LBNL-47977

Developing and Evaluating Distributions for Probabilistic Human Exposure Assessments

Randy L. Maddalena
Indoor Environment Department
Environmental Energy Technologies Division

Thomas E. McKone
Indoor Environment Department
Environmental Energy Technologies Division
and
School of Public Health
University of California, Berkeley

Funded by
U.S. Environmental Protection Agency
Office of Emergency and Remedial Response (OERR)

University of California
Ernest Orlando Lawrence
Berkeley National Laboratory
Berkeley, California

August, 2002

<Blank Page>

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory
is an equal opportunity employer.

Acknowledgments

For help with data acquisition and data evaluation, the authors gratefully acknowledge

Agnes Bodnar	Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory and The University of California, Berkeley
Janet Jacobson	Earth Science Division, Lawrence Berkeley National Laboratory
Michael D. Sohn	Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory

For review and helpful comments on this report, we thank

Elizabeth H. Margosches	Office of Pollution Prevention and Toxics, USEPA

For reviews and comments on an earlier draft of this report, we acknowledge

Tim Barry	Office of Policy
Robert Fares	Environmental Standards, Inc.
Henry Kahn	Office of Water
Elizabeth Margosches	Office of Toxic Substances
Tom McCurdy	Office of Research and Development/NERL
Jackie Moya	Office of Research and Development/NCEA
Paul Price	Ogden Environmental and Energy Services
Curtis Travis	Project Performance Corporation
Shree Whitaker	National Institute of Environmental Health Sciences

For helpful discussions and comments on various sections in this report, we thank

Steve Chang	Office of Emergency and Remedial Response, USEPA
Phil Price	Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory

Abstract

This report describes research carried out at the Lawrence Berkeley National Laboratory (LBNL) to assist the U. S. Environmental Protection Agency (EPA) in developing a consistent yet flexible approach for evaluating the inputs to probabilistic risk assessments. The U.S. EPA Office of Emergency and Remedial Response (OERR) recently released Volume 3 Part A of Risk Assessment Guidance for Superfund (RAGS), as an update to the existing two-volume set of RAGS. The update provides policy and technical guidance on performing probabilistic risk assessment (PRA). Consequently, EPA risk managers and decision-makers need to review and evaluate the adequacy of PRAs for supporting regulatory decisions. A critical part of evaluating a PRA is the problem of evaluating or judging the adequacy of input distributions PRA. Although the overarching theme of this report is the need to improve the ease and consistency of the regulatory review process, the specific objectives are presented in two parts. The objective of Part 1 is to develop a consistent yet flexible process for evaluating distributions in a PRA by identifying the critical attributes of an exposure factor distribution and discussing how these attributes relate to the task-specific adequacy of the input. This objective is carried out with emphasis on the perspective of a risk manager or decision-maker. The proposed evaluation procedure provides consistency to the review process without a loss of flexibility. As a result, the approach described in Part 1 provides an opportunity to apply a single review framework for all EPA regions and yet provide the regional risk manager with the flexibility to deal with site- and case-specific issues in the PRA process. However, as the number of inputs to a PRA increases, so does the complexity of the process for calculating, communicating and managing risk. As a result, there is increasing effort required of both the risk professionals performing the analysis and the risk manager reviewing it. For deterministic risk assessments, the use of default inputs has improved the ease and the consistency of both performing and reviewing assessments. By analogy, it is expected that similar advantage will be seen in the field of probabilistic risk assessment through the introduction of default distributions. In Part 2 of this report, we consider when a default distribution might be appropriate for use in PRA and work towards development of recommended task-specific distributions for several frequently used exposure factors. An approach that we develop using body weight and exposure duration as case studies offers a transparent way for developing task-specific exposure factor distributions. A third case study using water intake highlights the need for further study aimed at improving the relevance of "short-term" data before recommendations on task-specific distributions of water intake can be made.

Contents

Acknowledgments ii
Abstract iii
Contents iv
Tables vi
Figures vii
Acronyms and Symbols viii
Executive Summary 1
Part 1: Critical Attributes for Evaluating Distributions 1
Part 2: On the Development of Recommended Exposure Factor Distributions 4
1.0 Introduction 8
1.1 Probability Distributions in Exposure Assessment 8
1.2 The Need for an Evaluation Process for Existing Distributions 9
1.3 The Need for a Set of Recommended Exposure Factor Distributions 10
1.4 Objectives of this Study 11
1.5 Overview of the Report 11
2.0 Background 12
2.1 Probabilistic Inputs and Models 12
2.2 Variability and Uncertainty 13
2.3 The Link between Data and PDFs 15
2.4 A Tiered Approach to Uncertainty/Variability Analysis 17
2.5 Justification for Evaluation Process and Recommended Distributions 17
Part 1: Critical Attributes forEvaluating Distributions 19
3.0 Category One: The Sensitivity Analysis 21
3.1 The Myriad of Sensitivity Analysis Techniques 22
3.1.1 Local Sensitivity Analysis Techniques 22
3.1.2 Exact Analytical Methods 23
3.1.3 Global Sensitivity Analysis Techniques 25
3.2 Interpreting the Sensitivity Analysis Results 30
3.3 Relating Sensitivity Analysis Results to the other Attributes 31
4.0 Category Two: Adequacy of the Data 32
4.1 Data Quantity Attribute 32
4.1.1 Relating Sample Size to Precision in Distributions 33
4.1.2 Interpreting Sample Size 36
4.2 Data Quality Attribute 37

4.3 Data Relevance Attribute 38
5.0 Category Three: Adequacy of the Distribution 39
5.1 Statistical Goodness-of-Fit Attribute 40
5.2 Graphical Goodness-of-Fit Attribute 40
5.3 Distribution Performance Attribute 41
6.0 A Process for Evaluating Inputs in Probabilistic Risk Assessment 42
Part 2: On the Development of Recommended Distributions 47
7.0 Body Weight 50
7.1 Source of data 51
7.1.1 CDC Revised Growth Charts for Children from Birth to 20 years 51
7.1.2 Continuing Survey of Food Intakes by Individuals, CSFII 52
7.1.3 National Health and Nutrition Examination Survey, NHANES III 53
7.1.3 The Combined NHANES III, CSFII data set 53
7.2 Key demographic factors 55
7.3.1 U.S. Census Data Summary Tables 56
7.3 Constructing Age/Sex Baseline Distributions 57
7.3.1 Baseline distribution for children (birth through 19 years) 57
7.3.2 Baseline distributions for adults (≥ 20 years) 59
7.4 Relating the Demographically Based Distributions to Site-specific Distributions 61
8.0 Exposure Duration 62
8.1 Sources of data 63
8.2 Identification of key demographic factors 64
8.3 Overview of Existing Methods 67
8.4 Recommendation for Estimating Distributions of Exposure Duration 69
9.0 Water Intake 70
9.1 The Implications of using Short-term Data to Estimate Long-term Averages 71
9.2 A Final Recommendation for Water Intake 76
References 79
Appendix 1: Estimation of Body Weight Percentiles for Children 85
Appendix 2: Body Weight Percentile Data for Adults 91

Tables

Table 1: Relating sample size to precision in order statistics 34
Table 2: Best fit parameters for each age/sex category for children 58
Table 3: Best fit parameters for each age/sex category for adults.....61
Table A2.1: Empirical percentiles of body weight (kg) for adult women.....91
Table A2.2: Empirical Percentiles of body weight (kg) for adult men.....92
Table A2.3: The smoothed percentiles of body weight (kg) for adult women. 93
Table A2.3: The smoothed percentiles of body weight (kg) for adult men. 94

Figures

Figure 1: The family of cumulative probability plots reflecting both variability and uncertainty. Each curve expresses a realization of a variability distribution at a different level of uncertainty. 14

Figure 2a: Probability density function for a Normal distribution with mean of 1 and standard deviation of 0.3. 16

Figure 2b: Cumulative density function for a Normal distribution with mean of 1 and standard deviation of 0.3. 16

Figure 3: The 95% confidence interval for the ratio of the estimated sample variance to the true (unknown) variance illustrating the relationship between sample size and precision .35

Figure 4: Comparison of empirical cumulative distribution functions from two national surveys where the measured values from NHANES III are shown as smooth curves and the self-reported values from CSFII are the irregular or jagged curves. Figure 4 also illustrates the difference in body weight related to *sex*. 54

Figure 5: An illustration of difference between the empirical cumulative distribution of reported body weight from CSFII for females identified with different racial groups. 56

Figure 6: Illustration of the 5th (lower dash), 50th (solid line) and 95th (upper dash) percentile of body weight for adult males as a function of age from birth to 85 years and above. 60

Figure 7: The relationship between the average reported current residence time for the national population and the age of the sample person showing a bimodal distribution that arises from correlation between residence time of children and their parents. 65

Figure 8: Annual geographic mobility data showing percent of population that reported moving during the year that proceeded each Annual Demographic Supplement to the Current Population Survey. 66

Figure 9: Comparison of the estimates for occupancy period (total time that an individual lives in current residence) from different methods. 68

Figure 10: Results for Case 1. Comparison of the "true" distribution of average values (solid line) with an empirical distribution constructed using 250 hypothetical individuals where two samples are drawn for each individual and used to estimate the average intake for that person. The variance structure for individuals in the population is assumed to be constant from day-to-day for all members of the population; $\sigma = 0.9$. 73

Figure 11: Results for Case 2. Comparison of the "true" distribution of average values (solid line) with an empirical distribution constructed using short-term data (2-day) where the day-to-day variance structure in the population is assumed to be distributed randomly with lognormal distribution; LogNorm(1.35,1.00). 74

Figure 12: Results from Case 3. Comparison of the "true" distribution of average values with an empirical distribution constructed using short-term data (2-day) where the day-to-day variance structure in the population is assumed to be represented by a constant relative standard deviation, i.e., day-to-day variation for each member of the population has the same coefficient of variation; CV(0.5). 75

