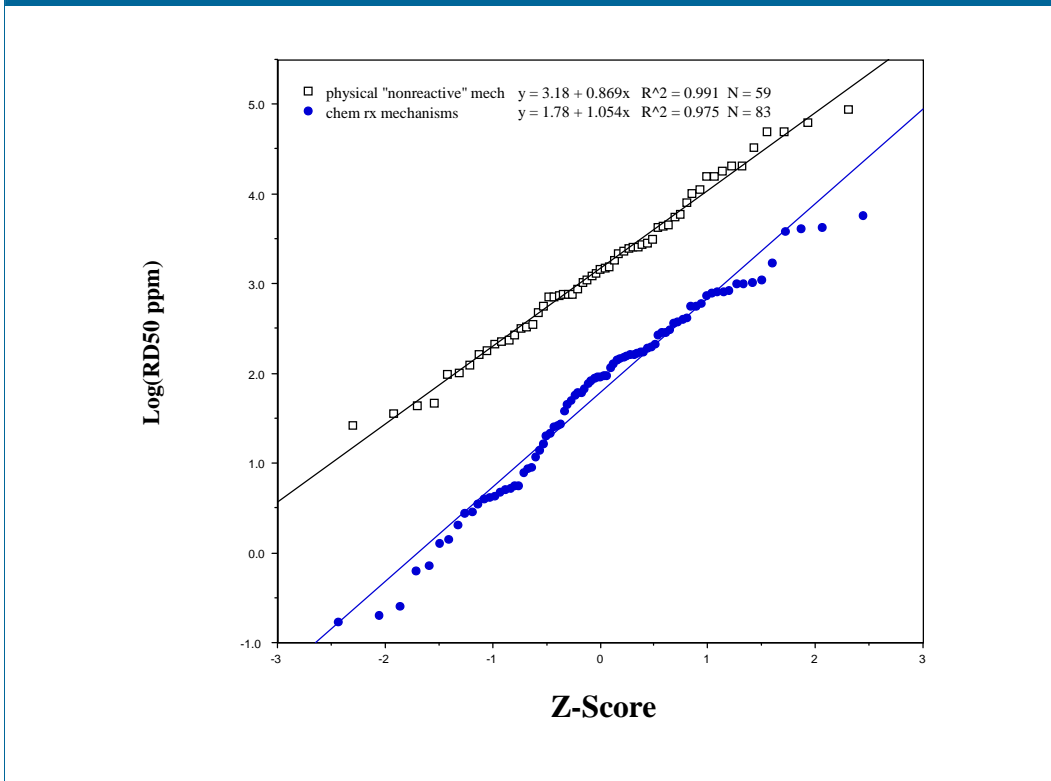


Table 3-1: Variance of RD₅₀s Explained by the Physical versus Chemical Reaction Categories of Alarie et al. (1998)

Class	N	Mean log	log(GSD)	log Variance	Gmean (ppm)	GSD
Chemical Reaction	83	1.78	1.027	1.06	61	10.7
Physical Mechanism	59	3.18	0.859	0.74	1,500	7.23
				0.92 average within-group log variance		
All Chemicals	142	2.36	1.193	1.42	8.1	14.00
				0.65*		

*0.92/1.42 = 0.65 = fraction of total variance within “physical” and “chemical” reaction categories.

Alarie et al. (1998) also provide data for several other physical/chemical properties for the 145 chemicals in their database:

- Excess molar refraction
- Chemical dipolarity/polarizability

- Chemical overall or effective hydrogen-bond acidity
- Chemical overall or effective hydrogen-bond basicity
- Chemical Ostwald partition coefficient on hexadecane at 25°C
- Chemical Ostwald partition coefficient on olive oil at 30°C
- Chemical vapor pressure at 22–25°C

The author does not provide any clear theoretical mechanistic rationale why each of these physical or chemical properties should matter either to the RD_{50s} in general or to RD_{50s} for a specific subset of chemicals. However, the availability of these data does present the opportunity for empirical analyses of the predictive value of these parameters.

There is also extensive categorization by recognizable chemical groups¹⁵ and regression analysis results. Unfortunately, these are not presented in a form that allows ready evaluation of the statistical significance or confidence limits of different independent variables. Nevertheless, much of this other material in this paper could contribute to the development of a predictive model of RD_{50s} for those HAPs that have not been specifically tested in the Straw Man system.

The chemical structural categories used are listed in Table 3-2. This categorization has the advantage of being far simpler and easier to evaluate than the Cramer decision tree. Still, it leaves open the possibility of ambiguities—for example, if there are both halo groups and aldehydes or ketones in the same molecule. Additional improvements would include the ability to score a chemical as having more than one halo, aldehyde, or other chemical group where indicated (perhaps with some further adjustment, for large molecules, with the molecular weight or the vapor pressure). This would allow a basis for multiple regression analyses to sort out the contributions of different groups to irritant potency.

¹⁵ Three chemicals were excluded from the reactive versus physical mechanism categories, and therefore do not appear in Table 3-1.

Table 3-2: Chemical Structural Categories Used By Alarie et al. (1998)

Chemical Structural Categories
Aliphatic acetate (saturated)
Aliphatic acetate (unsaturated)
Aliphatic acid (saturated)
Aliphatic alcohol (saturated)
Aliphatic aldehyde (saturated)
Aliphatic aldehyde (unsaturated)
Aliphatic amine (saturated)
Aliphatic amine (unsaturated)
Aliphatic ether (saturated)
Aliphatic ether (unsaturated)
Aliphatic halogenated (unsaturated)
Aliphatic hydrocarbon (saturated)
Aliphatic ketone (saturated)
Aliphatic ketone (unsaturated)
Aromatic aldehyde saturated
Aromatic alkylbenzene (halogenated)
Aromatic alkylbenzene (saturated)
Aromatic alkylbenzene (unsaturated)
Aromatic benzene (halogenated)
Aromatic ketone (saturated)
Aliphatic isocyanate
Aromatic isocyanate
Other

3.2 Short-Term LC₅₀ Concentration/Time Tradeoff Data Assembled by ten Berge

An important consideration in assessing risks for airborne toxicants is the interplay between exposure duration and intensity (concentration). There are many examples in the literature where empirical data are not compatible with a simple Haber’s rule¹⁶ interpretation. Table 3-3 shows the classical results of ten Berge et al. (1986), who performed probit analyses using a CⁿT transformation of dose (C) and time (T) in modeling inhalation lethality observations across a wide variety of animal species.¹⁷ If Haber’s rule were followed, *n*, the empirically estimated power of concentration that best fits the available data, would equal 1. However, it can be seen that the lower 95 percent confidence limit on *n*, appears to exceed a value of 1 in 14/20 or 70 percent of the cases shown, although the best-fitting *n* is 1.2 or less in 8/20 or 40 percent of the cases.

¹⁶ Haber’s rule states that time-weighted exposures should have equal toxic potential. In other words, exposure at 1 mg/m³ for 10 days is equally toxic to exposure at 10 mg/m³ for 1 day.

¹⁷ The equation fit by ten Berge (1987) is:

$$\text{Probit of response} = \text{standard normal deviate} + 5 = B0 + B1 \ln(C) + B2 \ln(T)$$

The values of *n* reported in Table 3-3 are the ratios of the B1 and B2 coefficients. In modern applications, the 5 term can be omitted. (It was originally added to avoid negative numbers, which could not be processed using the mechanical calculators of the 1930s).

Table 3-3: Exponents of CⁿT Dose Metrics from log Probit Fits to Inhalation Lethality Data for a Variety of Chemicals

Gas or Vapor	Exponent (n)	95% Confidence Limits
<i>Local Irritants</i>		
Ammonia	2.0	1.6-2.4
Hydrogen chloride	1.0	0.7-1.3
Chlorine pentafluoride	2.0	1.4-2.6
Nitrogen dioxide	3.5	2.7-4.3
Chlorine	3.5	2.5-4.4
Perfluoroisobutylene	1.2	1.1-1.4
Crotonaldehyde	1.2	1.1-1.3
Hydrogen fluoride	2.0	1.2-2.8
Ethylene imine	1.1	0.8-1.3
Bromine	2.2	2.0-2.4
Dibutylhexamethylenediamine	1.0	0.6-1.4
<i>Systemic Action</i>		
Hydrogen cyanide	2.7	1.8-3.7
Hydrogen sulfide	2.2	1.6-2.7
Methyl t-butyl ether	2.0	1.0-2.9
Chlorobromomethane	1.6	1.4-1.8
Ethylene dibromide	1.2	1.1-1.2
Tetrachloroethylene	2.0	1.4-2.6
Trichloroethylene	0.8	0.3-1.4
Carbon tetrachloride	2.8	1.9-3.7
Acrylonitrile	1.1	1.0-1.2

Source: ten Berge et al. (1986).

Note: Data combined from rats, mice, dogs, monkeys, guinea pigs, and rabbits, as available.

The CⁿT transformation is a useful empirical convenience for summarizing dose-response data collected for a range of exposure durations. However, it carries the inherent difficulty that, for traditional toxic processes, it must fail at extremes of dose and time outside the bounds of available data. For example, taking the ten Berge et al. (1986) probit lethality equation quoted in footnote 17, it should theoretically be possible to produce 50 percent mortality at any concentration of a toxicant, so long as we extend the exposure duration (T) long enough.

Where there is good reason to believe that some toxic damage that mediates ultimate end effects is accumulating in the face of a repair process, it will sometimes be possible to model the dynamics of the repair process from data on dose multiplied by time response. For example, Hattis and Shapiro (1990) used a model assuming a simple linear repair of damage, and the production of a toxic effect when a critical level of internal damage is reached.¹⁸ However, detailed pharmacodynamic modeling of this type is beyond the scope of this generic methodology paper.

¹⁸ When the simple linear repair rate theory was later applied to predict the dynamics of recovery of neurobehavioral functions following the end of acrylamide exposures, the model seriously underpredicted the times required for recovery (Hattis and Crofton 1995). This was taken to indicate that the original assumption of fully reversible damage used in the original analysis was likely to be incorrect—and that some of the delay in the manifestation of neurotoxic injury was likely to be due to redistribution of function from more- to less-damaged neural paths. However, this very example illustrates the scientific advantage of theorizing based on a quantitative mechanism over a simple empirical summarization of data as represented by the CⁿT transformation of dose data. Mechanistic theories can be the basis of predictions that can extend

The time dynamics for local toxic action of nonreactive chemicals in the respiratory system will also depend on simple physiochemical factors such as vapor pressure, octanol/water partition coefficient, and water solubility/Henry's law constant. These help determine rates of uptake and loss from lung tissue. As a next step, this large database of RD_{50} s assembled by Kuwabara et al. (2007) should be analyzed to detect any regularities in changes in RD_{50} values with time that may be associated with these parameters. Other things being equal, lower vapor pressures, and greater octanol/water partition coefficients should be associated with slower approach to equilibrium local concentrations and therefore lower RD_{50} s with longer durations of exposure. There has been historical use of octanol/water partition coefficients to understand not only partitioning but also potency of unreactive chemicals such as alkanes and aliphatic alcohols for inducing anesthesia (Hau et al. 2002).

3.3 Prospects for a Compilation of LC_{50} s and Analyses of LC_{50}/LD_{50} Correlations

Our literature search turned up a few other references to collections of potency data that may prove useful in developing a priori reference distributions for the toxic potency of different HAPs. These include:

- A compilation by Grant et al. (2007) of inhalation NOAELs and LC_{50} data for 97 chemicals. Using both of these types of data, they determined NOAEL/ LC_{50} ratios. They also utilized a “globally harmonized system of classification and labeling of chemicals”—yielding five categories of chemicals based on potencies for acute toxic effects.
- A series of measurements by Frantfk et al. (1994) of air concentrations of 48 common solvents (hydrocarbons, alcohols, ketones, and acetates) causing defined neurological changes. These effect concentrations were reportedly several fold lower than those needed to induce behavioral inhibition and one or two orders of magnitude lower than concentrations needed to induce narcosis. Reportedly, potencies for these kinds of effects were not highly correlated with octanol/water partition coefficients.

It would be straightforward to analyze the correlations of inhalation potencies in these data sets with LD_{50} s from the large Rhomberg and Wolff data set with physico-chemical parameters to partially correct for pharmacokinetic determinants of internal concentrations per unit of external exposure.

3.4 Potencies and Interindividual Variability/Susceptibility Distributions for Bronchoconstricting Agents—Guidance for a Common Mode of Action for Several Air Toxics—Acting in Asthmatics

Short-term bronchoconstriction responses in asthmatic people are important considerations in standard setting for criteria air pollutants—sulfur dioxide (SO_2) is discussed in this section, and ozone is discussed in **Section 3.5**. Moreover, it happens that we have extensive information about interindividual variability in susceptibility for this response from large epidemiological studies in which graded doses of methacholine have been administered to people to identify the lowest dose

beyond the previously studied conditions of dosage, duration, or age of the subjects exposed. If these out-of-range predictions are refuted by subsequent data, the departures from theory-predicted values can be the basis for more advanced mechanistic inferences and hopefully better predictions in the future. On the other hand, success or failure of a C^bT model fit to predict further data does not naturally lead to new kinds of experimental observations that can shed light on relevant mechanisms and toxic risks outside the range of previous experimental studies.

causing a specific amount of reversible reduction of FEV1 (amount of air that can be breathed out in one second—a common measure of lung function) in different individuals, although the extent of variability observed can be quite different for observations of different populations by different investigators (Table 3-4). Individual variability in susceptibility to this response is generally well described by lognormal distributions (Figure 3-4), although it should be noted that the published data do not usually report the fraction of people responding at the extreme ends of the distributions. Such data might perhaps be obtained from direct contacts with the investigators.

Figure 3-5 and Figure 3-6 show similar plots of bronchoconstriction response data for SO₂, as compiled from several clinical studies of exercising asthmatics in the latest EPA “Risk and Exposure Assessment” document prepared as part of its periodic review of the adequacy of current standards for criteria air pollutants (EPA 2009, Appendix C). These data are an important basis for current consideration of a short-term exposure standard for SO₂, and therefore provide a possible basis for comparison with HAPs that induce similar short-term respiratory responses.

It is not clear at this point that there will be a simple relationship between potency for this kind of effect and RD₅₀ values, but this seems like a reasonable hypothesis for further exploration. If there is such a relationship, then the substantial body of RD₅₀ data could be translated into terms that can be related to extensive regulatory evaluation work with criteria air pollutants.

Table 3-4: Observations of Human Interindividual Variability in External Concentrations of Methacholine Needed to Produce Defined Short-Term Changes in Respiratory Parameters				
Response and Reference	Type of Population	log(GSD)	5%-95% log(GSD) Conf. Limits	Statistical Weight = 1/variance of log[log(GSD)]
100% increase in baseline specific airway resistance (Balmes et al. 1997)	66 Healthy athletic adults, 18-50	0.421	0.39-.046	1,986
PC20— 20% decrease in FEV1 (Tashkin et al. 1996)	5,733 smokers, mild/moderate obstruction	0.642	0.59-0.70	2,157
10%, 15%, and 20% decreases in FEV1 (Paoletti et al. 1995)	748 Females—general population	0.740	0.59-0.93	276
10%, 15%, and 20% decreases in FEV1 (Paoletti et al. 1995)	810 Males—general population	0.998	0.80-1.25	280
PC20— 20% decrease in FEV1 (Sears et al. 1986)	813 nine year old New Zealand children	1.128	0.88-1.45	225
PC20— 20% decrease in FEV1 (Bakke et al. 1991)	490 Norwegian adults, Age 18-73	0.974	0.81-1.17	434
PC20— 20% decrease in FEV1 (Hanania et al. 1998)	15 Allergic asthmatic patients	0.599	0.44-.82	149
PC20— 20% decrease in FEV1 (O'Connor et al. 1987)	468 Male veterans	1.088	0.92-1.32	440
Summary, All Methacholine Data		0.704	0.67-0.74	5,947

Source: Hattis (2008).