Acronyms and Symbols

AT	Averaging Time for exposure
BW	Body weight
CDC	Center for Disease Control
CDF	Cumulative distribution function
CRT	Current residence time
CSFII	Continuing Survey of Food Intakes by Individuals
$\%CTV_i$	Percent contribution to variance for input *i*
CV	Coefficient of variation
DG	Demographic group
ED	Exposure Duration (assumed duration of time in current residence for this report)
EDF	Empirical distribution function
EF	Exposure frequency
EPA	Environmental Protection Agency
GM	Geometric mean
GOF	Goodness of fit
GSD	Geometric standard deviation
HUD	U.S. Department of Housing and Urban Development
IUw	Intake/uptake rate for water
LBNL	Lawrence Berkeley National Laboratory
LSRO	Life Science Research Office
NCHS	National Center for Health Statistics
NHANES III	Third National Health and Nutrition Examination Survey
OERR	U.S. EPA Office of Emergency and Remedial Response
PCC	Partial correlation coefficient
PDF	Probability distribution function
PRA	Probabilistic risk assessment
RAGS	Risk Assessment Guidance for Superfund
RSA	Regional sensitivity analysis
SA	Sensitivity analysis
SRC	standardized regressions coefficient
SS_{xy}	sum of the squared difference between x and y
σ	Arithmetic standard deviation
σ^2	Variance
USDA	U.S. Department of Agriculture
USDOC	U.S. Department of Commerce Economics and Statistics Administration
$\bar{x}$	Arithmetic mean

Executive Summary

This report describes research carried out at the Lawrence Berkeley National Laboratory (LBNL) to assist the U. S. Environmental Protection Agency (EPA) in developing a consistent yet flexible approach for evaluating the inputs to probabilistic risk assessments. As part of the EPA Superfund reform activities, the U.S. EPA Office of Emergency and Remedial Response (OERR) is updating its 1989 Risk Assessment Guidance for Superfund (RAGS). Volume 3 Part A of RAGS was recently released in draft form as an update to the existing two-volume set of RAGS. The update provides policy and technical guidance on conducting probabilistic risk assessment for both human and ecological receptors. Risk assessors now have the option of using probabilistic risk assessment (PRA) to provide supporting information for risk management decisions at Superfund hazardous waste sites. Consequently, EPA risk managers and decision-makers need to review and evaluate the adequacy of PRAs for supporting regulatory decisions. One of the key components in the evaluation of a PRA, and the main focus of this report, is the problem of evaluating or judging the adequacy of input distributions used in PRA.

This report has two objectives. The first objective is to identify critical attributes of input distributions. This objective is carried out from the perspective of a risk manager or decision maker and provides general recommendations for relating critical attributes to the distribution's task-specific adequacy, that is to how it will be used in decisions. The second objective is to identify when default distributions are appropriate for use in PRA. This is done by example through development of recommended distributions for several frequently used exposure factors. Both objectives have in common the goal of improving the ease and consistency of the regulatory review process for PRAs. But the objectives are sufficiently different to justify separating the report into two parts—one dealing with "Critical Attributes for Evaluating Distributions" and the other "Toward Development of Recommended Exposure Factor Distributions". Summaries for each part are provided below.

Part 1: Critical Attributes for Evaluating Distributions

This part of the report draws on numerous studies, reports, workshops and our experience to identify and summarize a set of critical attributes for probabilistic input distributions. Even though there are many different ways to develop probability distributions, we find that the information needed to judge its task-specific adequacy can be summarized with seven key attributes. These attributes are grouped into three categories—(1) the sensitivity of the outcome of a PRA to the input distribution, (2) data evaluation, and (3) the distribution or statistical model evaluation category. Our approach, justification, and application of each category are summarized in the following paragraphs.

Sensitivity Analysis Category

The sensitivity-analysis category considers the importance of the distribution in the context of the overall probabilistic risk analysis, that is, the sensitivity of the outcome of a PRA to the input distribution. There are an increasing number of methods available for investigating the relationship between variation in the inputs of a model and variation in the model outcome (Campolongo *et al*., 2000). These sensitivity analyses (SA) play a vital role in the set-up, use and evaluation of probabilistic risk models. Therefore, all PRAs should include results and discussion from some form of sensitivity analysis. These results will likely be reported as elasticity, sensitivity score, contribution to variance, correlation coefficients or regression coefficients. If SA results are lacking, the relative contribution to variance can be approximated for the simple product/quotient models that are typically used at Superfund hazardous waste sites. Distributions that contribute less than a few percent to the outcome variance of the model need only be checked to insure that the central tendency and the approximate range of the distribution are appropriate. The remaining distributions, those that contribute more than a few percent, should be given more scrutiny in the review process. In these cases, evaluation of the input would need to consider both the underlying data and the selected distribution.

Data Evaluation Category

Data evaluation considers as key attributes the quantity, quality and relevance of the information that is the foundation of any constructed distribution. Methods have been described that relate sample size to uncertainty (Cullen and Frey, 1999). The issue of relating data quality and relevance to uncertainty has also been discussed (Morgan and Henrion, 1990; USEPA, 1999a), although in more subjective terms.

Data Quantity Attribute – Typically the uncertainty in a distribution is inversely related to the sample size. The actual sample size needed to characterize a distribution depends on both the range of data and the precision in the measurement of each sample value. Uncertainty can often be ignored when sample size is large but small sample size will require explicit treatment of uncertainty in the analysis. For distributions based on moderately small samples the decision to explicitly include uncertainty due to the size of the dataset will need to be judged on a case-by-case basis.

Data Quality Attribute – Experimental error due to flaws in sampling design or imprecision in measurements and/or self-reported information are reflected in the quality of data and these errors propagate through to the distribution development process. The relative quality of information used in developing distributions for PRA is as follows: direct measurements of the exposure factor of interest > self reported values for the exposure factor of interest > direct measure of surrogate data > self-reported surrogate data. Surrogate information (e.g., estimated breathing rate from caloric intake or exposure duration based on population mobility) may often be the only information available.

Data Relevance Attribute – Even large amounts of high quality data cannot compensate when the data is not relevant to the population of interest. Relevant data adequately captures temporal and spatial variability as well as the demographic characteristics of the target population. Judging the relevance of a distribution requires a clearly defined population or cohort, a well-defined exposure scenario, a well-documented source of data and a clear description of the steps, if any, taken by the risk assessor to adjust for data quality or to improve data relevance (USEPA, 1999a).

Distribution Evaluation Category

The distribution evaluation considers how well the input distribution simulates the original information and the degree to which the performance of the distribution can be validated using alternate models, assumptions, theoretical underpinnings and/or independent data. The three attributes of the distribution-evaluation category include both statistical and graphical goodness-of-fit (GOF) and cross-validation performance or verification with competing models/theory.

Goodness-of-Fit Attributes– Many statistical GOF tests are available to the risk assessor for evaluating a hypothesis that data come from a particular distribution. Examples include the Chi-squared, Anderson-Darling, Kolmogorov-Smirnov and various measures of regression and correlation. Although positive results from a GOF test can contribute to the reviewer's belief that the distribution adequately summarizes the data, passing a statistical GOF test should not be taken as proof that a distribution is adequate, nor should failing the test be taken as a fatal flaw in the distribution. Graphical methods of performance testing are often preferred in cases where curve fitting is important in a particular region of the distribution. Informative graphical methods for visualizing the performance of a distribution include residual plots, percentage-percentage probability (P-P) plots or quantile-quantile (Q-Q) plots (USEPA, 2000b). A reviewer should be able to use these visual tests to evaluate the performance of a distribution relative to a specific target population and exposure scenario.

Performance Evaluation Attribute – Types of information that might be considered in the performance evaluation include the theoretical basis for a selected distribution, verification using alternate models or assumptions, validation using independent data, cross-validation using samples drawn from the original data and different sources of expert or professional advise. The performance evaluation attribute provides supporting information for decisions made by the risk assessor during the distribution development process.

Synthesizing the Attributes: The Evaluation Wheel

Evaluating the attributes described above is often, if not always, subjective on the part of both the risk assessor and the risk manager. Thus, we introduce a qualitative paper-based tool for guiding the risk manager through the evaluation process and for tracking the critical

attributes of a distribution. The evaluation wheel is introduced and used in this report as a balance sheet to track each attribute in a simple yet informative manner. Each attribute of a given input corresponds to a different ray or spoke on the wheel. After using sensitivity analysis results to prioritize the distributions, the reviewer systematically evaluates each of the key attributes and uses that information to either justify accepting the input or to facilitate discussion with the risk assessor about where improvements to the input might be needed. By following the recommendations presented in this report, and working closely with the risk assessor and other risk professionals, risk managers can arrive at consistent and transparent conclusions regarding the adequacy of the probabilistic input distributions used in a PRA.

Part 2: On the Development of Recommended Exposure Factor Distributions

For deterministic risk assessments, default inputs have improved the ease and the consistency of both performing and reviewing assessments. The challenge for developing "default" exposure factor distributions is the need to reduce the number of distributions that are considered for a given input while maintaining enough flexibility to adequately represent specific target populations and/or risk scenarios in the PRA.

We consider two initial options for developing recommended distributions for PRA. The first option begins with a baseline or "prior" distribution for the general U.S. population then updates the distribution to better represent the target population using Bayesian statistical methods and a minimum amount of sampling within the target population. The second option begins with a set of default distributions (building blocks) that are relevant to each key demographic category. These demographically based distributions are then used to construct a case-specific distribution, with uncertainty and variability, based on readily available information about the size and demographic composition of the target population. Although certain aspects of the Bayesian approach may be necessary, the case studies described in this report focus primarily on the second option because of the need for transparency in the distribution development process and the desire to work with existing and readily available information.

As a first step toward recommended distributions for PRA, we focus on three common exposure factors: body weight, exposure duration and water intake. Each of these exposure factors has available a significant amount of nationally representative data.

Body Weight

Although body weight can be important for relating exposure to risk, it is often one of the least important contributors to variance in probabilistic risk assessments. A large amount of

useful data is readily available for a number of demographic categories within the U.S. population. The general approach described in this report for developing recommended body weight input distributions for PRA is to identify and acquire data, use available data analysis methods to identify important demographic factors and develop distributions of body weight for each demographic category. Finally, we suggest a transparent method for sampling from these distributions, using information from census data, to construct relevant case-specific distributions.

Source of data: Measured body weights from the most recent National Health and Nutrition Examination Survey, NHANES III, (NCHS, 1996) are combined with the continuing survey of food intakes by individuals, CSFII, (USDA, 1998) and the Centers for Disease Control (CDC) data on children's bodyweights. The combined data are used to construct demographically relevant distributions that are sampled from to construct input distributions for specific target populations.