Figure 3-4: Lognormal Probability Plots of the Distribution of Methacholine Concentrations Needed To Cause a 20 Percent Reduction in FEV1 in Different Populations

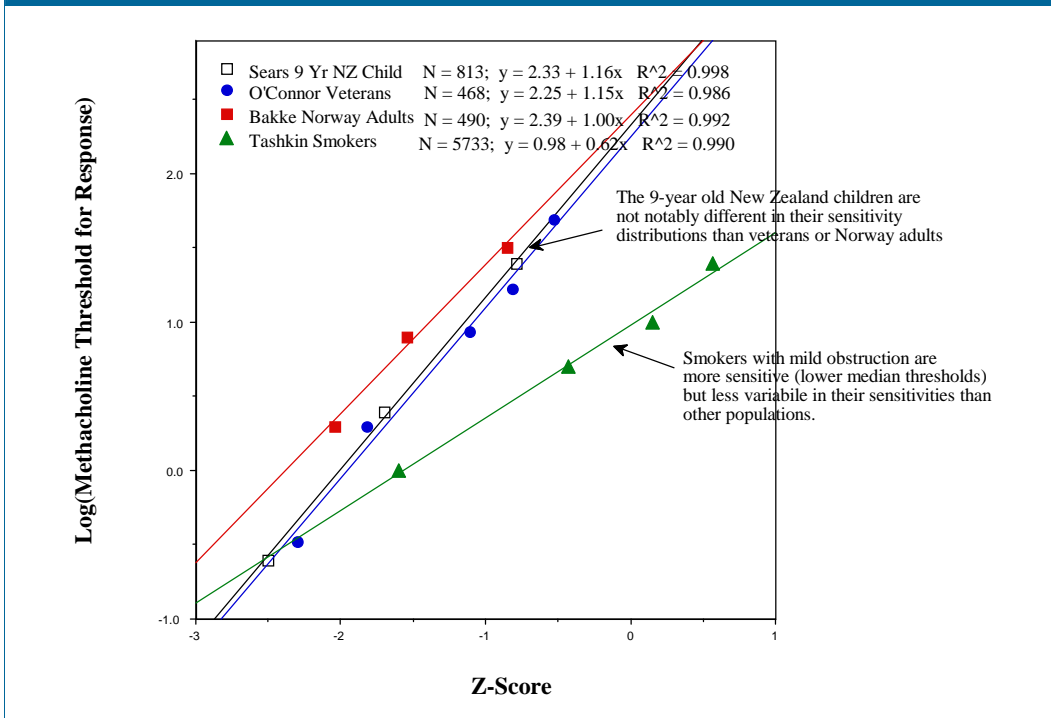
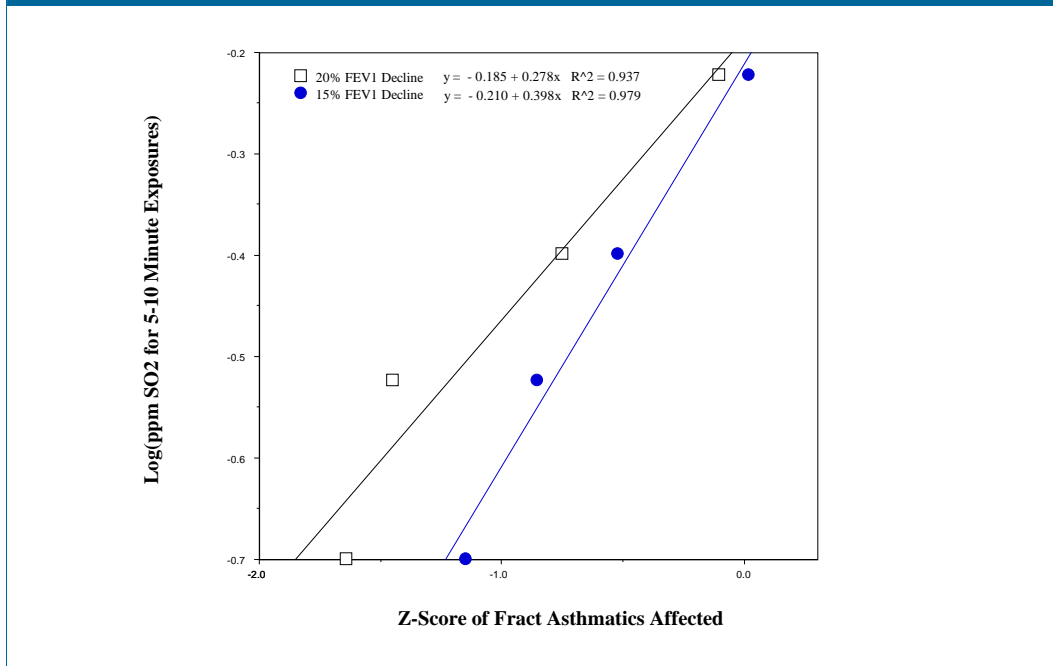
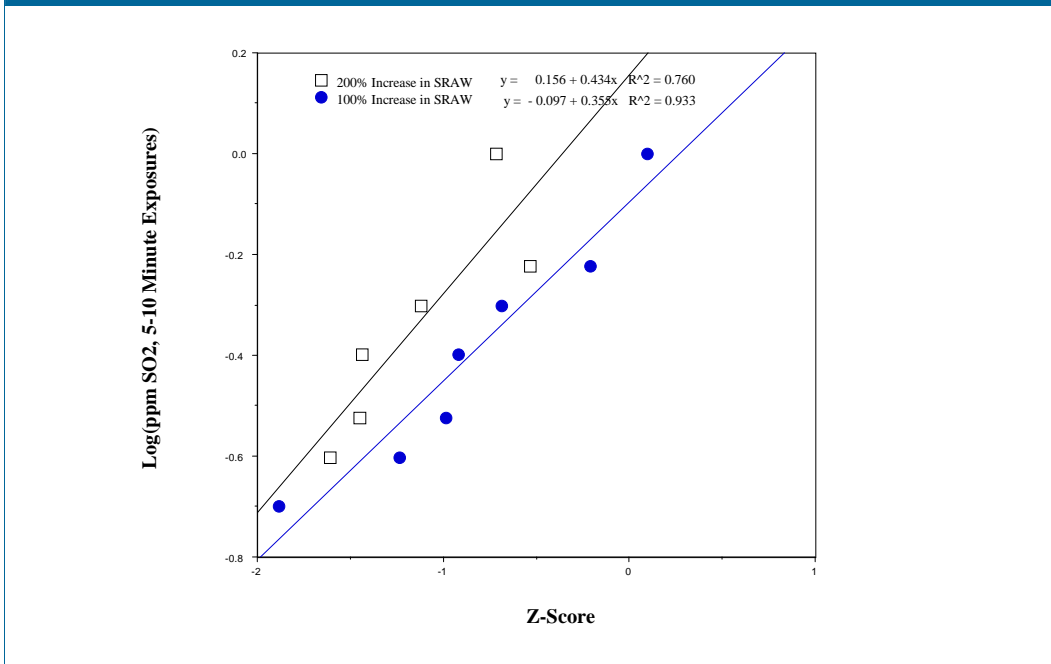


Figure 3-5: Lognormal Probability Plots of the Distributions of SO₂ Concentrations Needed to Cause a 15 or 20 Percent Reduction in FEV1 in Exercising Asthmatics



Source: Combined Plot of Data Compiled by EPA (2009) from Several Clinical Studies

Figure 3-6: Lognormal Probability Plots of the Distributions of SO₂ Concentrations Needed to Cause a 100 or 200 Percent Increase in Specific Airway Resistance in Exercising Asthmatics



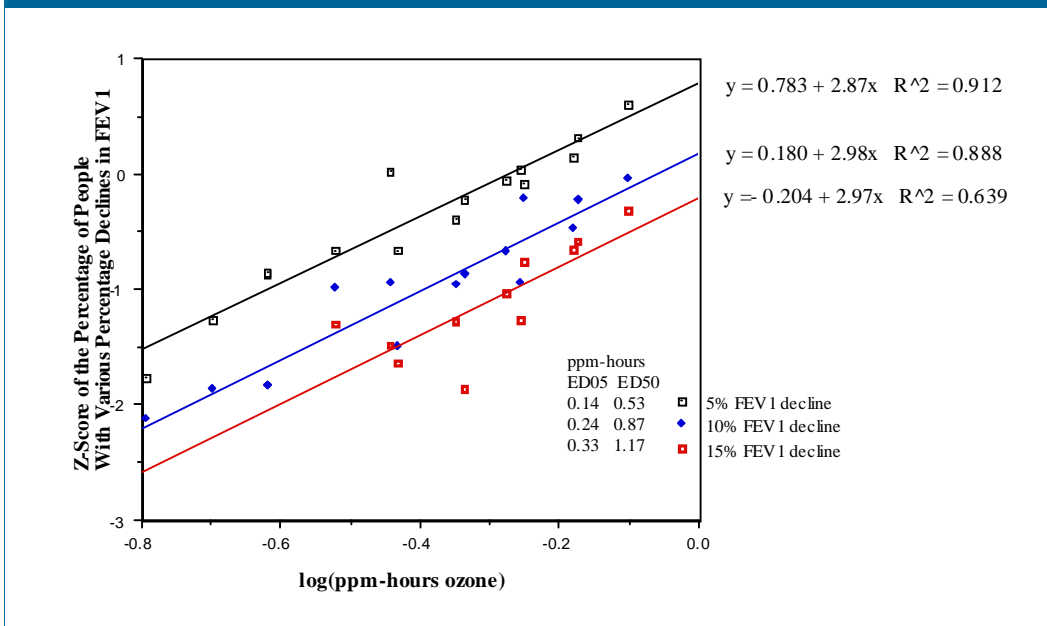
Source: Combined Plot of Data Compiled by EPA (2009) from Several Clinical Studies

3.5 Actions of and Human Susceptibility Distributions for Ozone—Another Basis for Analogies with Oxidant Chemicals?

Ozone is another criteria air pollutant for which there are extensive human clinical observations of changes in FEV1 in response to short-term exposures. In this case, however, the response occurs over a period of hours, and is therefore better seen as a delayed reaction to injury from this strong oxidant than an immediate neurogenic reaction to the presence of the pollutant. Figure 3-7 shows probability plots of the distributions of individual “thresholds” for different degrees of FEV1 response to multi-hour exposures to different combinations of time and exposure concentration for ozone. These plots also indicate that lognormal distributions are reasonable descriptions of the data.¹⁹ The measure of interindividual variability indicated by these data and those compiled by EPA (2009) for SO₂ are both toward the lower end of the observations recorded in Table 3-4 for methacholine—log(GSD) of 0.28–0.43 or so, indicating lower interindividual variability.

¹⁹ The plots are in the opposite orientation from the ones we usually produce (Z-scores on the y-axis rather than the x-axis), and therefore the slopes of the lines are the inverse of our usual measure of interindividual variability—the log(GSD).

Figure 3-7: Composite log Probit Plot of the Ozone Dose-Time-Severity-of-Response Relationship



Source: Hattis 1998, based on data of McDonnell et al. 1995.

3.6 Preliminary Categorization and Grouping of the Officially Designated Set of Air Toxics by a Priori Recognition of Basic Chemical Structural Alerts or Prominent Concerns for Carcinogenesis

In preparation for possible later work on HAPs, we have taken a preliminary look at their chemical structures to see what major chemical groupings could be used to organize an overall risk assessment effort with a view toward facilitating cross-chemical inferences on related properties. This section summarizes our early conclusions about potentially relevant categories that may be helpful in organizing such efforts for the officially designated HAPs. One step in such an analysis for the general organic chemicals on this list would be to review these structures in connection with a series of analyses of the implications of different structure groupings within the RD₅₀ data sets discussed in Section 3.1 above. The inorganic substances and previously established carcinogenic chemicals should be subject to distinct analyses oriented to available chronic carcinogenicity data for those chemicals in animals and people. In several places we have included single chemicals in more than one category where the same chemical had multiple potentially reactive groups or other characteristics.

In reviewing HAPs for organizing categories, the largest distinct category that emerged can be labeled as “famous carcinogens”—chemicals that are reasonably well known for carcinogenic activity (Table 3-5) and for which we must anticipate that significant analytical effort will be needed to evaluate what standards would be indicated to protect against possible cancer risks. This does not mean that noncancer effects will necessarily be secondary in the determination of needs for evaluation of risks and protective standards for these chemicals. Overall, we place 32 HAPs in this category—26 of these are organic chemicals, and 6 others are inorganic substances or minerals.

Table 3-5: Famous Carcinogens	
Organic Chemicals	Inorganic Substances or Minerals
2-Acetylaminofluorene	Asbestos (Friable)
Acrylamide	Beryllium Compounds
4-Aminobiphenyl	Inorganic Arsenic Compounds
Benzene	Nickel Compounds
Benzidine	Radionuclides
Bis(2-Chloroethyl) Ether	Chromium Compounds (Hexavalent Only)
Bis(Chloromethyl) Ether	
1,3-Butadiene	
2-Chlor-1,3-Butadiene	
Chloromethyl Methyl Ether	
Chlorobenzilate	
Coke Oven Emissions	
1,2-Dibromo-3-Chloropropane (Dbcp)	
1,2-Dibromoethane	
3,3'-Dichlorobenzidine	
1,4-Dioxane	
Epichlorohydrin	
Ethyleneimine	
Formaldehyde	
Hexamethylphosphoramide	
4,4'-Methylenebis(2-Chloroaniline)	
Polychlorinated Biphenyls	
Polycyclic Organic Matter (POM)	
Propane Sultone	
Urethane	
Vinyl Chloride	

Another large grouping (26 chemicals) is chlorinated aliphatic compounds with small molecular weight (Table 3-6); or compounds with an aromatic ring and a chlorinated side chain. Many of these will also need evaluation for possible carcinogenic effects. Also meriting related evaluation are 12 compounds that are possible or known metabolic precursors of epoxides. Similarly, possible carcinogenesis evaluations are needed for eight compounds that are either preformed epoxides themselves or are lactones or anhydrides, which are also reactive.

Table 3-6: Chemicals That Need Evaluation for Possible Carcinogenic Effects		
Chlorinated Aliphatic Compounds with Small Molecular Weight	Possible or Known Metabolic Precursors of Epoxides	Preformed Epoxides, Lactones, or Reactive Anhydrides
Allyl Chloride	Allyl Chloride	1,2-Butylene Oxide
Benzoic Trichloride	Campechlor (Mixture of Chlorinated Camphenes Containing 67-69% Chlorine)	Beta-Propiolactone
Benzyl Chloride	1,1-Dichloroethylene	1,2-Butylene Oxide
Campechlor (Mixture of Chlorinated Camphenes Containing 67-69% Chlorine)	1,3-Dichloropropene (Mixed Isomers)	Ethylene Oxide
Captan	Acrolein	Maleic Anhydride
Carbon Tetrachloride	Acrylic Acid	Phthalic Anhydride
Chloroacetic Acid	Acrylonitrile	Propylene Oxide
2-Chloroacetophenone	Ethyl Acrylate	Styrene Oxide
Chloroethane	Methyl Methacrylate	
Chloroform	Propyleneimine	
Chloromethane	Styrene	
1,1-Dichloroethane	Vinyl Acetate	
1,2-Dichloroethane	Vinyl Bromide	
1,1-Dichloroethylene		
Dichloromethane		
1,2-Dichloropropane		
1,3-Dichloropropene (Mixed Isomers)		
Dimethylcarbamoyl Chloride		
Hexachloro-1,3-Butadiene		
Hexachloroethane		
1,1,1,2-Tetrachloroethane		
Tetrachloroethylene		
Trans-1,3-Dichloropropene		
1,1,1-Trichloroethane		
1,1,2-Trichloroethane		
Trichloroethylene		

Finally, a chemical category that is associated with bladder carcinogenesis (Table 3-7) in many cases is the aromatic amines, of which there are 12.

Table 3-7: Chemicals Associated with Bladder Carcinogenesis
Aromatic Amines
2-Acetylaminofluorene
Aniline
(1,1'-Biphenyl)-4,4'-Diamine, 3,3'-Dimethyl- Chloramben
2,4-Diaminotoluene
3,3'-Dimethoxybenzidine
4-Dimethylaminoazobenzene
4,4'-Methylenedianiline
N,N-Dimethylaniline
O-Anisidine
O-Toluidine
P-Phenylenediamine
Quinoline

More generic large categories (Table 3-8) include 14 ring-chlorinated aromatic compounds, 8 “other aromatics” (no amino or halo group on the ring), 5 simple phenols (aromatic alcohols without a nitro or amino group) and 5 small molecular weight ketones and ethers.

Table 3-8: Other Generic Large Categories			
Ring-Chlorinated Aromatic Compounds	“Other Aromatics” (no amino or halo group on the ring):	Simple Phenols (aromatic alcohols without a nitro or amino group)	Small Molecular Weight Ketones and Ethers
Chloramben Chlordane Chlorobenzene Chlorobenzilate 2,4-D DDE 1,4-Dichlorobenzene Hexachlorobenzene Pentachlorophenol Quintozone 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) 1,2,4-Trichlorobenzene 2,4,6-Trichlorophenol 2,4,5-Trichlorophenol	Biphenyl Cumene (Isopropyl Benzene) Ethylbenzene Methoxychlor Naphthalene Propoxur Toluene Xylene (Isomers And Mixture)	Catechol Cresol (individual and mixed isomers) Phenol	1,4-Dioxane 2-Chloroacetophenone Acetophenone Methyl Tert-Butyl Ether Isophorone

In addition to these, we also created a substantial number of categories that proved to contain a smaller number of members in the air toxics listing. Even a category with only a few members on the list may nonetheless eventually be helpful for evaluation using the RD₅₀ database or other compilations of information. The smaller categories of organic compounds (Table 3-9) include five aliphatic azo, nitroso, or other reactive nitrogen compounds (many posing likely carcinogenic hazards), four isocyanates (many probably posing immune sensitization hazards), four hydrazines and alkylhydrazines, four chlorinated aliphatic ring compounds (at least three primarily used as insecticides), one small molecular weight nitroaliphatic compound, six brominated or iodinated aliphatic compounds (at least some with major concerns for carcinogenesis because of direct or indirect alkylating activity), four phthalates, three bi- or polycyclic aromatics and mixtures, three organic sulfur compounds (not elsewhere classified) of which the first two are related but the third might well be substantially distinguished, three acetylcholinesterase inhibitors (nerve agents primarily used as insecticides), three nitrophenols, three esters with possible alkylating activity, two small molecular weight aliphatic amides, and one small molecular weight aliphatic amine.

Table 3-9: Smaller Categories of Organic Compounds	
Category	Organic Compounds
Aliphatic Azo, Nitroso, Other Reactive Nitrogen Compounds	Ethyleneimine Diazomethane Methanamine, N-Methyl-N-Nitroso N-Nitroso-N-Methylurea N-Nitrosomorpholine
Isocyanates	Hexamethylene-1,6-Diisocyanate Methyl Isocyanate 1,1'-Methylenebis(4-Isocyanatobenzene) Toluene-2,4-Diisocyanate
Hydrazines And Alkylhydrazines	1,1-Dimethyl Hydrazine 1,2-Diphenylhydrazine Hydrazine Methyl Hydrazine
Chlorinated Aliphatic Ring Compounds	Chlordane Gamma-Lindane Heptachlor Hexachlorocyclopentadiene
Small Molecular Weight Nitroaliphatic Compound	2-Nitropropane
Brominated or Iodinated Aliphatic Compounds	1,2-Dibromo-3-Chloropropane (Dbcp) 1,2-Dibromoethane Vinyl Bromide Methyl Bromide Methyl Iodide Tribromomethane
Phthalates	Captan Bis(2-Ethylhexyl)Phthalate Dibutyl Phthalate Dimethyl Phthalate
Bi- or Polycyclic aromatics and Mixtures	Coke Oven Emissions Polycyclic Organic Matter (Pom) Dibenzofuran
Organic Sulfur Compounds	Carbon Disulfide Carbonyl Sulfide Ethylene Thiourea
Acetylcholinesterase Inhibitors	Carbaryl Dichlorvos Parathion
Nitrophenols	4,6-Dinitro-O-Cresol 2,4-Dinitrophenol 4-Nitrophenol
Esters With Possible Alkylating Activity	Dichlorvos Diethyl Sulfate Dimethyl Sulfate
Small Molecular Weight Aliphatic Amides	Acetamide N,N-Dimethylformamide
Small Molecular Weight Aliphatic Amine	Triethylamine

Among inorganic categories (Table 3-10), we found four metal kidney and/or cardiovascular toxins, two compounds with special immune-mediated effects, three other metals and metallic substances, one other nonmetal/nonmetallic substance, and one other/miscellaneous substance.

Table 3-10: Inorganic Categories				
Metal kidney and/or cardiovascular toxins	Special immune mediated effects	Other metals and metallic substances	Other nonmetal/nonmetallic	Other/miscellaneous
Antimony Compounds Cadmium Compounds Cobalt Compounds Lead Compounds	Beryllium Compounds Toluene-2,4-Diisocyanate	Manganese Compounds Mercury Compounds Titanium Tetrachloride	Selenium Compounds	Mineral Fibers

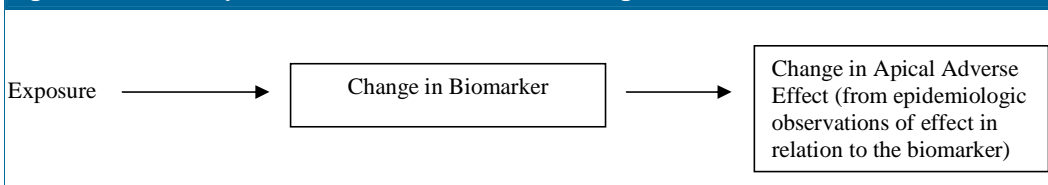
4 The Promise of Early-Effect Biomarkers for Two-Step Assessments of Risks

One of the most notable accomplishments in the past quarter century of risk assessment has been the use of physiologically based pharmacokinetic (PBPK) modeling to relate exposures to a pollutant of interest to a corresponding toxic response. Use of early-effect biomarkers to predict risks of apical effects²⁰ of ultimate concern is an additional two-step process in the same tradition:

- (1) Assess relationships between exposure to the chemical of interest and the intermediate biomarker of effect using limited chemical-specific information.
- (2) Assess relationships between the biomarker of effect and ultimate endpoints of concern from general epidemiological data.

This risk assessment process is also depicted in Figure 4-1.

Figure 4-1: Two-Step Risk Assessment Process Utilizing Biomarkers



Modeling risk using early-effect biomarkers continues the process of reducing a complex causal pathway to a series of discrete steps that can be separately observed and analyzed with quantitative modeling tools. Changes in apical effects from a toxicant exposure are inferred by first assessing the effect of the toxicant on changes in the distribution of the early-effect biomarker in the exposed population. Then the assessor uses pre-established relationships between the biomarker and risks of apical adverse effects to infer what changes in the incidence and/or severity of the ultimate effects of concern are likely to result from the toxicant-induced changes in the population distribution of the levels of the biomarker. For example, effects of an air toxic on the population distribution of the levels of a traditional cardiovascular risk factor such as serum cholesterol are combined with relationships between cholesterol levels and rates of myocardial infarctions to estimate the heart attack risk likely to be associated with the air toxic at specific levels of exposure. Additional examples of early-effect biomarkers that can be useful predictors of important human health endpoints include:

- Fetal growth restriction indicated by birth weight differences as predictors of infant mortality (Ananth and Platt 2004; Pedersen et al. 2007) and other adverse effects later in life (Lapidus et al. 2008; Kajantie et al. 2005).
- Semen quality differences as predictors of male fertility differences (Ibérico et al. 2004; Steinberger and Rodrigues-Rigau 1983; Rhemrev et al. 2001).

²⁰ An apical endpoint is an observable outcome in a whole organism, such as a clinical sign or pathologic state that is indicative of a disease state that can result from exposure to a toxicant (NRC 2007).

- Differences in traditional cardiovascular risk factors (e.g., blood pressure, serum cholesterol) and some newer risk factors (heart rate variability, serum fibrinogen, and some lung functions) as predictors of cardiovascular mortality (Kannel et al. 2003; Scheltens et al. 2008; Stein et al. 2008; Knuiman et al. 1999; Folsom et al. 1997; Tsuji et al. 1994).
- Iodide deficiency and thyroid hormone parameters in young children as predictors of neurodevelopmental impairment (Delange 2001; Santiago-Fernandez et al. 2004; Pop et al. 2003; Haddow et al. 1999; Pineda-Lucatero et al. 2008).

Early-effect biomarkers have another important role that they do not share with PBPK modeling. They serve to aggregate the influences of multiple stressors—both chemical and non-chemical—that act on the same biological system. Thus, birth weight can reflect effects of cigarette smoke, nutritional deficiencies, and effects from living at high altitudes, as well as any restriction of fetal growth from toxicity. Blood pressure is likely to reflect the integrated effects of salt intake, impairments of kidney function, and psychological stress from working as an air traffic controller. Motile sperm counts likely reflect both hormonal effects and the adverse effects of elevated testicular temperature from tight-fitting underwear. Thus, the use of intermediate effect biomarkers can enable risk assessors to implicitly take account of the host of interacting processes that affect health outcomes in the diverse human population, at least for a portion of the pathway to ultimate end effects. In **Section 4.1** we discuss some of the caveats to the biomarker approach, and in **Section 4.2** we illustrate an example of applying this approach to HAPs.

4.1 Caveats to the Biomarker Approach

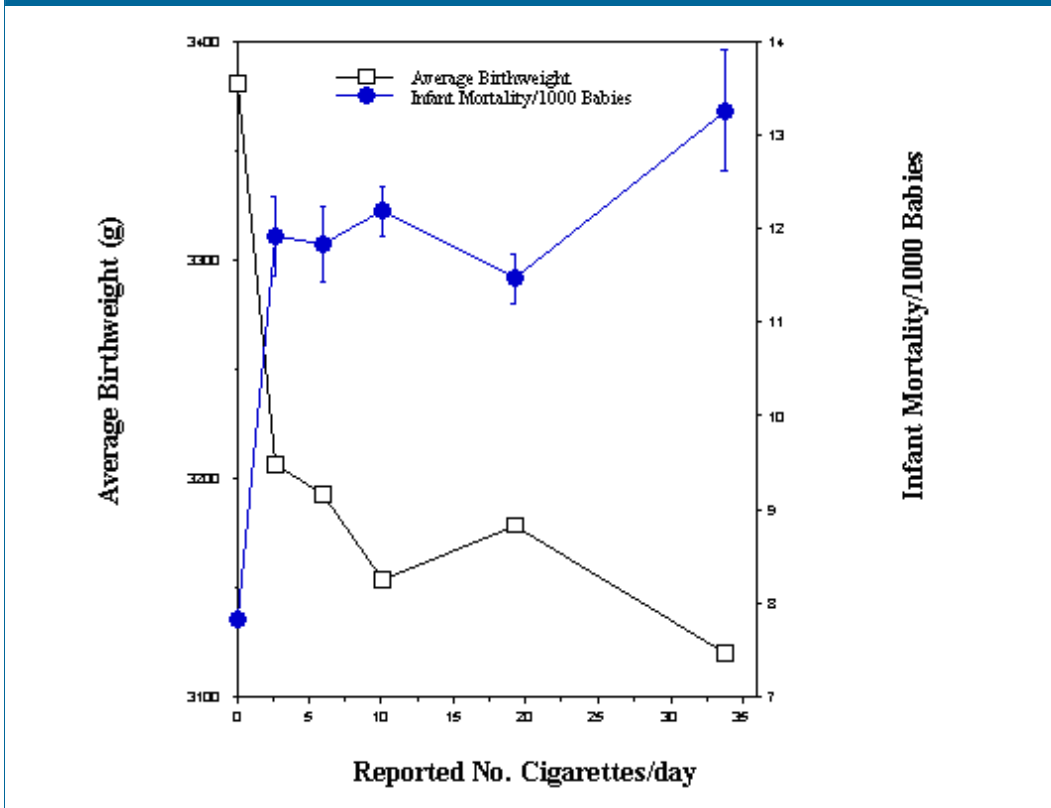
There is a need for caution and caveats to using this approach. The accuracy of any projection of risk with the aid of intermediate effect biomarkers depends on how well existing observations and inferred quantitative relationships for each of the two steps reflect real causal processes. For the most part, epidemiological studies yield measures of the association between variables. These associations may reflect causal processes, but may also be the result of confounding or other bias, thus causing an overestimate of the adverse effects from the exposure. On the other hand, quantitative relationships derived from epidemiological studies are notoriously subject to distortion via imperfection in the measurement of exposures (or, in the case of relationships between a biomarker and an apical response, measurements of the biomarker). There has been some use of “errors in variables” models (Stayner et al. 2003; Richardson et al. 2004; Brown et al. 2004; Choi 2000; Kulathinal et al. 2002; Siebert et al. 2001) to attempt to correct for this kind of problem, but such innovative analytical models are not yet common in epidemiological studies. These types of errors would tend to cause risk assessors to underpredict adverse effects from biomarker exposures. Selection biases—most notably the “healthy worker effect” (Steenland and Stayner 1991) and the “healthy worker survivor effect” (Steenland et al. 1996; Kolstad and Olsen 1999)—can cause similar distortions. Thus, although we think the “two-step” risk assessment approach outlined here is promising, and may be a reasonable approach for initial estimates of effects, risk conclusions derived using this type of method will be subject to extensive modification with the development of improved observations and modeling tools.