Key demographic factors: Body weight is primarily dependent upon age, followed by sex and race with the latter two factors only important for adults. Although race was found to have a measurable influence on body weight, even the combined CSFII and NHANES III data set was not large enough to allow development of distributions for all age/sex/race categories. Therefore, the demographic categories are constrained to age and sex.

Demographically specific distributions: The approach taken here for development of age/sex specific distributions for body weight is similar to the methods described in Kuczmarski et al., (2000) and recent work characterizing distributions of body weight using data from NHANES II (Burmaster and Crouch, 1997). Smooth percentiles are developed from the raw data or the CDC growth charts then a three-parameter lognormal model is fit to the percentiles for each demographic category.

Development of case specific distribution: Information about the demographic composition of the target population can be accessed through the U.S. Census interactive data retrieval service or by direct survey of the target population. This information is used to design a case-specific sampling strategy for reconstructing the target population from the demographic distributions. A representative number of values are drawn from each age/sex distribution then combined in a way that is relevant to the PRA (e.g., children ages 7-13 or adult women). An appropriate parametric model, typically a two-parameter lognormal or extreme value model, is fit to the simulated data. For small populations or cohorts the sampling and fitting process is repeated several times to estimate uncertainty about the model parameters.

Exposure Duration

Exposure duration (defined here as total duration within a specified region) is typically an influential input to a PRA. But exposure duration is rarely measured. Rather, surrogate data

such as mobility, mortality, current residence time and/or tax records have been used to estimate occupancy period, which is used as a surrogate for exposure duration (Israeli and Nelson, 1992; Johnson and Capel, 1992; Price *et al*., 1992; Sedman *et al*., 1998). Given the inherent uncertainties about the demographic makeup of the target population, all of the previously published methods give comparable results . The general approach taken for exposure duration in this report is to identify and acquire appropriate surrogate data, evaluate and identify the key demographic factors in that data and select or develop an appropriate model for relating the surrogate data to exposure duration.

Source of data: The U.S. Census Bureau conducts regular national housing surveys. These surveys provide comprehensive housing statistics with a range of demographic factors. The data from 1995 (USDOC, 1995) were used in a previous report (Maddalena *et al*., 1999) to identify key demographic factors in the population. The Census Bureau (USDOC, 1999) also reports population mobility on a regular basis. Population mobility has remained relatively constant over the last few decades and as such provides a useful surrogate for estimating exposure duration for selected cohorts or populations.

Identification of key demographic factors: Current residence time (CRT) has been used previously as a surrogate for exposure duration and to identify important factors relating to occupancy period (Maddalena *et al*., 1999). Results indicate that exposure duration is primarily dependent upon tenure (rent or own) and age. We found a significant difference between the probabilities of changing residences within the same county and that of moving to a different county.

Development of occupancy period model: Because of the difficulty of obtaining occupant-age, occupant-number and tenure within a given housing stock or target region, we assume no prior knowledge of age or mortality rate. Instead of individuals we focus the estimate of exposure duration on a specified number of homes within the target region. Exposure duration is simulated using only the tenure-based mobility that results in an occupant moving out of geographic region. A simple model is developed to simulate occupancy period over a specified time period based on the number of owner occupied and renter occupied homes within the target region. This model was found to be in good agreement with existing models that use a range of surrogate data.

Final recommendation for exposure duration distribution: Readily available demographic information is used to design a sampling strategy for simulating the target population. An appropriate parametric model, typically a two-parameter exponential model, is fit to the simulated data for occupancy period across the housing stock in the exposure district. For a small target region the sampling and fitting process can be repeated several times to estimate uncertainty about the model parameter.

Water Intake

Water intake can also be an influential input to a PRA. Data for water intake are from the same national survey that includes self-reported values of body weight but the values for water intake are from "recall" data, which may impact the quality of the data in ways that are difficult to predict. More importantly, long-term average water intakes for the population are estimated from short-term (2-day) sampling periods in these surveys. Although short-term data relates well to the overall population mean it greatly inflates the variance in the distribution of average daily intake values for individuals in the target population. Without correcting the inflated variance in the water intake data it is not possible to improve the relevance of the distribution using more complicated statistical models or demographic characteristics of the population. Therefore, we do not provide detailed information about possible demographic differences across the population. Rather, we evaluate the implications of using short-term data and develop recommendations for addressing these limitations with further study.

Recommendations for water intake: Lognormal distributions appear to provide a reasonable fit to water intake data, (Burmaster, 1998; Roseberry and Burmaster, 1992), but a recent report suggests that a more complex 5-parameter generalized F-distribution may be warranted (USEPA, 2000b). Our own research indicates that a mixture model is appropriate. However, without addressing the issue of variance inflation caused by using short-term survey data, it is not possible to identify the most appropriate parametric model or to recommend a transparent approach for constructing task-specific distributions. Developing a method for improving the relevance of short-term data is currently beyond the scope of this project but recommendations for future research are provided.

1.0 Introduction

This report describes research carried out at the Lawrence Berkeley National Laboratory (LBNL) to assist the U. S. Environmental Protection Agency (EPA) in developing a framework for supporting and evaluating probabilistic risk assessments. Among the current needs of the exposure-assessment community is the need to provide data for linking exposure, dose, and health information in ways that improve environmental surveillance, improve predictive models, and enhance risk assessment and risk management (NAS, 1994). The U.S. Environmental Protection Agency (EPA) Office of Emergency and Remedial Response (OERR) plays a lead role in developing national guidance and planning future activities that support the EPA Superfund Program.

The OERR is in the process of updating its 1989 Risk Assessment Guidance for Superfund (RAGS) as part of the EPA Superfund reform activities. Volume 3 Part A of RAGS was recently released in draft form as an update to the existing two-volume set of RAGS (USEPA, 1999b). The update provides policy and technical guidance on conducting probabilistic risk assessment for both human and ecological receptors. Risk assessors now have the option of using probabilistic risk assessment (PRA) to provide supporting information for risk management decisions at Superfund hazardous waste sites. Consequently, risk managers and decision-makers will need to review and evaluate the adequacy of PRAs for supporting regulatory decisions.

In judging the adequacy of a PRA, it is essential that the reviewer first understand the adequacy of the input distributions used in the assessment. An approach is needed for evaluating input distributions, from the risk manager's perspective, using information that is easily calculated or readily available in the risk assessment documentation. In addition, given the level of effort required of both the risk assessor and the risk manager when performing and reviewing a PRA, a set of recommended distributions for commonly used and well-characterized exposure factors would help reduce the effort and improve the consistency of the PRA process.

1.1 Probability Distributions in Exposure Assessment

Estimating potential human exposure involves collection and analysis of large amounts of data coupled with the use of models. Because these data and models must be used to characterize individual behaviors, contaminant transport, human contact, and uptake among large and often heterogeneous populations, there can be large variability and uncertainty associated with exposure predictions.

One common approach to address variability and uncertainty in exposure and risk assessments is the practice of compounding upper bound estimates in order to make decisions based on a highly conservative estimate of exposure. Such compounding of upper bound estimates leaves the decision maker with no flexibility to address margins of error; to consider reducible versus irreducible uncertainty; to separate individual variability from true scientific uncertainty; or to consider benefits, costs, and comparable risks in the decision-making process. Because the compounding of conservative estimates does not serve the exposure assessment process well, there has been a growing effort to include variance propagation and uncertainty analyses into the risk assessment process. This effort has culminated in the recent release of RAGS Volume 3 - Part A, *Process for Conducting Probabilistic Risk Assessment* (draft).

For human populations, total exposure assessments that include time-activity patterns and micro-environmental data reveal that an exposure assessment is most valuable when it provides a comprehensive view of exposure pathways and identifies major sources of variability and uncertainty (McKone, 1994; McKone and Bogen, 1992; USEPA, 1999b, pg. xx). Monte Carlo sampling methods using carefully developed exposure factor distributions provide one of the most versatile and informative means for characterizing uncertainty and variability in exposure estimates (Cullen and Frey, 1999; Morgan and Henrion, 1990). The draft guidance in RAGS 3A describes the application of these methods to the risk assessment process. However, the effectiveness of probabilistic methods for characterizing and communicating risk is largely dependent on the risk assessor's ability to characterize the type and degree of uncertainty and variability associated with inputs in the risk model, particularly for those inputs that have a strong influence on the modeling outcome.

A significant amount of discussion and debate has focused, from the risk assessor's perspective, on the methods that are used to construct distributions for probabilistic risk assessment. However, little attention has been given to the issue of evaluating the input distributions that are ultimately used in the PRA from the risk manager or decision maker's perspective.

1.2 The Need for an Evaluation Process for Existing Distributions

Constructing a distribution function for a given parameter begins by assembling information from the literature, from surveys and/or from experience. This information or data is selected, if possible, to be consistent with the model and its particular application. The values will vary as a result of measurement error, spatial and temporal variability, extrapolation of data from one scenario to another and lack of knowledge. When the data is limited, imprecise or imperfect – which is most often the case – then the process of constructing a distribution can become highly subjective. The process becomes less subjective as the amount of data for a given

parameter increases. However, a large set of data does not necessarily imply that the data is directly relevant to the risk scenario or that a suitable distribution function exists.

Once the necessary information is acquired, the next step in the distribution development process is to identify an appropriate statistical model and to fit that model to the data. A number of papers, reports and books are available that describe – often in great detail – methods for selecting and/or constructing input distributions for probabilistic risk assessment (Conover, 1999; Cullen and Frey, 1999; D'Agostino and Stephens, 1986; USEPA, 1997b; USEPA, 1998; USEPA, 1999a; USEPA, 2000b). In addition, computer software is readily available for automating much of the distribution development process. When these methods are applied, one obtains a distribution that provides an optimum fit to the available data. However, the resulting distribution does not automatically provide the risk manager with a clear understanding of how well the distribution replicates either the underlying data or the true range of the exposure factor being represented.

The quality or validity of the input distributions used in a probabilistic analysis directly influences the reliability of the model outcome, and the credibility of decisions that are based on that outcome. A significant amount of information is available to the risk assessor for constructing input distributions for the PRA. However, the problem of judging the task-specific adequacy of these distributions from the risk manager or decision-maker's perspective has not been sufficiently addressed. There is a need for a consistent yet flexible evaluation approach that can be used by regional risk managers when evaluating PRAs.

In order to develop an approach for evaluating existing exposure factor distributions, we first identify the key attributes that directly influence the adequacy of a given distribution for a specified task. At a minimum, the attributes must characterize (i) the importance of the input distribution relative to other sources of uncertainty, (ii) the adequacy of the information that is used to develop the distribution and (iii) the performance of the final distribution relative to the original data and supporting evidence. Once identified, these attributes can be used to provide a systematic approach for evaluating distributions and judging their adequacy in the context of a particular risk scenario.

1.3 The Need for a Set of Recommended Exposure Factor Distributions

As the number of inputs in a probabilistic risk assessment increases, so does the complexity of the analysis, and subsequently, the amount of effort required of both the risk professional performing the analysis and the risk manager reviewing it. For deterministic risk assessments, default inputs have improved the ease and the consistency of both performing and reviewing assessments. However, the idea of providing default distributions for exposure factors used in PRA can be inconsistent with probabilistic analysis. The problem is that the desire for a

small number of "conservative" or "health protective" distributions is often at odds with the need to insure adequate representation of the target population and explicit treatment of variability and uncertainty. This tradeoff between simplicity and relevance suggests the need for a process that identifies and incorporates the important inter- and intra-individual characteristics of the target population in a flexible and transparent manner.

There are a number of tools and methods available for constructing input distributions for PRA. However, given the need for task-specific relevance, any single distribution will almost certainly be inadequate for representing different populations and exposure scenarios. What is needed is a transparent approach that balances consistency with flexibility. The approach taken in this report is to develop, for a given exposure factor, a set of recommended distributions that are relevant to each key demographic category. These demographically based distributions are then used to construct task- or cohort-specific distributions, with uncertainty and variability, based on readily available information about the size and demographic characteristics of the target population. A validated set of demographically based distributions can provide the necessary consistency for the distribution development process and a sampling scheme based on census data can provide the flexibility needed to address the range of target populations and exposure scenarios.

1.4 Objectives of this Study

The overall focus of this report is to improve the ease and consistency both of performing PRAs and evaluating the adequacy of existing risk assessments. This overall focus is divided into two primary objectives. The first objective is to provide a clear and concise monograph on the critical attributes of distributions, from the risk manager's perspective, along with general recommendations for an evaluation process that relates these attributes to the task-specific adequacy of a distribution. The second objective is to consider when "default" distributions might be appropriate for use in PRA and to move toward development of recommended distributions for a small set of frequently used and well characterized exposure factors.

1.5 Overview of the Report

The remainder of this report consists of eight sections. In the next section, Section 2, we provide background on the development and use of distribution functions in exposure and risk assessment. The remaining chapters are divided into two parts. Part 1 (Sections 3-6) introduces the critical attributes for evaluating existing distributions, beginning in Section 3 with a discussion of sensitivity analysis, followed by the attributes that are specific to the data (Section 4) and those that are specific to the distribution (Section 5). The last section in Part 1 (Section 6) describes a process for systematically considering each of the critical attributes in a way that

both organizes the information for judging the adequacy of the distribution and facilitates of focuses dialogue with the risk assessor and other risk professionals. In Part 2 of this report we begin to develop recommendations for task-specific distributions for several commonly used and well characterized exposure factors beginning with *body weight* (Section 7) followed by *exposure duration* (Section 8) and *water intake* (Section 9).

2.0 Background

An important step in the process of conducting an uncertainty analysis is defining distributions for model inputs that contribute to variance in the outcome. These distributions can be in the form of probability density functions (PDF), cumulative distribution functions (CDF), or empirical distribution functions (EDF). The purpose of these distributions is to define the range of values that a variable can take on and to assign a probability of obtaining any particular value within that range. One step in the process of constructing a probability distribution is to define the range and moments of the input data. Once this is done, various methods (subjective, graphical, and statistical) can be used to select and parameterize an appropriate statistical model or distribution and judge the effectiveness with which the distribution describes the data.

2.1 Probabilistic Inputs and Models

Exposure models are used to describe the relative magnitude and variation in human contact with environmental contaminants. An important feature of exposure models is the ability to account for factors that influence variation in human contact (i.e. age, sex, location, activity patterns and physiological characteristics). Uncertainties limit the ability of models to fully characterize these relationships. Uncertainty in model predictions can be from a number of different sources, including specification of the problem; formulation of the conceptual model; formulation of the computational model; estimation of input values; and calculation, interpretation, and documentation of the results. Of these, only uncertainties due to estimation of input values can be quantified in a straightforward manner using error propagation techniques. Uncertainties that arise from errors associated with the specification of the problem or the model formulation are clearly important but fall outside of the scope of this report.

Point value inputs to models fail to express exposure variation and the uncertainty that arises from the use of incomplete and proxy data. Such issues can be addressed in part with the use of probability distributions as inputs to models. The value of information derived from a probabilistic analysis is very much dependent on the care given to the process of constructing the input parameter distributions in the context of the overall assessment. The data, scenarios, and/or

models used to represent this overall assessment of exposures to environmental contaminants include at least five important relationships:

(i) The magnitude of the source medium concentration, that is, the level of contaminant in the air, water, soil, and food with which the population has contact;

(ii) the contaminant concentration ratio, which defines how much a source medium concentration changes as a result of transfers, transformation, partitioning, dilution etc. prior to human contact;

(iii) the rate of human contact, which includes the frequency (days per year or hours per day) and magnitude (m^3/day, L/day or kg/day) of human contact with a potentially contaminated exposure medium;

(iv) the duration of potential contact for the population of interest as it relates to the fraction of lifetime during which an individual is potentially exposed; and

(v) the averaging time for the type of health effects under consideration, i.e. is the appropriate averaging time the cumulative duration of exposure (as is typical for cancer and chronic diseases) or some relatively short time period (as is the case for acute effects).

Taken together, these relationships define the risk scenario and must be considered both individually and in concert when developing and evaluating distributions. Not only should the distributions provide adequate representation of each exposure factor, it is also important that the various relationships that make up a PRA receive homologous treatment (i.e., the level of detail and sophistication should be comparable across the entire process). For example, developing a sophisticated treatment of variability and uncertainty for intakes may not improve the credibility of the assessment if a comparable level of knowledge is not available for characterizing the concentrations in exposure media. As a result, acceptable levels of uncertainty will differ from one exposure scenario to the next and a distribution that is adequate for one situation is not necessarily adequate for the next.

2.2 Variability and Uncertainty

One of the issues in uncertainty analysis that must be considered is how to distinguish between variability and uncertainty during the characterization of inputs and modeling results. Variability refers to quantities that are distributed empirically–such factors as rainfall, soil characteristics, weather patterns and human characteristics that come about through processes that we expect to be stochastic because they reflect actual variations in nature. These processes are inherently random. In contrast, true uncertainty or model-specification error (e.g., statistical

estimation error) refers to an input that, in theory, has a single value but the value cannot be known with precision due to measurement or estimation error.

In many situations, an exposure model is used to characterize the relative magnitude and importance of parameter uncertainty (lack of information) versus parameter variability (inter-individual variation). It is important to distinguish between the two forms of variance because each plays a unique role in decision making. Uncertainty can be reduced by further study while variability can only be better characterized (i.e., uncertainty about the variability might be reduced through further study).

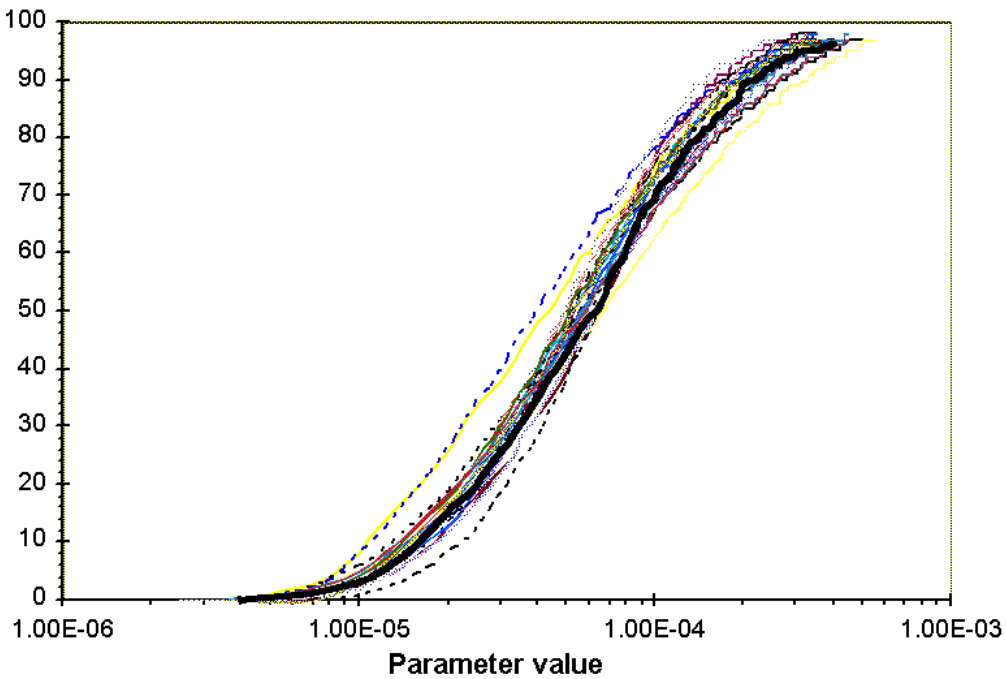


Figure 1: The family of cumulative probability plots reflecting both variability and uncertainty. Each curve expresses a realization of a variability distribution at a different level of uncertainty. Various methods are available for transforming the family of curves into confidence regions about the distributions.

To fully express the combined impact of uncertainty and variability, it is sometimes necessary to carry out a two-dimensional (2-D) Monte Carlo simulation consisting of an inner set of calculations embedded within an outer set. Bogen and Spear (1987) first described this approach. In the first phase, a single realization is obtained from the distribution of each uncertain parameter, followed by repeated sampling from the variable parameters. This process

is repeated until a large number of uncertain parameter value sets are taken in the outer phase and a larger number of the variable parameter values are used in the inner set of calculations. The simulation results are plotted as either a two-dimensional surface or as a family of variability curves representing uncertainty. Typical results for this type of simulation are shown in Figure 1.