Decision-making for pharmaceuticals has developed a generic solution to the problem of resolving causal versus non-causal associations: the double-blind, randomized, controlled clinical trial. Selection effects are minimized via the randomization process, and distortions from dosage measurement errors are avoided by the deliberate supply of measured amounts of test chemicals to the clinical subjects.

Such measures will seldom if ever be practically available for risk assessment and decision-making on environmental chemical exposures. However, environmental health researchers can take a few different approaches to evaluate and reduce uncertainties in the two-step quantitative inferences of risk:

- Take advantage of “experiments of nature” such as episodes of unusually large or distinctive population exposures (e.g., the air pollution episodes observed in Donora, Pennsylvania, and the infamous “London fog,” which first strongly indicated effects of air pollutants on short-term mortality changes) or the temporary closing of a specific facility responsible for an important fraction of the pollutant exposures of known types in a specific area.
- Look for situations in which an exposure causes measurable changes in both the intermediate biomarker and an apical outcome of ultimate concern. Ask, “Are the changes in the incidence of the apical effect accurately predicted by the changes brought about in the early-effect biomarker by the responsible exposure?” and “Is the pattern of change in the apical effect with dose similar enough to the pattern of change in the biomarker with dose to suggest a causal relationship? For example, Figure 4-2 juxtaposes the dose-response relationship between reported cigarette smoking and population-average birth weights, and the similar dose-response relationship between reported cigarette smoking and infant mortality. Both appear to show saturation-type dose-response relationships, suggesting that the causal factor(s) for both effects may be the product of a saturable active metabolism process or a saturable receptor binding process. Quantitatively, we have found that the infant mortality effect of cigarette smoking can be understood as a combination of effects of smoking on fetal growth restriction and a shortened period of gestation (unpublished observations). More formally, an analysis of 10 million singleton live U.S. births between 1998 and 2000 by Ananth and Platt (2004) finds that “the effect of maternal smoking on neonatal mortality is largely mediated through reduced fetal growth.”
- Quantitatively evaluate the experience with the predictiveness of cardiovascular risk factors for cardiovascular mortality and morbidity. Beginning in the 1950s and 1960s there were a substantial number of prospective cardiovascular disease studies that developed predictive risk relationships using multiple logistic risk equations (Truett et al. 1967). Subsequently, the drug industry has developed a large number of pharmaceutical interventions designed to reduce the levels of specific risk factors. Generally, these interventions are likely to have been evaluated both in terms of their effects on risk factors and in terms of mortality/morbidity reductions. Risk assessors of today therefore have an opportunity to ask, for each biomarker/“risk factor,” “Was the degree of mortality/morbidity reduction achieved via the pharmaceutical intervention larger or smaller than what would have been predicted from the change in the risk factor(s) brought about by the intervention?”

Figure 4-2: Relationships Between Reported Cigarettes Smoked per Day, Average Birth Weight, and Infant Mortality



Data Source: National Center for Health Statistics, 1996. 1990 Birth Cohort Linked Birth/Infant Death Data Set, NCHS CD-ROM Series 20 No. 6, SETS Version 1.22a, issued May 1996.
 Figure previously appeared in Hattis, D. "Role Of Dosimetric Scaling And Species Extrapolation In Evaluating Risks Across Life Stages. IV. Pharmacodynamic Dosimetric Considerations." Draft Report to the U.S. Environmental Protection Agency Under RFQ No. DC-03-0000, January 2004.

4.2 Application of the Biomarker Approach to HAPs

While we do not explore in detail the potential of all the various early-effect biomarkers that may shed light on the risks of HAPs, studies of birth weight changes in relation to air pollutant exposures deserve special mention and are presented as an example illustrating the application of this promising approach.

Recently, Bell et al. (2007) published an epidemiological study that extends previous work from California and elsewhere on apparent relationships between exposures to various criteria air pollutants at ambient levels and birth weights in a large number of babies born in Connecticut and Massachusetts. This cohort is much larger than has been previously evaluated. Perhaps because of the large cohort size, even though exposure assessments were relatively basic (county averages), apparent effects of various pollutants were quantified with reasonably narrow confidence limits (Table 4-1). The findings can be brought together into a consistent framework for comparative analysis of

potencies and approximate indicated population-level effects (Table 4-2). This type of information should be combined with other data through a thorough meta-analysis based on multiple good studies. Specifically, if the fetal growth restriction potency of exposures to different air toxics can be related to the potencies of these criteria pollutants through comparative animal experiments, then it seems possible to make at least approximate evaluations of population fetal growth restriction hazards of different air pollutants and comparative cost effectiveness of regulations and other policy actions (e.g., changes to the mix of fuels used in automobiles or changes in emissions from greater fuel efficiency).

Appendix A, drawn from our previous work (Hattis 2004), provides more extensive background information on the implications of birth weight changes for infant mortality and longer-term risks in adult life, such as Type 2 diabetes. Appendix B provides additional background on implications of iodide uptake changes (as might be produced by perchlorate exposure) for neurodevelopmental impairment.

Table 4-1: Basic Birth Weight Reduction Results Based on County-Average Air Pollutant Exposures During Gestation*

Air Pollutant	Grams Reduction Birth Weight per Interquartile Exposure Range	Lower 95% Confidence Limit (g)	Upper 95% Confidence Limit (g)	Mean and Std Dev Exposure	Interquartile Exposure Range	Exposure units
NO ₂	8.9	7	10.8	17.4 ± 5.0	4.8	ppb
CO	16.2	12.6	19.7	656 ± 180	303	ppb
SO ₂	0.9	-2.6	4.4	4.7 ± 1.2	1.6	ppb
PM ₁₀	8.2	5.3	11.1	22.3 ± 5.3	7.4	µg/m ³
PM _{2.5}	14.7	12.3	17.1	11.9 ± 1.6	2.2	µg/m ³

*for 358,504 Babies in Massachusetts and Connecticut, Evaluated with Single-Pollutant Models, Controlling for Confounders

Source: Bell et al. (2007).

Table 4-2: Implications for Population Aggregate Birth Weight Changes*

Air Pollutant	Indicated Potency in g Birth Weight Reduction per ppb Gas or (for Particles) µg/m ³	Lower 95% Confidence Limit on Potency	Upper 95% Confidence Limit on Potency	Suggested Population Aggregate Effect (g/baby) (Potency x Mean Exposure)
NO ₂	1.85	1.46	2.25	32
CO	0.053	0.042	0.065	35
SO ₂	Non-significant			
PM ₁₀	1.1	0.7	1.5	25
PM _{2.5}	6.7	5.6	7.8	80

Source: Bell et al. (2007)

* Results for Pollutant Potencies (gram Reduction in Mean Baby Weights Per Unit Exposure During Gestation) and Suggested Population Aggregate Impacts on Birth Weights

5 Conclusions

The suggestions and illustrations in the previous sections indicate that probabilistic modeling tools can be used to systematically analyze the risks posed by HAPs. Even in the best of cases there will be considerable uncertainty about the quantitative level of risk posed by specific agents. However, some of the important uncertainties can be identified and approximately quantified. Moreover, the uncertainty quantification itself can be used to design targeted research activities aimed at clarifying the most potentially significant opportunities for public health protection, as assessed by (1) existing exposures, (2) analyses of likely relative potencies, and (3) analyses of the potential for uncertainty reduction via specific types of research.

Most importantly, the adoption of this kind of approach, as recently recommended by the NRC (2008), will open up a regulatory system to inputs from new kinds of information. This openness itself will attract new investigative resources to the field of quantitative risk estimation. Thus, although the adoption of this approach will not provide “final” quantitative answers with narrow confidence limits on the posed by specific air toxics, there is every prospect that it will allow the system for assessing and managing health risks from toxicant exposures to develop in ways that cannot be fully anticipated at present.

6 References

- Ad Hoc Working Group on Risk Assessment, Federal Coordinating Council for Science, Engineering and Technology (FCCSET). 1992. Draft report: A cross-species scaling factor for carcinogen risk assessment based on equivalence of mg/kg^{3/4}/day. Federal Register 57:24152-24173, June 5
- Alarie Y. 1966. Irritating properties of airborne materials to the upper respiratory tract. Arch Environ Health. 1966 Oct;13(4):433-49.
- Alarie Y. 1981. Bioassay for evaluating the potency of airborne sensory irritants and predicting acceptable levels of exposure in man. Food Cosmet Toxicol 19(5):623–626.
- Alarie Y. 1998. Computer-based bioassay for evaluation of sensory irritation of airborne chemicals and its limit of detection. Arch Toxicol. 72(5):277-82.
- Alarie Y, Schaper, M, Nielsen GD, Abraham MH. 1998. Structure-activity relationships of volatile organic chemicals as sensory irritants. Arch Toxicol. 72:125-40.
- Ananth CV, Platt RW. 2004. Reexamining the effects of gestational age, fetal growth, and maternal smoking on neonatal mortality. BMC Pregnancy Childbirth. 4(1):22, doi:10.1186/1471-2393-4-22.
- Bakke, S., Baste, V., and Gulsvik, A. 1991. Bronchial responsiveness in a Norwegian community. Am. Rev. Respir. Dis. 143:317-322.
- Baird SJS, Cohen JT, Graham JD, Shlyakhter AI, Evans JS. 1996. Noncancer risk assessment: A probabilistic alternative to current practice, Hum. Ecol. Risk Assess. 2:78-99.
- Balmes, J. R., Aris, R. M., Chen, L. L., Scannell, C., Tager, I. B., Finkbeiner, W., Christian, D., Keller, T., Hearne, P. Q., Ferrando, R., and Welch, B. 1997. Airway inflammation and responsiveness to ozone in normal and asthmatic subjects. In *Effects of ozone on normal and potentially sensitive subjects*, Research Report Number 78, June. Cambridge, Mass: Health Effects Institute.
- Bell ML, Ebisu K, Belanger K. 2007. Ambient Air Pollution and Low Birth Weight in Connecticut and Massachusetts. Environmental Health Perspectives 115(7): 1118-1124
- Bos PM, Busschers M, Arts JH. 2002. Evaluation of the sensory irritation test (Alarie test) for the assessment of respiratory tract irritation. J Occup Environ Med. 44(10):968-976.
- Boylstein LA, Luo J, Stock MF, Alarie Y. An attempt to define a just detectable effect for airborne chemicals on the respiratory tract in mice. Arch Toxicol 70(9):567–578.
- Brown SC, Schonbeck MF, McClure D, Baron AE, Navidi WC, Byers T, Rutenber AJ. 2004. Lung cancer and internal lung doses among plutonium workers at the Rocky Flats Plant: a case-control study. Am J Epidemiol. 160(2):163-72.
- Choi BC. 2000. A technique to re-assess epidemiologic evidence in light of the healthy worker effect: the case of firefighting and heart disease. J Occup. Environ Med. 42(10):1021-34
- Cramer GM, Ford RA. 1978. Estimation of toxic hazard—A decision tree approach. Food and Cosmetics Toxicology 16:255-276.
- Danesh, J., Collins, R. Appleby, P., Peto, R. “Association of Fibrinogen, C-reactive Protein, Albumin, or Leukocyte Count with Coronary Heart Disease,” JAMA, V. 279, #18, p. 1477-82, May 13, 1998.
- Delange F. 2001. Iodine deficiency as a cause of brain damage. Postgraduate Medical Journal 77:217-220.

- Evans JS, Baird SJS. 1998. Accounting for missing data in noncancer risk assessment, *Human and Ecological Risk Assessment* 4:291-317.
- Evans JS, Rhomberg LR, Williams PL, Wilson AM, Baird SJ. 2001. Reproductive and developmental risks from ethylene oxide: a probabilistic characterization of possible regulatory thresholds. *Risk Anal.* 21(4):697-717.
- Folsom, A. R., Wu, K. K., Rosamond, W. D., Sharrett, A. R., and Chambless, L. E. 1997. "Prospective study of hemostatic factors and incidence of coronary heart disease. The Atherosclerosis Risk in Communities (ARIC) study." *Circulation* 96:1102-1108.
- Frantík E, Hornychová M, Horváth M. 1994. Relative acute neurotoxicity of solvents: isoeffective air concentrations of 48 compounds evaluated in rats and mice. *Environ Res.* 66(2):173-185.
- Grant RL, Kadlubar BJ, Erraguntla NK, Honeycutt M. 2007. Evaluation of acute inhalation toxicity for chemicals with limited toxicity information. *Regul Toxicol Pharmacol.* 47(3):261-273.
- Haddow JE, Palomaki GE, Allan WC, Williams JR, Knight GJ, Gagnon J, O'Heir CE, Mitchell ML, Hermos RJ, Waisbren SE, Faix JD, Klein RZ. 1999. Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. *N Engl J Med.* 341(8):549-55.
- Hanania, N. A., Tarlo, S. M., Silverman, F., Urch, B., Senathirajah, N., Zamel, N., and Corey, P. 1998. Effect of exposure to low levels of ozone on the response to inhaled allergen in allergic asthmatic patients. *Chest* 114,:752-756.
- Hattis D. 1998. Strategies for assessing human variability in susceptibility, and using variability to infer human risks. In *Human Variability in Response to Chemical Exposure: Measures, Modeling, and Risk Assessment*, D. A. Neumann and C. A. Kimmel, eds., CRC Press, Boca Raton, FL, pp. 27-57, 1998.
- Hattis, D. 2004. Role of dosimetric scaling and species extrapolation in evaluating risks across life stages. IV. Pharmacodynamic dosimetric considerations. Draft Report to the U.S. Environmental Protection Agency Under RFQ No. DC-03-0000.
- Hattis, D. 2008. Distributional analyses for children's inhalation risk assessments. *Journal of Toxicology and Environmental Health*, 71(3):218-226.
- Hattis D, Banati P., Goble R. 1999. Distributions of individual susceptibility among humans for toxic effects—for what fraction of which kinds of chemicals and effects does the traditional 10-fold factor provide how much protection?" *Annals of the New York Academy of Sciences* 895:286-316.
- Hattis, D., Baird, S., and Goble, R. 2002. "A Straw Man Proposal for a Quantitative Definition of the RfD," in Final Technical Report, U.S. Environmental Protection Agency STAR grant # R825360, "Human Variability in Parameters Potentially Related to Susceptibility for Noncancer Risks," Full version available on the web at <http://www2.clarku.edu/faculty/dhattis>; shortened version *Drug and Chemical Toxicology*, Vol. 25, pp. 403-436 (2002).
- Hattis D, Crofton K. 1995. Use of Biological Markers of Causal Mechanisms in the Quantitative Assessment of Neurotoxic Risks. Chapter 53 in *Handbook of Neurotoxicology*, Vol. 3 *Approaches and Methodologies*. Lewis Chang and William Slikker, eds., Academic Press, 1995, pp. 789-803.
- Hattis D, Lynch MK. 2007. "Empirically Observed Distributions of Pharmacokinetic and Pharmacodynamic Variability in Humans—Implications for the Derivation of Single Point Component Uncertainty Factors Providing Equivalent Protection as Existing RfDs." In