In short, two-dimensional Monte Carlo simulations require that probabilistic inputs be parameterized not only with the statistical descriptors of the model (i.e., mean, standard deviation, mode, scale, etc.) but also with distributions representing uncertainty about those parameters as well.

2.3 The Link between Data and PDFs

Although it is occasionally appropriate to use raw data or a custom distribution to characterize the inputs for a probabilistic analysis, it is often useful to present the range of values in terms of a standard probability distribution. It is important that the selected distribution be matched to the range and moments of the available data. Although many different distributions are available to choose from (Evans *et al*., 1993; Johnson *et al*., 1994a; Johnson *et al*., 1994b), some of the more commonly used statistical models include the Normal, the Lognormal, the Uniform, the Log-uniform, and Triangular (Cullen and Frey, 1999; McKone *et al*., 1996).

Probability distributions are typically displayed as either probability density functions (PDFs) or as cumulative distribution functions (CDFs). For a continuous distribution, the PDF is a smooth function, $f(x)$, which represents the density of probability that a variate $\boldsymbol{X}$ has a value between $x - \Delta x/2$ and $x + \Delta x/2$, where Δx is an infinitely small interval (Morgan and Henrion, 1990). In a CDF, $F(x)$ represents the probability that a value realized by the variate $\boldsymbol{X}$ is less than or equal to x. Figure 2a shows a PDF for a continuous distribution and Figure 2b shows the corresponding CDF for this distribution. These two diagrams illustrate the PDF and CDF that would be used to represent a normal distribution with mean value, 1, and standard deviation, 0.3. Continuous distributions, like the examples in Figure 2, may be separated into three categories: (1) Those that represent only variability, (2) those that represent only uncertainty and (3) those that represent both variability and uncertainty.

There are a number of methods for constructing a parametric PDF from observations (data). These include moments matching, graphical methods, goodness of fit approaches, maximum likelihood methods and Bayesian methods, among others (Cullen and Frey, 1999; D'Agostino and Stephens, 1986; McKone *et al*., 1996; USEPA, 2000b). When these methods are applied, one obtains a distribution that has been optimized to fit the available observations. Often, more than one parametric model can satisfactorily fit a data set but the function used to optimize the fit rarely provides decisive information about which of two models is better. Thus,

the final choice of model (distribution type) is often based on a combination of theory, convention and/or convenience (Thompson, 1999). Each step in the distribution development process must be considered in the context of the overall risk scenario.

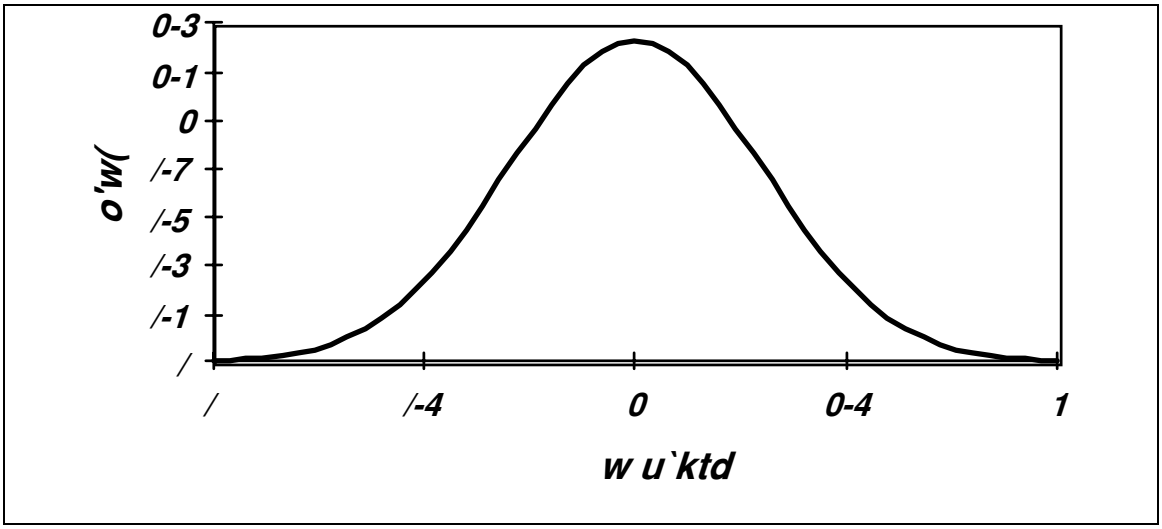


Figure 2a: Probability density function for a Normal distribution with mean of 1 and standard deviation of 0.3.

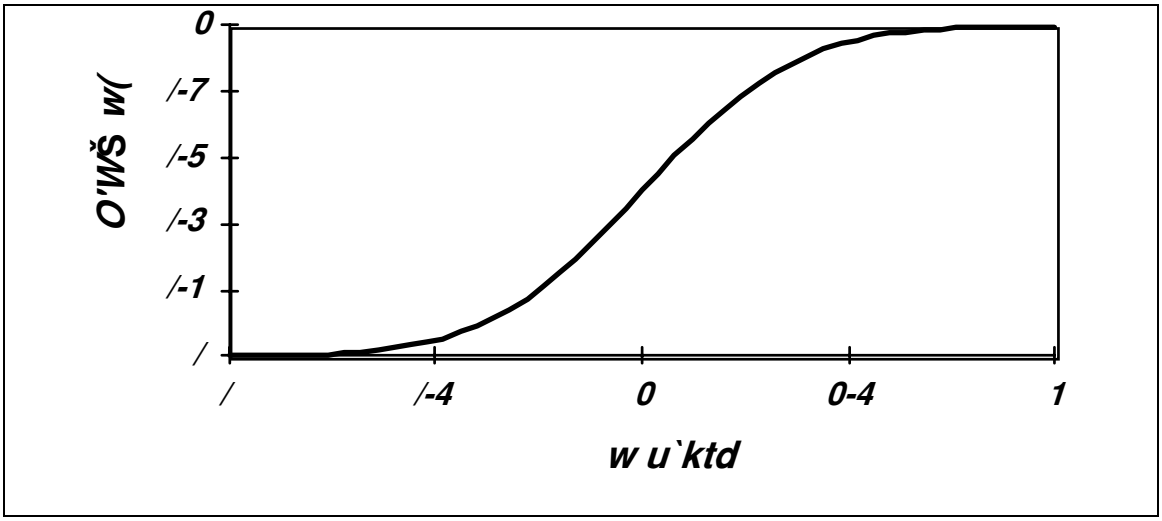


Figure 2b: Cumulative density function for a Normal distribution with mean of 1 and standard deviation of 0.3.

2.4 A Tiered Approach to Uncertainty/Variability Analysis

As discussed earlier (section 2.1) compounding of upper bound estimates for exposure factors is no longer the generally accepted approach to exposure and risk assessment. A more informative approach is one that provides the decision maker with flexibility to address margins of error; to consider reducible versus irreducible uncertainty; to separate individual variability from true scientific uncertainty; and to consider benefits, costs, and comparable risks in the decision making process. In order to make an exposure assessment consistent with such an approach, it should have both sensitivity and uncertainty analyses incorporated directly into an iterative process by which premises lead to measurements, measurements lead to models, models lead to better premises, and better premises lead to additional, better-informed measurements. In 1996, the U.S. EPA Risk Assessment Forum held a workshop on Monte Carlo Analysis (USEPA, 1997b). Among the many useful discussions at this meeting was a call for a "tiered" approach for probabilistic risk analysis (PRA) that is iterative and progressively more complex. Such a tiered approach required new methods in the exposure assessment community and greater demands on the number and types of exposure measurements that must be made.

According to the summary report of this workshop (USEPA, 1997b), at least three tiers are needed. First, the variance associated with all input values should be clearly stated and the impact of these variances on the final estimates of risk assessed. At a minimum, the estimation error or the experimental variance associated with the parameters should be included when these values or their estimation equations are defined. A clear summary and justification of the assumptions used for each aspect of a model should be provided. In addition, it should be stated whether these assumptions are likely to result in representative values or conservative (upper bound) estimates. Second, a sensitivity analysis should be used to assess how model predictions are impacted by model reliability and data precision. The goal of a sensitivity analysis is to evaluate and rank the input parameters, and modeling assumptions, on the basis of their contribution to variance in the output. Third, variance propagation or error analysis techniques (including but not necessarily limited to Monte-Carlo methods) should be used to carefully map how the overall precision of risk estimates are tied to the variability and uncertainty associated with the models, inputs, and risk scenarios.

2.5 Justification for an Evaluation Process and Recommended Distributions

The numerous modeling relationships that make up a risk scenario (section 2.1), the need to separate uncertainty and variability (section 2.2), the subjective nature of the distribution development process (section 2.3), and the iterative or tiered approach to PRA (section 2.4), combine to increase the difficulty associated with evaluating PRAs. Guidance is readily available for constructing distributions (Cullen and Frey, 1999; D'Agostino and Stephens, 1986; Seiler and Alvarez, 1996; Thompson, 1999; USEPA, 2000b) and for performing PRA

(Burmaster and Anderson, 1994; Cullen and Frey, 1999; Morgan and Henrion, 1990; Smith, 1994; Thomspon *et al*., 1992; USEPA, 1997b; USEPA, 1999b). However, the guidance is often written specifically for the risk assessor or statistician, with little or no guidance available for the decision maker or risk manager who must evaluate a PRA and judge the adequacy of this information before using it to support regulatory decisions.

This report is different from most others in the probabilistic risk analysis literature in that it approaches the issue from the perspective of the "Risk Manager" charged with using the output from a stochastic risk model. The authors recognize that judgement about the task-specific adequacy of a PRA, and the individual distributions within the PRA, is best left to the regional risk manager. Even if it were possible to entirely remove subjectivity from the review process, it would not be advisable to reduce the process to a simple scorecard. The risk manager brings a level of familiarity and understanding of the specific risk scenario to the evaluation process that cannot be captured in a scorecard or checklist. However, we do think that a concise description of the critical attributes that most influence the adequacy of a distribution can improve the consistency of the review process without removing the human element. Systematically working through the attributes can facilitate constructive dialogue between the risk manager, the risk assessor and other risk professionals focusing attention on the key elements of the PRA and providing consistency to the overall process. Therefore, the first objective of this report is to provide a monograph on the critical attributes of distributions, from the risk manager's perspective, along with general recommendations for relating these attributes to the task-specific adequacy of an input distribution.