- Toxicokinetics in Risk Assessment*, J. C. Lipscomb and E. V. Ohanian, eds., Informa Healthcare USA, Inc., 2007, pp. 69-93.
- Hattis D, Shapiro K. 1990. Analysis of dose/time/response relationships for chronic toxic effects—the case of acrylamide, *Neurotoxicology* 11:219-236.
- Hau KM, Connell DW, Richardson BJ. 2002. A study of the biological partitioning behavior of n-alkanes and n-alkanols in causing anesthetic effects. *Regul Toxicol Pharmacol.* 35(2 Pt 1):273-279.
- Horstman D, Roger LJ, Kehrl H, Hazucha M. 1986. Airway sensitivity of asthmatics to sulfur dioxide. *Toxicol Ind Health.* 2(3):289-298.
- Ibérico G, Vioque J, Ariza N, Lozano JM, Roca M, Llácer J, Bernabeu R. 2004. Analysis of factors influencing pregnancy rates in homologous intrauterine insemination. *Fertil Steril.* 81(5):1308-1313.
- Kajantie E, Osmond C, Barker DJ, Forsén T, Phillips DI, Eriksson JG. 2005. Size at birth as a predictor of mortality in adulthood: a follow-up of 350 000 person-years. *Int J Epidemiol.* 34(3):655-663.
- Kannel WB, Vasan RS, Levy D. 2003. Is the relation of systolic blood pressure to risk of cardiovascular disease continuous and graded, or are there critical values? *Hypertension.* 42(4):453-456.
- Kasanen JP, Pasanen AL, Pasanen P, Liesivuori J, Kosma VM, Alarie Y. 1998. Stereospecificity of the sensory irritation receptor for nonreactive chemicals illustrated by pinene enantiomers. *Arch Toxicol.* 1998 72(8):514-523.
- Knuiman MW, James AL, Divitini ML, Ryan G, Bartholomew HC, Musk AW. 1999. Lung function, respiratory symptoms, and mortality: results from the Busselton health study. *Ann. Epidemiol.* 9:297-306.
- Kolstad HA, Olsen J. 1999. Why do short term workers have high mortality? *Am J Epidemiol.* 149(4):347-352.
- Kristiansen U, Nielsen GD. 1988. Activation of the sensory irritant receptor by C7-C11 n-alkanes. *Arch Toxicol.* 1988;61(6):419-25.
- Kulathinal SB, Kuulasmaa K, Gasbarra D. 2002. Estimation of an errors-in-variables regression model when the variances of the measurement errors vary between the observations. *Stat Med.* 2002 Apr 30;21(8):1089-101.
- Kuwabara Y, Alexeeff GV, Broadwin R, Salmon AG. 2007. Evaluation and application of the RD₅₀ for determining acceptable exposure levels of airborne sensory irritants for the general public. *Environ Health Perspect.* 115(11):1609-1616.
- Lapidus L, Andersson SW, Bengtsson C, Björkelund C, Rossander-Hulthén L, Lissner L. 2008. Weight and length at birth and their relationship to diabetes incidence and all-cause mortality—a 32-year follow-up of the population study of women in Gothenburg, Sweden. *Prim Care Diabetes.* 2(3):127-133.
- Lehman AJ, Fitzhugh OG. 1954. 100-fold margin of safety, *Assoc. Food Drug Off. U.S. Q. Bull.* 18:33-35.
- Lynch, M. K. (2007). Interindividual variability in response to xenobiotics: Implications for the derivation of the reference dose, Boston University AAT 3279941 134.

- McCarthy MC, O'Brien TE, Charrier JG, Hafner HR. 2009. Characterization of the Chronic Risk and Hazard of Hazardous Air Pollutants in the United States Using Ambient Monitoring Data. *Environmental Health Perspectives*, available at doi: 10.1289/ehp.11861 (<http://dx.doi.org/>) Online 9 January 2009.
- Mordenti J. 1986. Man versus beast: pharmacokinetic scaling in mammals. *J. Pharmaceut. Sci.*, 75:1028-1040.
- Munro IC, Ford RA, Kennepohl E, Sprenger JG 1996. Correlation of structural class with no-observed-effect levels: A proposal for establishing a threshold of concern. *Food and Chemical Toxicology* 34:829-867.
- National Center for Health Statistics. 1996. 1990 Birth Cohort Linked Birth/Infant Death Data Set, NCHS CD-ROM Series 20 No. 6, SETS Version 1.22a, issued May 1996.
- National Institute for Occupational Safety and Health (NIOSH). 1982. Registry of Toxic Effects of Chemical Substances (1980 edition, volume 1) Washington, D.C. 20402: U. S. Department of Health and Human Services DHHS (NIOSH) Publication No. 81-116, Government Printing Office, Washington, D.C. February 1982.
- National Research Council Committee on Toxicity Testing and Assessment of Environmental Agents, 2007. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. National Academy of Sciences Press, Washington, DC, 2007.
- National Research Council Committee on Improving Risk Analysis Approaches Used by the U.S. EPA, 2008. *Science and Decisions: Improving Risk Analysis*. National Academy of Sciences Press, Washington, DC, 2008.
- NTP (National Toxicology Program) 2006. Studies of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) (CAS NO. 1746-01-6) in Female Harlan Sprague-Dawley Rats (Gavage Studies). Carcinogenic endpoints.
- Nessel CS, Lewis SC, Stauber KL Adgate JL. 1995. Subchronic to chronic exposure extrapolation: Toxicologic evidence for a reduced uncertainty factor, *Human and Ecological Risk Assessment* 1:516-526.
- Nielsen GD, Vinggaard AM. 1988. Sensory irritation and pulmonary irritation of C3-C7 n-alkylamines: mechanisms of receptor activation. *Pharmacol Toxicol.* 63(4):293-304.
- O'Connor, G., Sparrow, D. Taylor, D., Segal, M., and Weiss, S. 1987. Analysis of dose-response curves to methacholine: An approach suitable for population studies, *Am. Rev. Respir. Dis.* 136:1412-1417.
- Paoletti, P., Carrozzi, L., Viegi, G., Modena, P., Ballerin, L., Di Pede, F., Grado, L., Baldacci, S., Pedreschi, M., Vellutini, M., Paggiaro, P., Mammini, U., Fabbri, L., and Giuntini, C. 1995. Distribution of bronchial responsiveness in a general population: effect of sex, age, smoking, and level of pulmonary function, *Am. J. Resp. Crit. Care Med.* 151:1770-1777.
- Pedersen CB, Sun Y, Vestergaard M, Olsen J, Basso O. 2007. Assessing fetal growth impairments based on family data as a tool for identifying high-risk babies. An example with neonatal mortality. *BMC Pregnancy Childbirth.* 7:28.
- Pineda-Lucatero A, Avila-Jiménez L, Ramos-Hernández RI, Magos C, Martínez H. 2008. Iodine deficiency and its association with intelligence quotient in schoolchildren from Colima, Mexico. *Public Health Nutr.* 11(7):690-698.

- Pop VJ, Brouwers EP, Vader HL, Vulsmat T, van Baar AL, de Vijlder JJ. 2003. Maternal hypothyroxinaemia during early pregnancy and subsequent child development: a 3-year follow-up study. *Clin Endocrinol (Oxf)*. 59(3):282-288.
- Price PS, Keenan RE, Swartout JC. 2008. Characterizing interspecies uncertainty using data from studies of anti-neoplastic agents in animals and humans. *Toxicology and Applied Pharmacology* 233:64-70.
- Rhemrev JP, Lens JW, McDonnell J, Schoemaker J, Vermeiden JP. 2001. The postwash total progressively motile sperm cell count is a reliable predictor of total fertilization failure during in vitro fertilization treatment. *Fertil Steril*. 76(5):884-891.
- Richardson D, Wing S, Steenland K, McKelvey W. 2004. Time-related aspects of the healthy worker survivor effect. *Ann Epidemiol*. 14(9):633-9.
- Rhomberg LR, Wolff SK. (1998). Empirical scaling of single oral lethal doses across mammalian species based on a large database. *Risk Anal*. 18(6):741-753.
- Samoli E, Touloumi G, Schwartz J, Anderson HR, Schindler C, Forsberg B, Vigotti MA, Vonk J, Kosnik M, Skorkovsky J, Katsouyanni K. 2007. Short-term effects of carbon monoxide on mortality: An analysis within the APHEA project. *Environ Health Perspect* 115:1578-1583.
- Santiago-Fernandez P, Torres-Barahona R, Muela-Martinez JA, Rojo-Martinez G, Garcia Fuentes E, Garriga MJ, Leon AG, Soriguer F. 2004. Intelligence quotient and iodine intake: A cross-sectional study in children. *Clinical Endocrinology & Metabolism* 89:3851-3857.
- Scheltens T, Verschuren WM, Boshuizen HC, Hoes AW, Zuithoff NP, Bots ML, Grobbee DE. 2008. Estimation of cardiovascular risk: a comparison between the Framingham and the SCORE model in people under 60 years of age. *Eur J Cardiovasc Prev Rehabil*. 15(5):562-566.
- Sears M, Jones DT, Holdaway MD, Hewitt CJ, Flannery EM, Herbison GP, Silva PA. 1986. Prevalence of bronchial reactivity to inhaled methacholine in New Zealand children. *Thorax* 41:283-289.
- Šrám RJ, Binková B, Dejmeek J, Bobak M. 2005. Ambient air pollution and pregnancy outcomes: a review of the literature. *Environ Health Perspect* 113:375-382.
- Siebert U, Rothenbacher D, Daniel U, Brenner H. 2001. Demonstration of the healthy worker survivor effect in a cohort of workers in the construction industry. *Occup Environ Med*. 2001 Dec;58(12):774-9.
- Stayner L, Steenland K, Dosemeci M, Hertz-Picciotto I. 2003. Attenuation of exposure-response curves in occupational cohort studies at high exposure levels. *Scand J Work Environ Health*. 29(4):317-24.
- Steenland K, Deddens J, Salvan A, Stayner L. 1996. Negative bias in exposure-response trends in occupational studies: Modeling the healthy workers survivor effect. *Am. J. Epidemiology* 143(2):202-210.
- Steenland K, Stayner L. 1991. The importance of employment status in occupational cohort mortality studies. *Epidemiology* 2(6):418-423.
- Stein PK, Barzilay JI, Chaves PH, Mistretta SQ, Domitrovich PP, Gottdiener JS, Rich MW, Kleiger RE. 2008. Novel measures of heart rate variability predict cardiovascular mortality in older adults independent of traditional cardiovascular risk factors: the Cardiovascular Health Study (CHS). *J Cardiovasc Electrophysiol*. 19(11):1169-1174.
- Steinberger E, Rodrigues-Rigau LJ. 1983. The infertile couple. *J. Androl*. 4:111-118.

- Steinhagen WH, Barrow CS. 1984. Sensory irritation structure-activity study of inhaled aldehydes in B6C3F1 and Swiss-Webster mice. *Toxicol Appl Pharmacol.* 72(3):495-503.
- Tashkin, D. P., Altose, M. D., Connett, J. E., Kanner, R. E., Lee, W. W., and Wise, R. A. 1996. Methacholine reactivity predicts changes in lung function over time in smokers with early chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 153:1802-1811.
- ten Berge WF, Zwart A, Appema LM. 1986. Concentration-time mortality response relationship of irritant and systemically acting vapors and gases. *J. Hazardous Materials* 13:301-309.
- Travis CC, White RK. 1988. Interspecific scaling of toxicity data. *Risk Anal.* 8: 119-125.
- Travis CC, White RK, Ward RC. 1990. Interspecies extrapolation of pharmacokinetics. *J Theor Biol.* 1990 Feb 9;142(3):285-304.
- Truett J, Cornfield J, Kannel W. 1967. A multivariate analysis of the risk of coronary heart disease in Framingham. *J Chronic Dis.* 20(7):511-524.
- Tsuji H, Venditti FJ, Manders ES, Evans JC, Larson MG, Feldman CL, Levy D. 1994. Reduced heart rate variability and mortality risk in an elderly cohort: The Framingham heart study. *Circulation,* 90:878-883.
- U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. 2009. Risk and Exposure Assessment to Support the Review of the SO₂ Primary National Ambient Air Quality Standards: Second Draft, U.S. Environmental Protection Agency EPA-452/P-09-003, Research Triangle Park, North Carolina, March 2009, Appendix C, p. 3-4.
- Watanabe, K, Bois FY, Zeise L. 1992. Interspecies extrapolation: A reexamination of acute toxicity data. *Risk Analysis* 12:301-310.
- Weil CS, McCollister DD. 1963. Relationship between short and long-term feeding studies in designing an effective toxicity test, *Agriculture and Food Chemistry* 11:486-491.
- Whitney M, and L Ryan. 2009 in press. Bayesian model averaging. In Cooke, R.M. (ed) (2009) *Uncertainty Modeling in Dose Response: Bench Testing Environmental Toxicity*, Wiley, New York, ISBN: 978-0-470-44750-5. Prepublication version available at <http://www.rff.org/Events/Pages/Cooke-Uncertainty-Workshop.aspx>.
- Woodruff TJ, Zeise L, Axelrad DA, Guyton KZ, Janssen S, Miller GG, Schwartz JM, Alexeef G, Anderson H, Birnbaum L, Bois F, Cogliano VJ, Crofton K, Euling SY, Foster PMD, Germolec DR, Gray E, Hattis DB, Kyle AD, Luebke RW, Luster MI, Portier C, Rice DC, Solomon G, Vandenberg J, Zoeller RT. 2008. Moving upstream: Evaluating adverse upstream endpoints for improved risk assessment and decision-making. *Environmental Health Perspectives* 116:1657-1675.

Appendix A – Birth Weight in Humans and Fetal Growth Reduction in Animals—Model Biomarkers for Assessing Generic Reproductive Toxicity in Animals and People

Birth Weight and Infant Mortality

The concept of homeostatic controls can support the plausibility of a NOAEL for a particular person (an individual threshold) for a specific mechanism of damage. However, as mentioned earlier, this does not necessarily imply the existence of a population threshold (a dose so low as to be at, or below, the lowest threshold dose for any individual in a mixed population with diverse sensitivities). Specifically, there may be a finite expectation for individuals to be affected by even marginal exposures in cases where even without additional exposure some people have no “functional reserve capacity” to act as a buffer between the base health state and a state of at least marginally worse health. Two specific types of cases where this can happen are:

- Some individuals in the diverse population are already suffering from various kinds of pathological dysfunction in key parameters that may be marginally affected by different toxicants (e.g., a person undergoing a myocardial infarction may have a marginally expanded area of heart muscle death if the oxygen-carrying capacity of his or her blood is reduced by a marginal exposure to additional carbon monoxide).
- Some individuals are presently engaged in a task (e.g., running a 100-yard dash) that taxes some physiological capabilities to their limit, and marginal exposures to a toxicant marginally reduce those physiological capabilities.

One possible example of the second type is the possibly limiting mobilization of metabolic energy to sustain growth and development during gestation. Briefly, dose-response relationships for fetal growth inhibition in relation to a variety of toxicant exposures in experimental animals, and cigarette smoking in people, suggest that marginal exposures to toxicants may impose a metabolic “tax” on the developing fetus that causes changes in developmental progress. These changes, at least on a population basis, can have potentially significant implications for public health.