Even with simple guidelines for reviewing distributions, PRAs can be very labor intensive for both the risk assessor and the risk manager. It has been shown that even when large, high quality data sets are available, risk practitioners will often select different statistical models to represent the data and the rational for the selection of different exposure factor distributions is often hard to discern (Binkowitz and Wartenberg, 2001). Default inputs have improved both the ease and the consistency of deterministic risk assessments. However, it remains unclear how appropriate default distributions might be for PRA. A single recommended distribution, or even a small set of "defaults" to choose from, would almost certainly limit the risk practitioner's ability to adequately model a specific target population and risk scenario. Thus, the second objective of this report is to attempt to strike a balance between the consistency of default inputs and the need to maintain flexibility and relevance in the PRA process. The approach developed in Part 2 of this report begins with raw data and readily available information about the target population. We then provide a transparent process for combining these two pieces of information in a way that is consistent across assessments but also provides enough flexibility to achieve relevance for a specified risk scenario.

Part 1

Critical Attributes for Evaluating Distributions

<Blank Page>

Part 1
Critical Attributes for Evaluating Distributions

For this part of the report, we draw on numerous studies, reports, workshops and our own experience, to identify and summarize the critical attributes of a probabilistic input distribution. Even though there are many different ways to develop a distribution, we find that the information needed to judge its task-specific adequacy can be summarized using seven key attributes. For simplicity, we group these attributes into three categories that only coincidentally resemble the three tiers introduced earlier for development of a probabilistic risk assessment (section 2.3). The first category, which directly influences the level of scrutiny one gives to an exposure factor input, considers the importance of the distribution in the context of the overall probabilistic risk analysis, or, put another way, the sensitivity of the PRA to the distribution. The second category considers the original information (data, summary statistics, expert judgement, etc.) that provides the foundation of any constructed distribution. The key attributes in the second category are quantity, quality and relevance of the data. The final category considers how well the final distribution represents the original information and the degree to which the performance of the distribution can be confirmed or validated using alternate models, assumptions, theoretical underpinnings and/or independent data. The attributes in the third category include statistical goodness-of-fit, graphical goodness-of-fit and performance evaluation or cross-validation results.

The risk manager's judgment about the overall adequacy of a distribution will likely consider all seven attributes in concert but for clarity we first summarize each category (and the specific attributes of each category), then we introduce a systematic evaluation process for tracking and communicating the information. Some discussion is provided for the individual attributes with respect to their overlap and interaction and how this can impact the adequacy of the final distribution. The last section in Part 1 provides a graphical paper-based tool and systematic approach for evaluating the attributes of an exposure factor distribution. This tool combines the various types of information in a way that facilitates judgement about the task-specific adequacy of the distribution. The process is designed to be flexible, allowing the risk manager to weigh information about the attributes and maintain the transparency that is necessary for constructive dialogue among the risk practitioners involved in the assessment.

3.0 Category One: The Sensitivity Analysis

There are many methods for investigating the relationship between variation in the inputs to a model and variation in the model outcome (Campolongo *et al*., 2000). These investigations,

collectively termed sensitivity analyses (SA) play a vital role in the set-up, use and evaluation of probabilistic risk models. Therefore, all PRAs should include results and discussion from some form of sensitivity analysis (USEPA, 1997b, USEPA, 1999 #101). These results will likely be reported as sensitivity scores, contribution to variance, correlation coefficients, or regression coefficients. Results from the SA can help the reviewer identify the inputs that contribute the most to variance in the model output and therefore require the closest scrutiny during the evaluation process. We begin with an overview of the types of sensitivity analysis that a reviewer might see in a PRA. We then discuss approaches for using results of the SA in the distribution evaluation process.

3.1 The Myriad of Sensitivity Analysis Techniques

The relationship between the variance (uncertainty and variability) in model inputs and variance in the model outcome is generally estimated using either analytical methods or simulation methods. Analytical methods use mathematics to derive an expression for the model outcome variance. If the analytical expression is exact, it can be related directly to the input variances. If approximate, some steps in the derivation may require alternative approaches. Simulation methods provide numbers, histograms and/or curves from which the model outcome variance is read.

In order to discuss sensitivity analysis, we first recognize that any mathematical model produces an outcome Y, such as risk, that is a function of w input variates, X_i, and time, t,

$$Y = f\,(X_1, X_2, X_3, \ldots X_w, t) \qquad (1)$$

The explanatory variables, X_i, represent the various inputs to the model. As applied to a mathematical model, sensitivity analysis involves measuring changes in model outcome (Y) that result from changes in individual model inputs (X_i). The goal of a sensitivity analysis with respect to the PRA evaluation process is to rank the X_is on the basis of their contribution to variance in the output.

Sensitivity analyses can be performed on either a local or global scale. A local sensitivity analysis examines the effects of small changes in X_i at some specified point in the range of input/outcome values. A global sensitivity analysis quantifies the effects of variation in X_is over their entire range of values. The global sensitivity analysis requires a probabilistic sampling method (e.g., Monte Carlo or Latin Hypercube sampling) as a starting point.

3.1.1 Local Sensitivity Analysis Techniques

Iman and Helton (1988) and others (Morgan and Henrion, 1990; Saltelli *et al*., 2000) describe different measures of local sensitivity. The simplest form is given as the rate of change

in the output with respect to change in the input evaluated at the nominal or base case given by $X°$ (typically in the central range or at the mean value of the output).

$$U_{S(x,y)} = \left[\frac{\partial y}{\partial x}\right]_{X°} \tag{2}$$

The problem with the simple sensitivity score is that the scale or units of the input influences the measurement of sensitivity. To normalize the sensitivity measure and remove its dependence on scale an "elasticity score" can be used, such as

$$U_{E(x,y)} = \left[\frac{\partial y}{\partial x}\right]_{X°} \frac{x°}{y°} \tag{3}$$

where $x°$ and $y°$ are the base case values or quantiles of the input and output. Elasticity is simply the ratio of percent change in the output caused by a given percent change in the input.

Although better than the simple measure, the elasticity measure still has the limitation that it does not account for uncertainty. Even an exposure factor with a very high elasticity score can be of little importance if we know the value of the input with a high degree of certainty. To account for the level of uncertainty in each input, the coefficient of variation (*CV*) for the input is used as a normalization factor instead of the ratio $x°/y°$ where *CV* is the ratio of the arithmetic standard deviation to the arithmetic mean.

3.1.2 Exact Analytical Methods

For many mathematical operations, including addition, subtraction, multiplication, division, logarithms, exponential, power relations, etc. there are exact analytical expressions for explicitly relating input variance and covariance to variance in the model predictions (Bevington, 1969). In analytical methods, the mean, variance, and covariance matrix of the input distributions are used to determine the mean and variance of the outcome. These methods are often referred to as error analysis. Textbooks such as *Data Reduction and Error Analysis for the Physical Sciences* by Bevington (1969) give variance propagation formulae for a range of mathematical operations. Many models used in risk assessment can be simplified to a chain of additive and multiplicative terms using these formulas.

The following is an example of the exact analytical approach. If ψ is the product of x times y times z, then the equation for the mean or expected value of ψ, $E(\psi)$, is:

$$E(\psi) = E(x) \times E(y) \times E(z) \tag{4}$$

The variance in ψ (the standard deviation squared) is given by:

$$\sigma_{\psi}^{2} = [\mathrm{E}(\psi)]^{2}\left\{\frac{\sigma_{x}^{2}}{[\mathrm{E}(x)]^{2}} + \frac{\sigma_{y}^{2}}{[\mathrm{E}(y)]^{2}} + \frac{\sigma_{z}^{2}}{[\mathrm{E}(z)]^{2}}\right.$$

$$\left. + \frac{2\sigma_{xy}^{2}}{[\mathrm{E}(x)\times\mathrm{E}(y)]} + \frac{2\sigma_{xz}^{2}}{[\mathrm{E}(x)\times\mathrm{E}(z)]} + \frac{2\sigma_{yz}^{2}}{[\mathrm{E}(y)\times\mathrm{E}(z)]}\right\} \tag{5}$$

where σ_{x}^{2}, σ_{y}^{2}, σ_{z}^{2} are the variances of the variates *X, Y, Z,* and σ_{xy}^{2}, σ_{xz}^{2}, σ_{yz}^{2}, are, respectfully, the covariance of *x, y* and *z,* (Bevington, 1969). The covariance terms are defined as:

$$\sigma_{xy}^{2} = \sum_{i=1}^{n}[(x_i - \mathrm{E}[x])(y_i - \mathrm{E}[y])] \tag{6}$$

The expected value, taken here as the arithmetic mean ($\bar{x}$), is used to represent all inputs derived from a set of measured values, assuming equal sampling probability—even those that might have geometric distributions. The $\bar{x}$ is computed by summing the *n* observations in the dataset and dividing this sum by *n*:

$$\textit{Arithmetic mean } \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{7}$$

The variance (unbiased) is calculated as:

$$\textit{variance}\left(\sigma^{2}\right) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^{2}}{n-1} \tag{8}$$

The standard deviation (σ) is the square root of the variance and the coefficient of variation (CV) is $\sigma/\bar{x}$. It should be noted that based on the central limit theorem of statistics, the confidence associated with the estimate of $\bar{x}$ becomes large as the number of samples used to estimate $\bar{x}$ also becomes large (Snedecor and Cochran, 1989). Consequently, the reliability of the mean and *CV* estimates of a variable is low when the sample size is small. An estimate of the error associated with estimating a mean from a small sample size is the standard error of the mean [*S.E.* $(\bar{x})$] given by:

$$\textit{Standard error of the mean } (S.E.(\bar{x})) = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^{2}}{n(n-1)}} \tag{9}$$

Bevington (1969) lists variance propagation solutions like the one illustrated in Equations 5 through 9 for several mathematical operations. These relationships hold regardless of the shape of the input distributions. For situations where covariance is negligible, variance propagation solutions can be used to develop simple estimates for each input's contribution to outcome variance. However, the analytical method can not be used to estimate confidence intervals or percentiles of the output variable unless we have prior knowledge of the shape of the output distribution.

3.1.3 Global Sensitivity Analysis Techniques

When random sampling methods such as Monte Carlo or Latin Hypercube sampling are used to estimate outcome variance by propagating input variance through a given model, several measures of sensitivity can be applied. These include graphical methods, approximate analytical methods, correlation- and regression-coefficients in both the units of the data and rank-transformed space. In addition, variations on the standard approaches such as regional sensitivity analysis and different forms of probabilistic sensitivity analysis are also available for characterizing different aspects of model dependencies on input variance. Each of these methods is briefly described below.

3.1.3.1 Graphical Methods of Analysis

It is often useful to begin the process of understanding the nature and strength of relationships between inputs and outputs of a model with visual observations. This can be accomplished using scatter plots of model output values plotted against the random values selected for each input/output pair in the analysis. These plots can provide insight, show nonlinear effects, thresholds, and other subtle but potentially important relationships. In addition, scatter plots are useful in diagnostic analyses to help determine if the model is working as intended (Iman and Helton, 1988) or if the sensitivity is dominant over a given range of the data (Helton *et al*., 1989).

When performing a graphical analysis, one has several options for viewing data. Two of the more common options are plotting values in the original units of the data or plotting in rank-transformed space. The rank-transformed values are determined by ranking them, and indexing them by their order position. (Cullen and Frey, 1999, pg. 259; Snedecor and Cochran, 1989, pg.194). Viewing the data according to their rank can remove the influence of outliers and can help to highlight nonlinear and continuously increasing or decreasing (i.e., monotonic) relationships. However, the advantages of working with rank-transformed data are often much more pronounced in the correlation and regression analyses summarized later in this section.

3.1.3.2 Empirical Methods

When a probabilistic analysis is performed, numerical methods are needed to estimate the individual contribution of various inputs to model outcome variance. The most commonly used

numerical methods for ranking uncertain parameters are correlation and regression methods. These methods are described in the Safety Series 100 report of the International Atomic Energy Association (IAEA, 1989) and in other reference books such as those by Draper and Smith (1981), Freund and Minton (1979) and Saltelli et al. (2000). A brief overview is provided here. The reader should refer to the original references for details on each method.

Standard Correlation Coefficients: A simple, but not necessarily reliable way to derive a ranking for model inputs in a Monte-Carlo simulation is to compute the correlation coefficient between model predictions and the random variates of each model input. The correlation coefficient, ρ_{YX}, is a measure of the degree of association between two variates that are distributed over a range such that

$$\rho_{YX} = \sigma^2_{YX} \Big/ [\sigma_Y \sigma_X] \tag{10}$$

where σ^2_{YX} is the covariance between input X and output Y, and σ_X and σ_Y are the standard deviation of the variates X and Y respectively.

For a Monte-Carlo type simulation with m realizations, the m pairs of model input/output values are correlated using the relationship between each X_j and associated outcome Y. From a given random sample $(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)$, of size m from the joint distribution of Y and X, the sample or empirical correlation coefficient ρ_{YX} is obtained from

$$\rho_{YX_j} = \frac{\sum_{i=1}^{m}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{m}(x_i - \bar{x})^2 \sum_{i=1}^{m}(y_i - \bar{y})^2}} \tag{11}$$

The sample correlation coefficient is a measure of the degree of linear relationship between the set or simulated X and Y variates. This correlation gives an estimate of the linear contribution of the input to uncertainty in the model prediction. By averaging the effect of each input over the joint probability distribution for all other inputs, it is inherently a global measure of uncertainty importance (Morgan and Henrion, 1990).

Partial Correlation Coefficients: Iman and colleagues (1980; 1988) have also suggested the use of partial correlation coefficients (PCC), which is a measure of the contribution of an input to the output variance, after removing the effects attributable to the other inputs. PCC are particularly useful when there is significant correlation among inputs. PCC indicate the degree of linear relationship between those portions of the model prediction and the input that cannot be explained by a linear relationship of each to the remaining inputs (Iman and Helton, 1988). The partial correlation coefficient is calculated by first developing a set of regression models

$$\hat{y} = b_0 + \sum_{k \neq j} b_k x_k$$

$$\hat{x}_j = c_0 + \sum_{k \neq j} c_k x_k$$

where $\hat{y}$ and $\hat{x}$ are the outcome and input of regressions for X_j, respectively, b_0 and c_0 are intercepts, and b_k and c_k are weighting coefficients. The partial correlation coefficient for the jth input is then calculated as the correlation between $Y - \hat{Y}$ and $X_j - \hat{X}_j$.

Correlation methods can over-predict the sensitivity of the outcome to a particular input when outliers are present in the input/output pairs. Likewise, the method can under-predict the importance when the relationship is nonlinear but monotonic. One way to reduce misleading results from the sensitivity analysis is to transform the values into ranks as described earlier (section 3.1…) and perform the correlation analysis on the ranks rather than the actual values.

Regression Models and Standardized Regression Coefficients: Another approach to evaluating the strength of the relationship between inputs and outputs is based on a regression analysis. Consider a least squares regression model fit to estimate the output, Y, as a linear function of the inputs, X_j. (Note that the index j is over all inputs, whereas k was over the m sample values from a probabilistic analysis for a given input)

$$\hat{y} = \beta_0 + \sum_{j \neq 1}^{w} \beta_j X_j \tag{12}$$

The β_j regression coefficients are measures of the linear sensitivity of Y to input X_j (Draper and Smith, 1981). Regression coefficients have the disadvantage that they depend on the units or scale of measurement of Y and X_j. A more useful measure of sensitivity, known as the *standardized regressions coefficient* (SRC), may be obtained by multiplying each coefficient by the ratio of the estimated standard deviations, s, of X_j to Y such that (Draper and Smith, 1981)

$$U_{src}\left(X_j, Y\right) = \frac{\beta_j \times s_y}{s_j} \tag{13}$$

It is often useful to perform stepwise regression, producing a sequence of linear models consisting of 1, 2, 3, up to the total number of the input variables respectively. Inputs are added one at a time to maximize the fit of the model according to an R^2 value. The sequence in which they are selected is a useful measure of their importance, as is the incremental increase in R^2 that they produce.

An alternative approach relates the changes in Y with respect to X_j, which is based on the partial derivative with respect to X_j. The coefficients a_j, j=1, 2, ..., w are from a multiple linear regression model of the results from a probabilistic analysis having w input variates

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_w X_w + \varepsilon \tag{14}$$

and represent the partial derivatives of Y with respect to X_j. The a_j indicate the change in Y associated with unit changes in X_j, all other X_k, $k \neq j$ remaining constant. If Y and X_j are standardized according to

$$Y' = \frac{y - \bar{y}}{s_Y} \text{ and } X'_j = \frac{x_j - \bar{x}_j}{s_j} \tag{15}$$

then the coefficients $\frac{\beta_j s_j}{s_Y}$, j = 1, 2, ..., w of the linear regression model with standardized variables

$$Y' = \frac{\beta_1 s_1}{s_Y} X'_1 + \frac{\beta_2 s_2}{s_Y} X'_2 + \ldots + \frac{\beta_w s_w}{s_Y} X'_w + \varepsilon \tag{16}$$

are called standardized partial regression coefficients. The standardized partial regression coefficients indicate how many standard deviation changes in Y are associated with one standard deviation change in X_j, all other X_k, k≠j, remaining constant.

As with the correlation methods, the regression methods are a measure of the linear relationship and do not perform well with nonlinear but monotonic input/output pairs or when outliers are present in the data. However, as with the correlation methods, regression methods can also be extended to evaluate rank-transformed data. For further details on these and other methods, the reader should consult the specific reference included in this section.

3.1.3.3 Approximate Analytical Method

In order to rank the relative importance of inputs that are truly uncertain versus variable inputs, McKone (1994) used an approximate analytical method to assess variance contributions from groups of parameters. Whereas exact analytical methods and numerical methods are used to rank the contribution of individual parameters, the approximate analytical method (McKone, 1994) ranks groups of parameters in exposure models.

Depending on the structure of the model, it is assumed that variance combines as either an additive, multiplicative, or geometric function. That is, if we consider that a given outcome, Y, comes about as a function of both uncertain (U) and variable parameters (V),

$$Y = f(U_1, U_2, \ldots, U_n; V_1, V_2, \ldots, V_n) \tag{17}$$

then the total variance (i.e., uncertainty and variability) in Y can be approximated from the mean and standard deviation of the U and V parameters according to following expressions, which are, respectively, additive, multiplicative, and geometric,

$$\sigma_Y^2 = \sigma_{U_1}^2 + \sigma_{U_2}^2 + \ldots + \sigma_{U_n}^2 + \sigma_{V_1}^2 + \sigma_{V_2}^2 + \ldots + \sigma_{V_n}^2 \tag{18}$$

$$\sigma_Y^2 = E^2[Y]\left\{\frac{\sigma_{U_1}^2}{\eta_1^2} + \frac{\sigma_{U_2}^2}{\eta_2^2} + \ldots + \frac{\sigma_{U_n}^2}{\eta_n^2} + \frac{\sigma_{V_1}^2}{\nu_1^2} + \frac{\sigma_{V_2}^2}{\nu_2^2} + \ldots + \frac{\sigma_{Vn}^2}{\nu_n^2}\right\} \tag{19}$$

$$\sigma_{\ln Y}^2 = \sigma_{\ln U_1}^2 + \sigma_{\ln U_2}^2 + \ldots + \sigma_{\ln U_n}^2 + \sigma_{\ln V_1}^2 + \sigma_{\ln V_2}^2 + \ldots + \sigma_{\ln V_n}^2 \tag{20}$$

where E[Y] is the expected value of Y and η_j and υ_j are expected values of the respective U and V parameters.