One of the most common types of effects observed in reproductive/developmental studies is an inhibition of fetal growth. The current practice is to express this effect in terms of the dose causing a particular increase in the percentage of animals that have growth that is more than a standard deviation below the control mean weight. This practice implies that this effect should be treated as if it were a classical individual threshold response, to be fed into the usual NOAEL/uncertainty factor system.

However, we first noticed in a 1988 analysis of developmental effects of glycol ethers (Ballew and Hattis 1989) that the observed pattern of dose-response relationships for fetal weight reductions differed considerably from the highly nonlinear observations of the frequency of typical teratogenic anomalies. Figure A-1 and Figure A-2, reproduced from that study, show the pattern of mean birth weight change in relation to dose observed for ethylene glycol ethyl ether (EGEE) and methylene glycol ethyl ether (MGEE) in different experimental animal systems.

Figure A-1: Loss in Fetal Weight in CD-1 Mice Exposed to EGEE via Oral Gavage (Data of Wier et al. 1987)

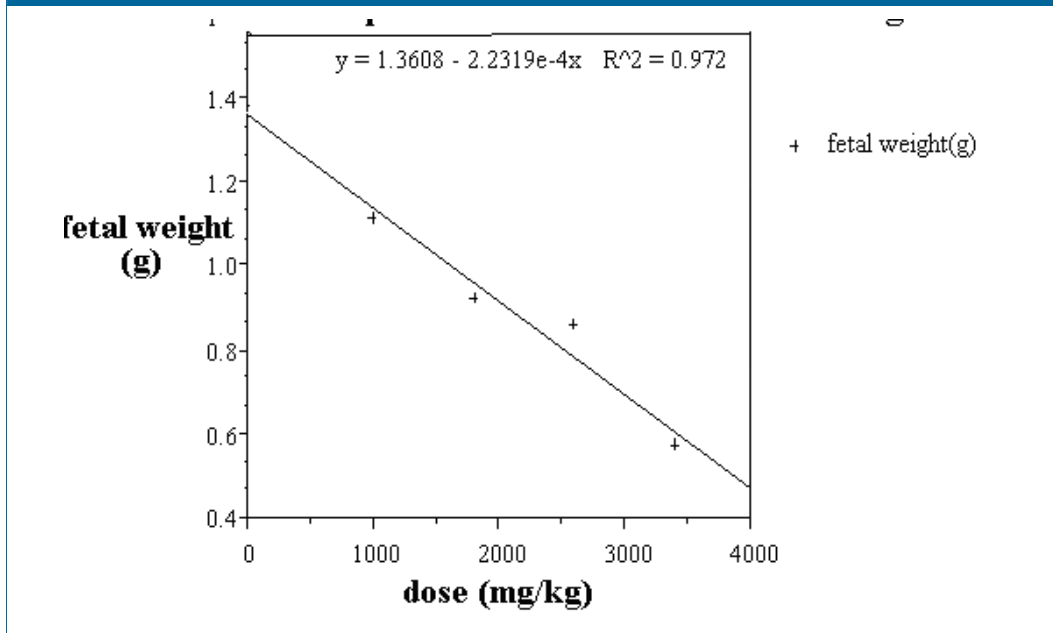
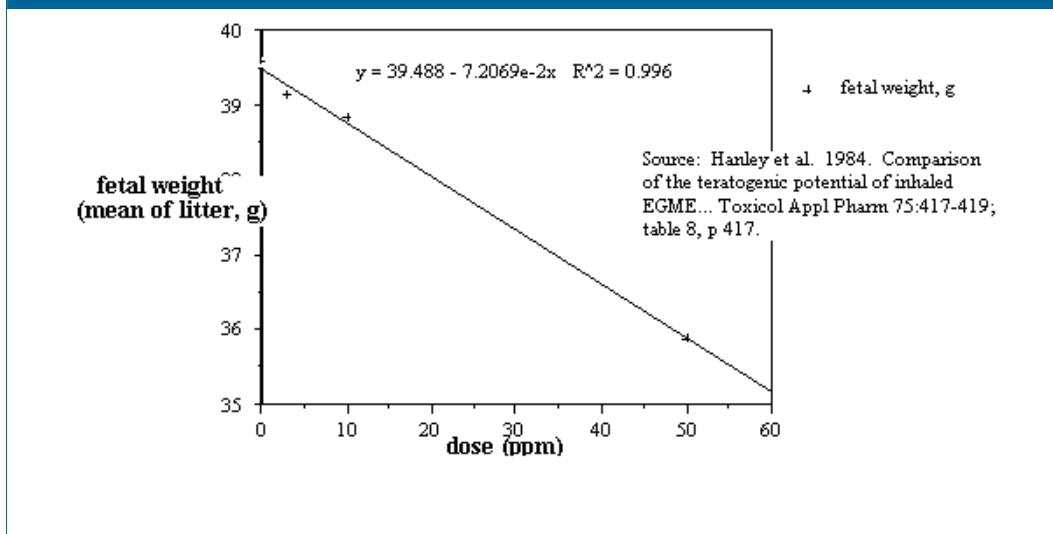


Figure A-2: Loss in Fetal Weight in New England White Rabbits Exposed to EGME via Inhalation

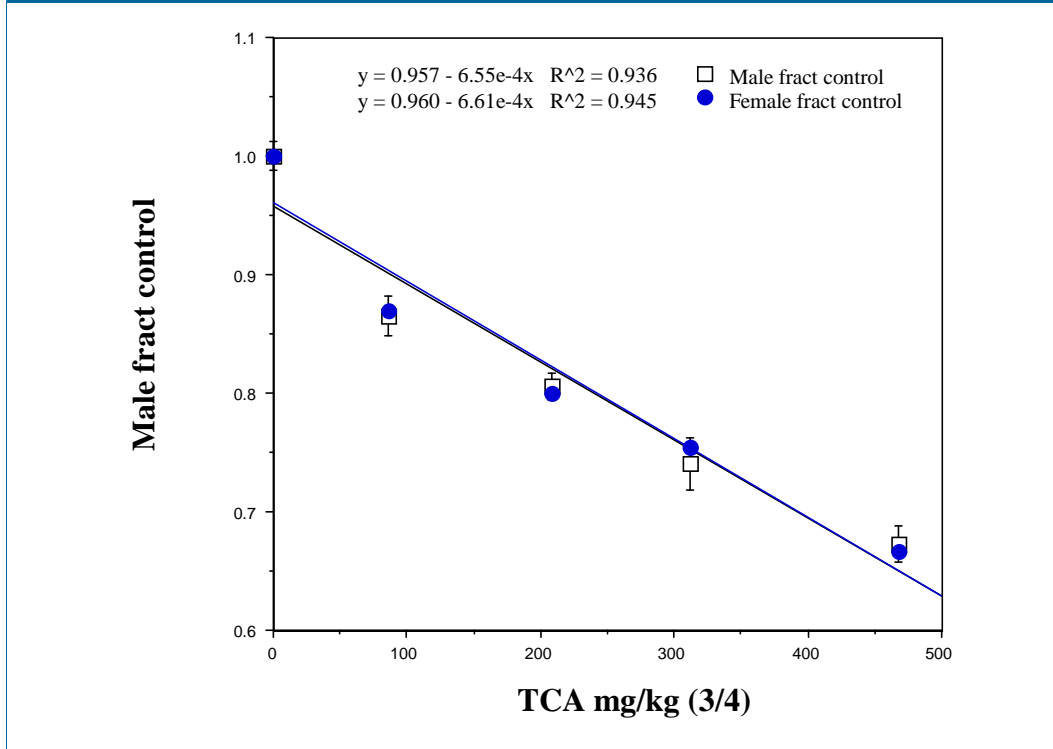


The impression that this category of effect often appears to be linear or nearly linear over a wide range of tested doses was strengthened by a series of observations of fetal weight responses to a variety of disinfection byproducts, prepared from data assembled by Syracuse Research for an EPA-sponsored workshop in 1999 (Figure A-3 through Figure A-8). To reduce some of the noise in individual dose-response data sets, Figure A-9 shows the results of combining these data (with inverse variance weighting) in terms of a dosimeter labeled “Trichloroacetic Acid (TCA) equivalents” constructed from the slopes of the fetal weight dose-response relationships for the individual agents. Overall, the data do not appear to be incompatible with low-dose linearity, although there is some indication of smaller effects than predicted by the straight line for the lowest dose points, and also some high-dose saturation of the response. The data are also compatible with simple empirical models that incorporate quadratic and cubic terms (Figure A-10 and Figure A-11).

The simplest overall interpretation of this type of finding for diverse toxicants is that the toxicants impose a kind of nonspecific “tax” on the resources that the organism would otherwise have available to grow and develop. This implies that for the case of the developing fetus, there simply may be little or no functional reserve capacity. The fetus, particularly in late pregnancy, may be mobilizing energy and other resources as fast as it can, such that even a marginal drain on these resources to counteract perturbations caused by a toxicant marginally reduces the growth and development of the organism.

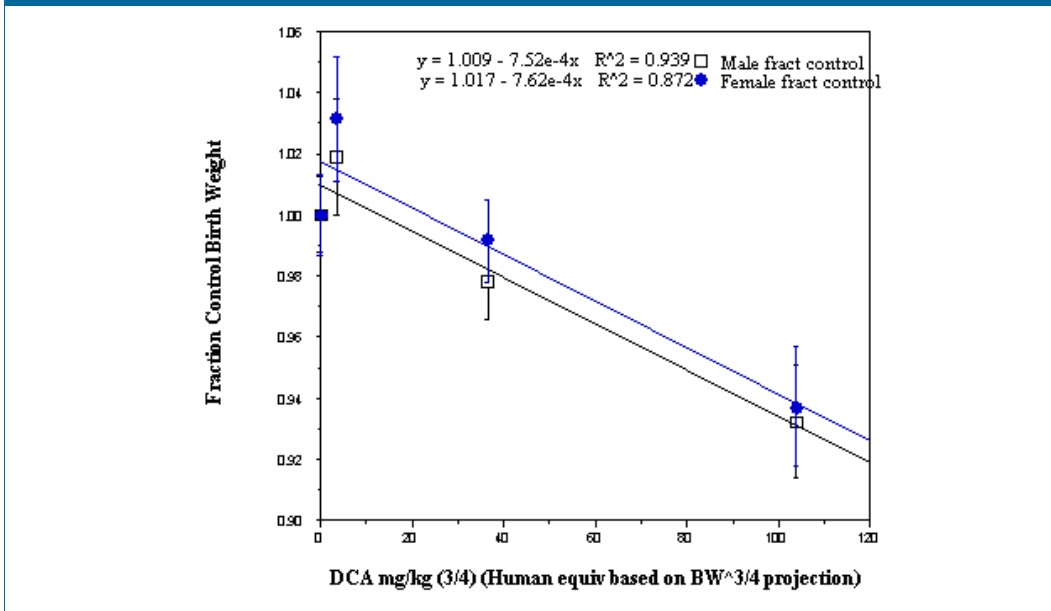
Such observations of fetal growth inhibition in animal systems could be of considerable public health significance if they turn out to be predictive of birth weight reductions in humans. Birth weights are strongly related to infant mortality, and the association between lower birth weights and higher infant mortality is not at all confined to the traditional category of “low birth weight” babies weighing under 2,500 g (Figure A-12). Certainly, this does not reflect a direct causal connection between lighter weight at birth and mortality, but it does seem likely that birth weight is a proxy for the energy resources that were available to the developing fetus to grow and develop the protective mechanisms that could be causally related to impairment of mechanisms that help reduce mortality risks in the first year of life.

Figure A-3: Fetal Weight Response of Rats to Trichloroacetic Acid



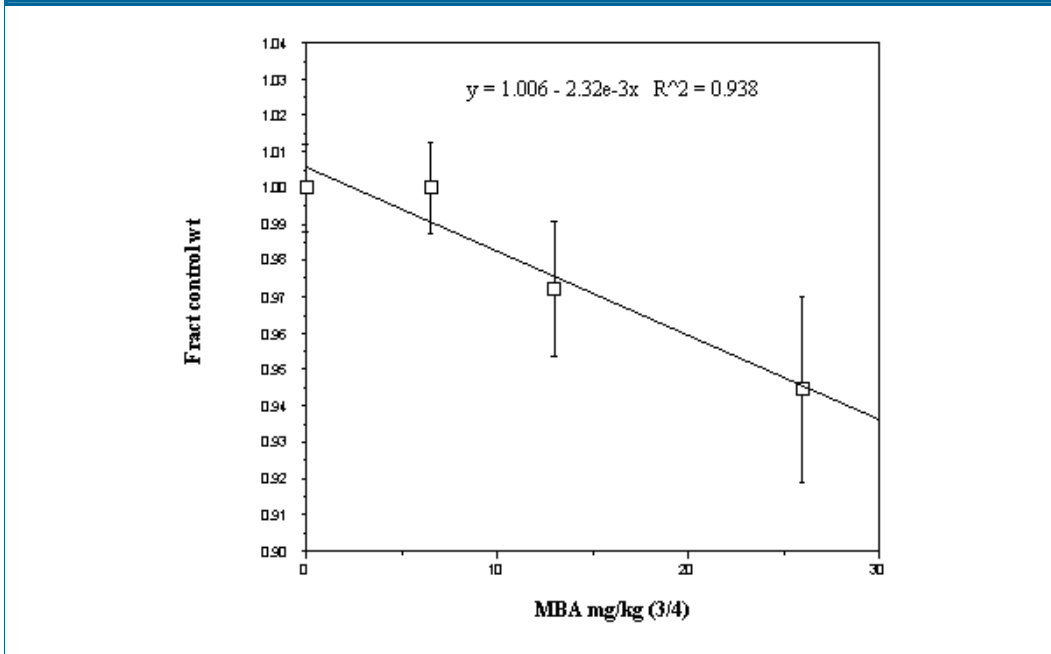
Source: Data of Smith et al. (1989a)

Figure A-4: Fetal Weight Response of Rats to Dichloroacetic Acid



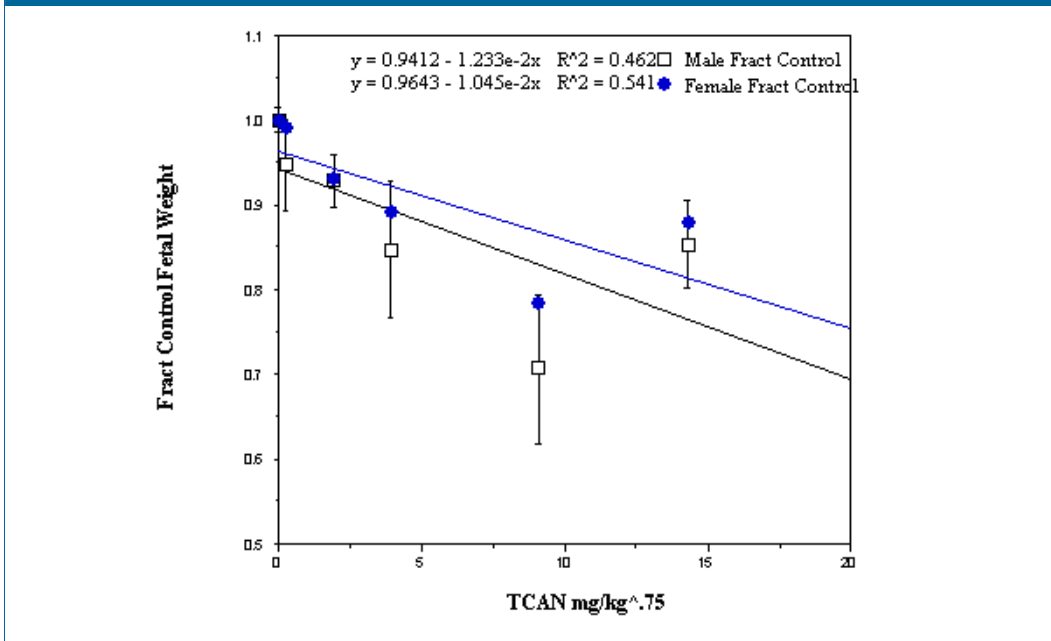
Source: Data of Smith et al. (1992)

Figure A-5: Fetal Weight Response of Rats to Monobromoacetic Acid



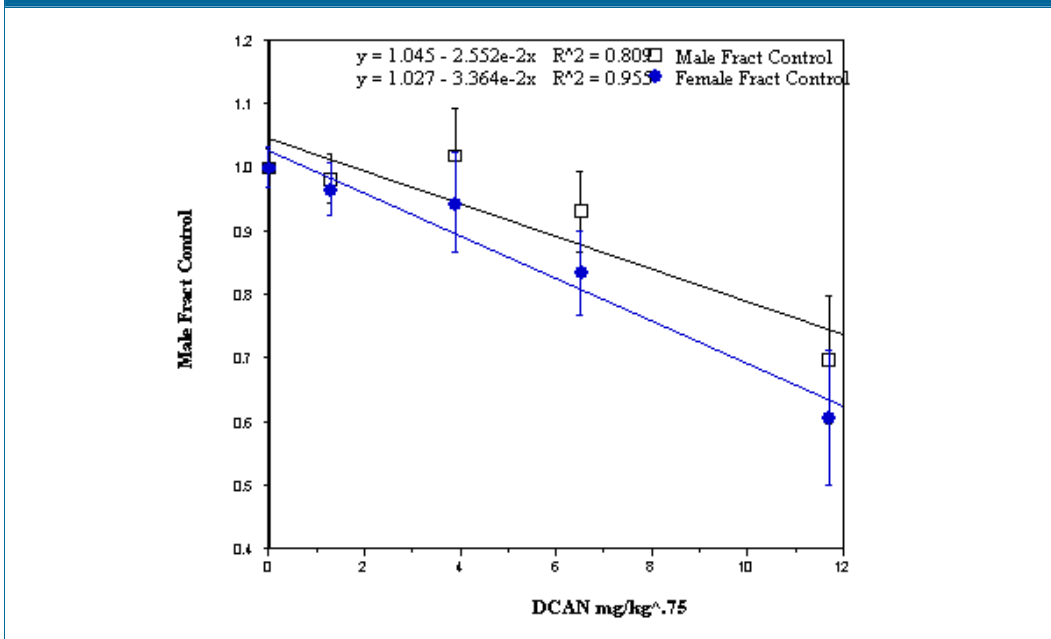
Source: Data of Randall et al. (1991)

Figure A-6: Fetal Weight Response of Rats to Trichloroacetonitrile



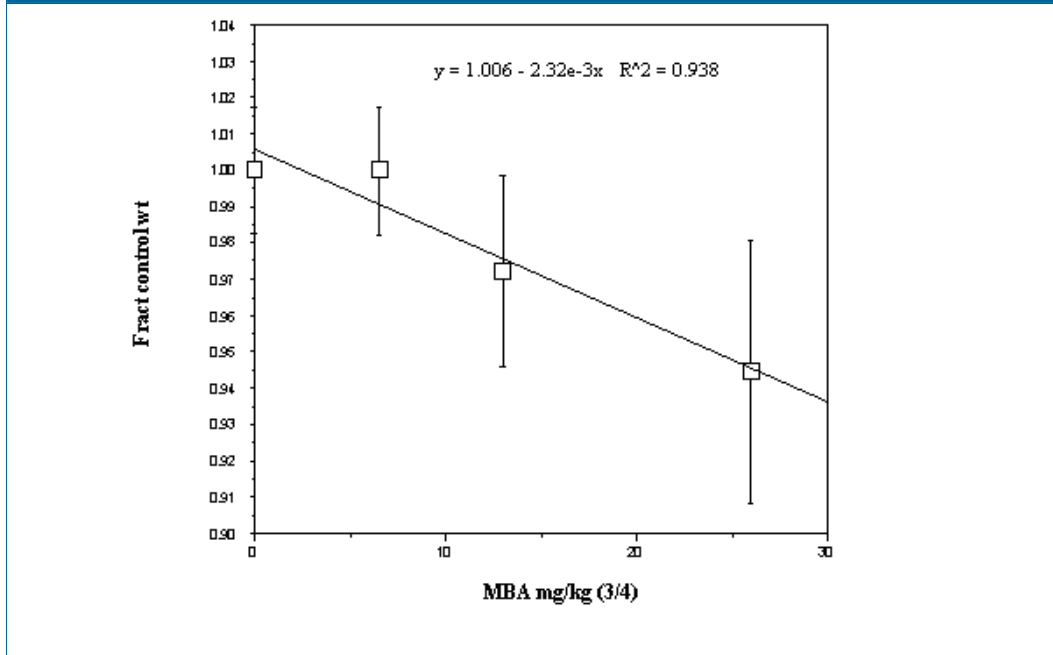
Source: Data of Smith et al. (1998)

Figure A-7: Fetal Weight Response of Rats to Dichloroacetylnitrile



Source: Data of Smith et al. (1989b)

Figure A-8: Fetal Weight Response of Rats to Monobromoacetic Acid



Source: Data of Randall et al. (1991)

Figure A-9: Results of Regression Analysis of the Fraction of Control Fetal Weight Response in Grouped Categories of TCA Equivalents—Interpretation with a Linear Model

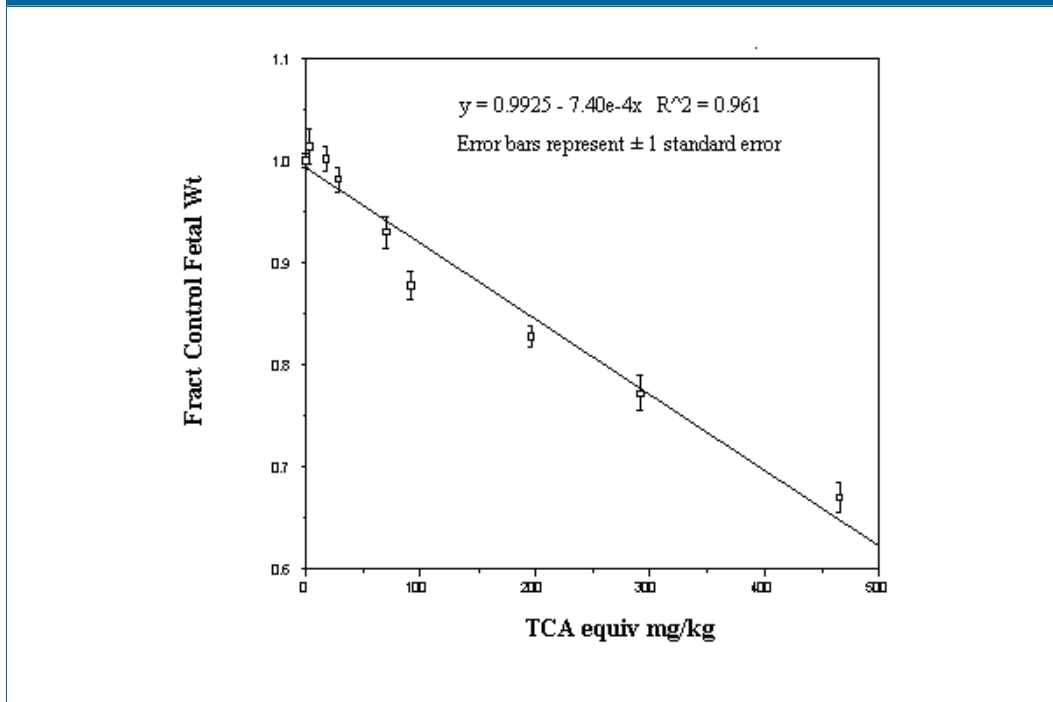


Figure A-10: Results of Regression Analysis of the Fraction of Control Fetal Weight Response in Grouped Categories of TCA Equivalents—Interpretation with a Quadratic Model

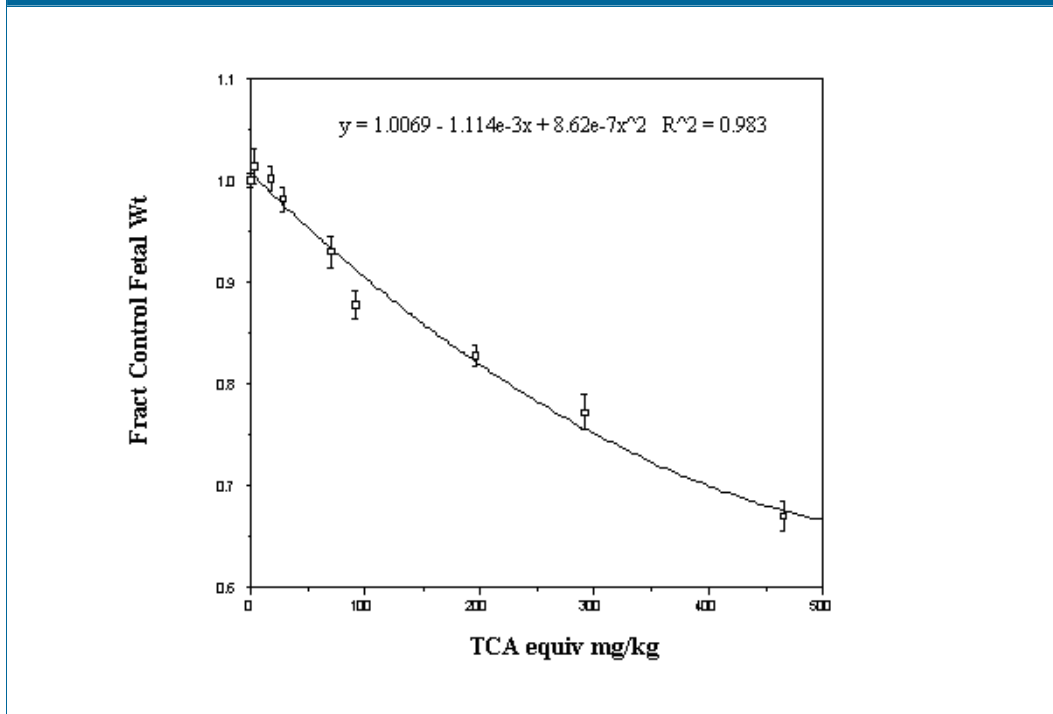


Figure A-11: Results of Regression Analysis of the Fraction of Control Fetal Weight Response in Grouped Categories of TCA Equivalents—Interpretation with a Third-Degree Polynomial

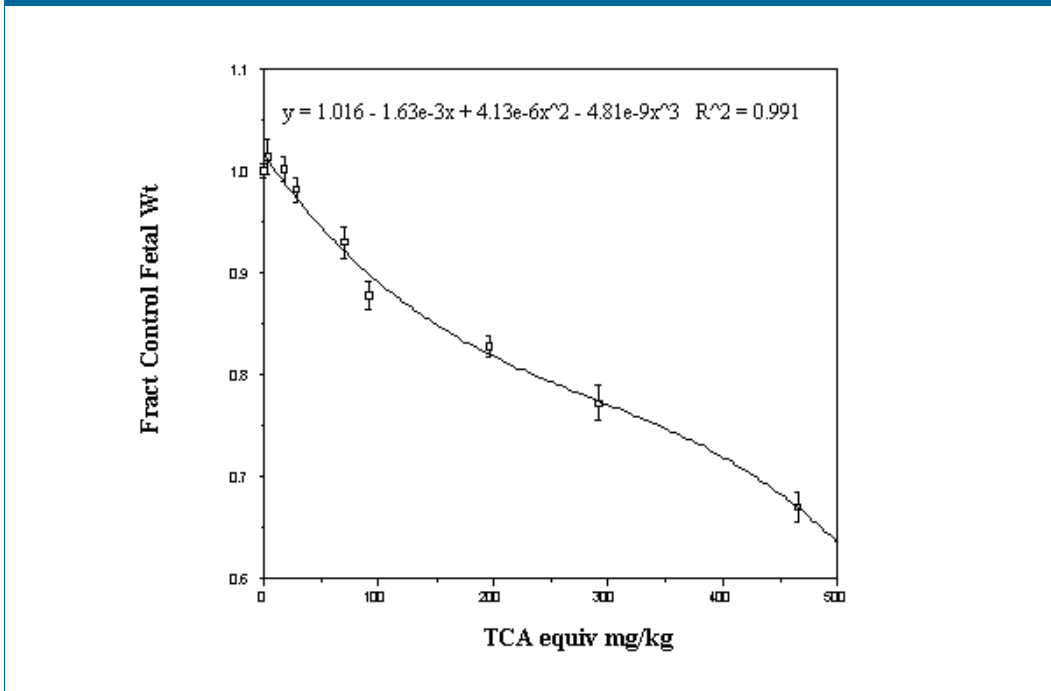
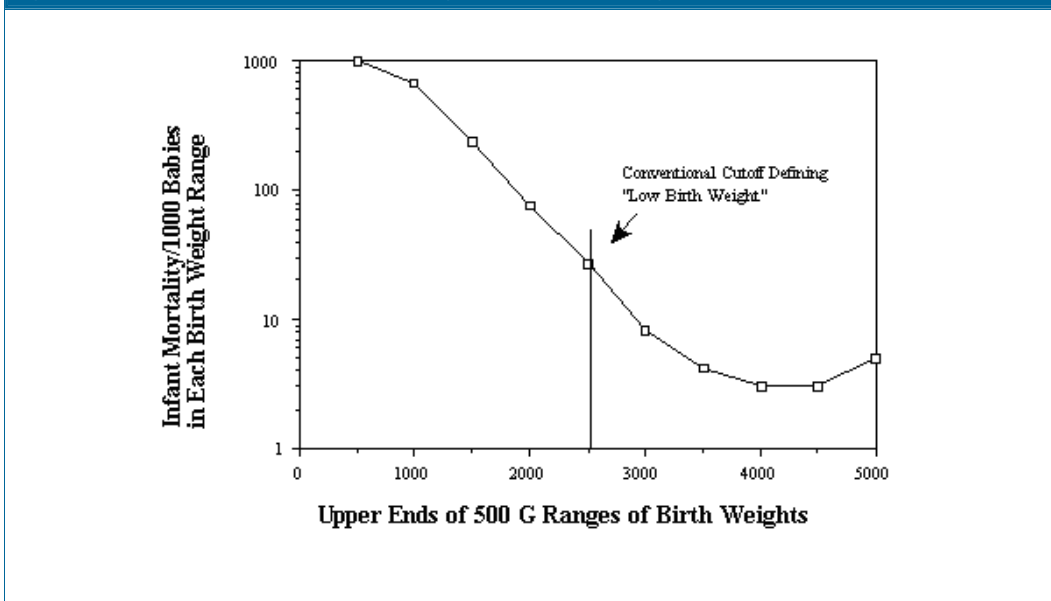


Figure A-12: Relationship Between Weight at Birth (in 500 g Increments) and Infant Mortality



Human data are available (National Center for Health Statistics 1996) for the associations of both birth weight reduction and infant mortality for reported exposures during pregnancy for one toxicant—cigarette smoking. It can be seen in Figure 4-2 (p. 49 in the main text) that the dose-response relationships for these effects bear a strong resemblance to each other. This suggests that inferences about incremental effects on infant mortality may not be unreasonable based on the predicted or observed effects of some toxicants on fetal growth and birth weights.

The Barker (2001) Observations and a Hypothesis—Latent Effects Appearing in Later Life From Early-Life Reductions in the Reserve Capacity To Accomplish Homeostatic Control

A repeated finding in recent years is that relatively low birth weights in humans are strongly associated with elevated risks of Type 2 diabetes 3-7 decades after birth (Forsen et al. 2000). The data in Figure A-13 indicate that the relationship is continuous over a wide range of birth weights, and does not appear to be confined to the traditional “low birth weight” babies (under 2,500 g).

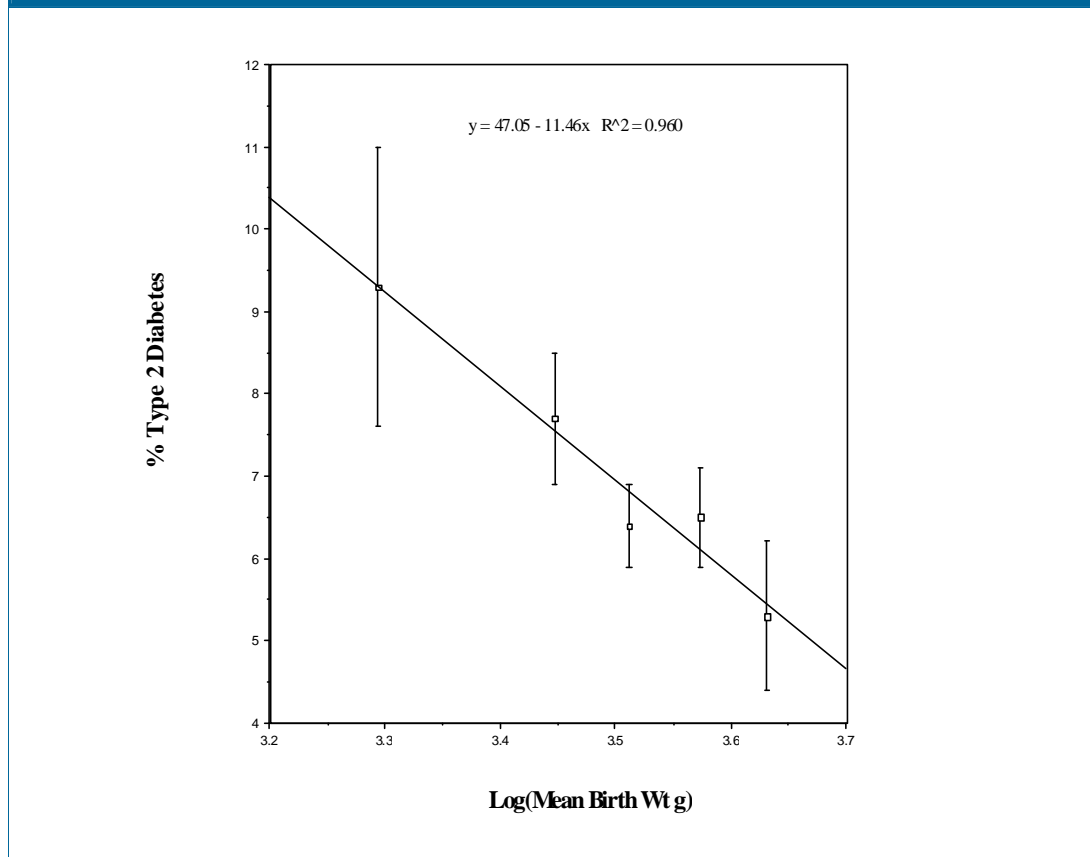
Hill and Duvillie (2000) review confirmatory mechanistic observations in rats—including findings that a 50 percent restriction of maternal calorie intakes from gestation day 15 through birth leads to a reduction in the mass of pancreatic beta cells—the cells that produce insulin and thereby control blood sugar levels. If dietary restriction is continued through weaning, the changes became irreversible in later life. Animals with continuing restriction through weaning are susceptible to glucose intolerance in later life. Thus, it appears that the size of the reserve capacity (pancreatic islets and beta cell numbers) built by the system for blood sugar control depends on what the system sees during developmental phases both before and after birth. A stressor (nutrient restriction in this case) can reduce the amount of reserve capacity that the system builds to provide for homeostatic control. The long-term consequence of this may be the excess incidence of Type 2 diabetes observed in people after normal aging processes have an opportunity to deplete the initially low reserve capacity to critical levels.

Similar findings have also been reported for another cardiovascular risk factor—blood pressure in human adolescents studied in the Philippines (Adair et al. 2001), even after correction for maternal nutrition variables.

In a guinea pig system, a period of maternal nutrient restriction as short as 48 hours in late pregnancy gives rise to detectable changes in pituitary-adrenal function in adult offspring (Lingas and Matthews 2000).

The presence of mutually supportive findings in this area in both human and experimental animal systems indicates that there is good potential to use the experimental animal models to work out detailed dose-time response relationships and mechanistic pathways for the likely actions of environmental toxicants in these systems—particularly toxicants such as the disinfection byproducts, glycol ethers, and valproic acid, where there are established effects on fetal growth.

Figure A-13: Plot of the Incidence of Type 2 Diabetes in Relation to log(Mean Birth Weight)



Source: Data of Forsen et al. (2000)

References - Appendix A

- Adair LS, Kuzawa CW, Borja J. 2001. Maternal energy stores and diet composition during pregnancy program adolescent blood pressure. *Circulation* 104: 1034-1039.
- Ballew M, Hattis D. 1989. Reproductive Effects of Glycol Ethers in Females--A Quantitative Analysis, M.I.T. Center for Technology, Policy, and Industrial Development, CTPID 89-7, Cambridge, Massachusetts, July.
- Barker JP, ed. 2001. *Fetal Origins of Cardiovascular and Lung Disease*, Marcel Dekker, Inc., New York.
- Forsen T, Eriksson J, Tuomilehto J, Runanen A, Osmond C, Barker D. 2000. The fetal and childhood growth of persons who develop type 2 diabetes. *Ann. Intern. Med.* 133: 176-182.
- Hanley TR Jr, Yano BL, Nitschke KD, John JA. 1984. Comparison of the teratogenic potential of inhaled ethylene glycol monomethyl ether in rats, mice, and rabbits. *Toxicol Appl Pharmacol.* 75(3):409-422.

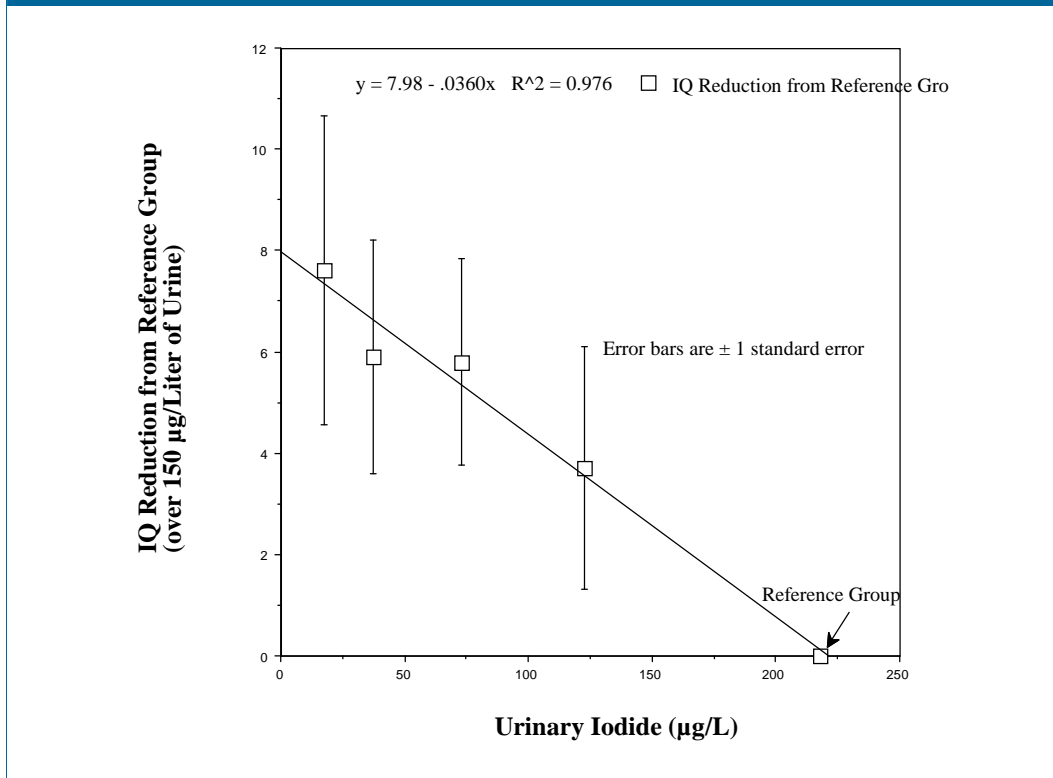
- Hill DJ, Duvillie B. 2000. Pancreatic development and adult diabetes. *Pediatric Research* 48: 269-274.
- Lingas RI, Matthews SG. 2001. A short period of maternal nutrient restriction in late gestation modifies pituitary-adrenal function in adult guinea pig offspring. *Neuroendocrinology*. 73(5):302-11.
- Randall JL, Christ SA, Hortonperez P, Nolen GA, Read EJ, Smith MK. 1991. Developmental effects of 2-Bromoacetic acid in the Long Evans rat *Teratology* 43: 454.
- Smith MK, Randall JL, Read EJ, Stober JA. 1989a. Teratogenic activity of trichloroacetic acid in the rat. *Teratology*. 40(5):445-451.
- Smith MK, Randall JL, Stober JA, Read EJ. 1989b. Developmental toxicity of dichloroacetonitrile: a by-product of drinking water disinfection. *Fundam Appl Toxicol*. 12(4):765-772.
- Smith MK, Randall JL, Read EJ, Stober JA. 1992. Developmental toxicity of dichloroacetate in the rat. *Teratology*. 46(3):217-223.
- Smith MK, Randall JL, Tocco DR, York RG, Stober JA, Read EJ. 1988. Teratogenic effects of trichloroacetonitrile in the Long-Evans rat. *Teratology*. 38(2):113-120.
- Wier PJ, Lewis SC, Traul KA. 1987. A comparison of developmental toxicity evident at term to postnatal growth and survival using ethylene glycol monoethyl ether, ethylene glycol monobutyl ether and ethanol. *Teratog Carcinog Mutagen*. 7(1):55-64

Appendix B – Case Study: Competitive Inhibition of Iodide Uptake by Perchlorate

Regulation of perchlorate in drinking water has been a subject of much recent controversy. An analysis by the National Research Council (NRC 2005) recommended an increase of about 20 fold in an earlier EPA recommended RfD for this substance, based on a group-average analysis of human data on perchlorate transport. Some of the reason for the controversy involves the likely interactions between perchlorate exposure and iodide deficiency (Ginsberg and Rice 2005). Iodide transport into the thyroid is a vital part of the causal chain of events that leads to synthesis of thyroid hormones, and perchlorate is a competitive inhibitor of the “symporter” protein that is responsible for active transport of the iodide ion. Particularly during the perinatal period, iodide deficiency is associated with risks of impaired neurological development (Delange 2001). Measurements associate even relatively mild iodide deficiency (as imperfectly indicated by the excretion of relatively low levels of iodide in spot urine samples) with detectable and apparently continuous decrements in IQ in school-age children (Figure B-1—based on our analysis of data from Santiago-Fernandez et al. 2004). This result suggests, contrary to arguments advanced by NRC (2005), that mild iodide deficiency present in developed countries is associated with detectable effects that are not fully offset by normal compensatory mechanisms (including, for example, increased levels of symporter molecules. (The NRC, in discussing the results of this study, complains that the authors did not adjust for parental socioeconomic or educational status.)

One objection to the NRC (2005) analysis raised by Ginsberg and Rice (2005) is over the analysis of the data from the key study (Greer et al. 2002) of low-moderate perchlorate exposures on a group, rather than an individual basis. (This is not the same as the “ecological fallacy” familiar in epidemiology, but it does represent a potential failure to use the data as fully as possible to most sensitively detect effects that may be most clearly manifest in a minority of people with different baseline iodide status.) Analysis on a group basis led the NRC to designate the lowest 0.007 mg/kg dose in the Greer et al. study as a NOEL on the basis that the 1.8 percent group average inhibition of the 7 individuals in that low-dose group was not statistically significantly different from the 0-dose group. Ginsberg and Rice (2005) note from the individual data that some individuals in this group with relatively higher observed baseline uptakes (possibly associated with lower internal iodide concentrations, although this was not measured) appeared to show a much larger inhibition when exposed to the low dose of perchlorate. A second objection is that the mathematical formula used for the original analysis by Greer et al. is a simple empirical log-linear model with no apparent mechanistic justification. We believe a better analysis is possible using a mechanistically-grounded equation using a Michaelis-Menten type formula to represent changes of the residual absolute amount of iodide transport as a function of perchlorate exposure. The fact that neither the original authors nor the NRC, nor, as far as we can tell, regulatory agencies have chosen this path for analysis of the individual data is testimony to the continuing dominance of the NOAEL/uncertainty factor paradigm based on empirical analyses of group average data for continuous variables, rather than a new paradigm of mechanistically driven assessments of the distribution of expected effects in a diverse human population.

Figure B-1: Inferences of the Population Change in Median IQ in Relation to Spot Urinary Iodide Measurements at Age 9



Note: Mean urinary iodide concentrations and median IQ in each group were reconstructed from percentile data provided in Santiago-Fernandez et al. (2004).

References - Appendix B

- Delange F. 2001. Iodine deficiency as a cause of brain damage. *Postgraduate Medical Journal* 77:217-220.
- Ginsberg G, Rice D. 2005. The NAS perchlorate review: questions remain about the perchlorate RfD. *Environ Health Perspectives* 113(9):1117-1119.
- Greer MA, Goodman G, Pleus RC, Greer SE. 2002. Health effects assessment for environmental perchlorate contamination: the dose response for inhibition of thyroidal radioiodine uptake in humans. *Environ Health Perspect.* 110(9):927-937.
- National Research Council (NRC). 2005. *Health Implications of Perchlorate Ingestion*. National Academies Press, Washington, DC.
- Santiago-Fernandez P, Torres-Barahona R, Muela-Martinez JA, Rojo-Martinez G, Garcia Fuentes E, Garriga MJ, Leon AG, Soriguer F 2004. Intelligence quotient and iodine intake: A cross-sectional study in children J. *Clinical Endocrinology & Metabolism* 89:3851-3857.