Because some inputs are correlated and the contribution to variance from a group of variables and not a single variable is modeled in Eqs. 18-20, it is not possible to determine *a priori* how the overall variance in Y is attributed to the uncertain and variable input groupings. In one method for evaluating contribution to variance, a number of realizations are used to generate a "base case" with all inputs treated as probabilistic. Next, a second set of realizations is generated using the same random number sequence but with either the uncertain set or variable set (exclusion case) treated probabilistically. Finally the same number of realizations is generated, again with the same random number sequence, but with only parts of input set allowed to vary (partial inclusion). This process is repeated for each parameter category.

The variances obtained from the base case, exclusion case, and partial inclusion are grouped as additive, multiplicative, and geometric to allow evaluation of variance attributable to uncertain, variable, or mixed parameters. The method is based on a predefined set of criteria that leads to a "robust" solution (McKone, 1994). The approach provides insight into model structure and interaction among various categories of inputs.

3.1.3.4 Regional Sensitivity Analysis

Spear and Hornberger (Hornberger and Spear, 1980; 1980) first described what they defined as a regional sensitivity analysis (RSA) to investigate parameter structure and parameter interaction in a complex environmental model. The method has since been used in a wide range of model systems ((Spear *et al*., 1994), and references therein). In the RSA method, random values are drawn from input PDFs and used to calculate a set of model outcomes. Each input is assigned a binary classification (yes-no, pass-fail) depending on whether or not the outcome is in agreement with some expected or observed behavior of the system (specified in advance). Spear

and Hornberger (1980) showed that, for a given parameter, the degree of separation between the CDF of the inputs classified as "pass" and the those classified as "fail", is an indication of the importance of that parameter toward the attainment of the specified outcome.

Beck and Chen (2000; Chen and Beck, 1998) recently modified the RSA approach to evaluate the sensitivity of a model at different regions of the outcome distribution. For example, the method presented by Chen and Beck can be used to identify inputs that influence the high end of the outcome distribution. In higher-order fate and exposure models (those having a large number of inputs) different inputs might be important to the outcome variance as one moves from the low end of the outcome distribution to the upper tail. This is particularly important when the assessment is directed at a specified outcome (i.e., estimate of risk in excess of some predetermined value).

3.2 Interpreting the Sensitivity Analysis Results

There are almost as many ways to interpret results from a sensitivity analyses as there are different ways for generating those results. Many of the sensitivity analysis techniques described above, and elsewhere (Saltelli *et al*., 2000), are designed as diagnostic and evaluative tools to help model developers gain insight into the level and source of uncertainty in a model and whether that uncertainty is acceptable given the modeling task (Scott *et al*., 2000). In addition, different types of sensitivity analyses will be used during different phases of the model development process. Simple screening level analyses will likely be used in the early phases of the modeling to get a general idea of the important inputs while more detailed analyses (Chen and Beck, 1998; Maddalena *et al*., 2001; McKone, 1994) might be performed during the final stages of model development and evaluation. Although absolute measures of sensitivity from the different methods do not always agree, the relative ranking of importance is typically conserved for the input distributions, particularly those that have a significant contribution to outcome variance (McKone *et al*., 1995).

Even though a value can be calculated for each probabilistic variable in the sensitivity analysis, not all inputs in a model contribute significantly to outcome variance. Although some exposure and risk scenarios require a large number of stochastic inputs to fully characterize the outcome variance, the majority of scenarios are controlled by a relatively small number of inputs (Maddalena *et al*., 2001; McKone and Bogen, 1992). This is particularly true in the higher order regulatory environmental models that have a large number of input variables. It is advantageous for both the risk assessor and risk manager to understand which inputs need to be treated as stochastic, which can be assigned point values and which can be excluded from the calculations. It is best to classify the inputs early in the model development process before investing resources into the development and evaluation of distributions that will not influence the outcome.

Therefore, results from a number of different sensitivity analysis exercises will likely be included in the dialogue between the risk assessor and risk manager as they work through the various tiers of a probabilistic risk assessment. The results reported in the final PRA will likely be given for each probabilistic variable as a measure of rank correlation (or partial rank correlation) generated directly from the final output of the probabilistic analysis. Rank correlation results are reported on either a scale from –1 to +1, (depending on whether increasing the input tends to decrease or increase the output) or as an approximate contribution to variance on a scale normalized to unity (or 100%). If some of the inputs in the risk model are correlated, the sensitivity analysis results should be generated and presented both with and without correlation among the inputs (or as partial rank correlation). When the model outcome is sensitive to one of the correlated inputs it can inflate or mask the importance of the other variable if correlation is included in the analysis.

Regardless of how the results are presented, dialogue during the risk assessment process and information in the final PRA should provide enough detail for the Risk Manager to have a clear understanding of the relative importance of the inputs in the model.

3.3 Relating Sensitivity Analysis Results to the other Attributes

While the remaining categories and attributes of the distribution (described in the following chapters) relate to the intrinsic properties of a particular distribution, the sensitivity analysis category relates the adequacy of those attributes to the overall assessment. Presumably, by the time the final PRA is completed, decisions about which inputs to treat as probabilistic, which to model as point values and which to exclude from the assessment will already have been made. In deciding that an exposure factor distribution is adequate for a given task, the risk professional is always making a subjective decision. Nevertheless, sensitivity analysis could and should provide valuable information for supporting the decision.

Distributions that contribute less than a few percent (normalized scale) to the outcome variance of the model will typically only require a quick check to insure that the central tendency and the approximate range of the distribution are appropriate. The remaining distributions should be taken to the next level of review and the information about the specific attributes should be weighted according to overall importance of the given input determined in the sensitivity analysis.

4.0 Category Two: Adequacy of the Data

Although there are many sources of information that can be used to construct distributions, the primary focus of this section is on the adequacy of raw data. For information on the adequacy of expert judgment in the distribution development process, the reader is referred to Morgan and Henrion (1990). For a discussion of the adequacy of summary statistics for development of distributions see the recent reports by Research Triangle Institute (Myers *et al*., 1998; USEPA, 2000b). However, the following discussion about the adequacy of raw data can be applied to summary statistics if information about the original source of data is available.

The three key attributes of the raw data used to develop distributions include quantity (number of samples in the data set), quality (precision and accuracy of the data collection process) and relevance (relationship of the data to the actual exposure factor within the population, cohort or scenario of interest). These attributes represent the knowledge base that provides the foundation for constructing a distribution. Lack of data or poor quality data does not necessarily mean one cannot construct an adequate distribution for the PRA. However, limitations in the data or gaps in the knowledge base must be acknowledged by explicitly including uncertainty in the distribution. Methods have been described that relate sample size to uncertainty in probability distributions (Cullen and Frey, 1999; Morgan and Henrion, 1990). Although much more subjective in nature, the issue of relating data quality and relevance to uncertainty has also been discussed (Morgan and Henrion, 1990; USEPA, 1999a).

4.1 Data Quantity Attribute

The quantity of data used to construct a distribution is important because uncertainty in the distribution is inversely related to the sample size. For very small samples, uncertainty about the selection of a particular parametric model increases greatly. In addition, precision of the various methods for estimating the value of model parameters (e.g., mean, standard deviation, mode, etc.) decreases with decreasing sample size. This lack of precision in the statistical parameters is much more pronounced for the moments representing the spread of the data than for the central tendency. Using a form of bootstrap sampling, which is a method for repeatedly sampling data to reduce bias and get estimates of variance, Cullen and Frey (1999, pp. 107-113) have also shown that skewness and kurtosis (i.e., the shape of the distribution) are influenced by sample size.

Although the central moments relating to shape can be useful to the risk assessor when choosing an appropriate statistical model to represent data, the primary concern for the risk manager is how sample size might influence uncertainty about the range and percentiles of the data (and subsequent distribution). Therefore, recommendations developed in this section for

evaluation of the data quantity attribute are focused primarily on the second central moment of the data and methods for estimating uncertainty about the percentiles.

4.1.1 Relating Sample Size to Precision in Distributions

Ershow and Cantor (1989) specify the minimum number of samples required to report various percentiles of water intake. The values were based on guidelines developed by the National Center for Health Statistics. Sample sizes needed to report the 50th, 75th, 90th, 95th and 99th percentiles were 6, 13, 36, 72 and 369, respectively (Ershow and Cantor, 1989, page 17). Another recent report on water ingestion (USEPA, 2000a) also gives estimates of necessary sample size for reporting percentiles based on minimum reporting requirements suggested by the Life Science Research Office (LSRO, 1995). Samples sizes recommended in the second report were estimated as {8×(variance inflation factor)/(1-percentile)} where the "variance inflation factor" (1.60) was specific to the survey being used (USEPA, 2000a, page 3-7). The resulting minimum sample sizes for the 50th, 75th, 90th, 95th and 99th percentiles were 26, 51, 128, 256 and 1280, respectively. The minimum sample requirements in the second of the two reports are approximately 3.5 times greater than the values recommended in the earlier report by Ershow and Cantor (1989).

One way to evaluate the relative precision provided by the recommendations for minimum sample size is to compare the approximate range in the order statistics that encompasses the specified percentile for a given sample size. Morgan and Henrion (1990, pages 83 and 202) derive a simple approximation for the number of random values (i.e. sample size) necessary to achieve a specified level of confidence that the actual fractile (percentile/100) of a distribution is within a specified range. The sample size (n) that is required to have a (1-α) degree of certainty that a specified fractile (p) is within a certain range Δp is

$$n = p(1-p)\left(\frac{Z_{(1-\alpha)}}{\Delta p}\right)^2 \qquad (21)$$

where $Z_{(1-\alpha)}$ is the standard normal deviate for the (1-α) probability. Derivation of Eq. 21 is based on order statistics and an assumption that the sample size is large enough to allow a normal distribution to adequately approximate the binomial distribution. For details on the derivation of Eq. 21, the reader is referred to the original text (Morgan and Henrion, 1990, pgs. 200 to 203).

Rearranging Eq. 21 to approximate the range encompassing p for various sample sizes we arrive at

$$\Delta p = \frac{Z_{(1-\alpha)}}{\sqrt{n/p(1-p)}} \tag{22}$$

Using Eq. 22 we estimate the Δp that encompasses a specified fractile with 95% confidence for various sample sizes and compare the estimates to those recommended in references (Ershow and Cantor, 1989) and (USEPA, 2000a). The estimated fractile range from Eq. 22 is independent of the shape of the original distribution giving rise to the data (Morgan and Henrion, 1990).

Table 1: Relating sample size to precision in order statistics [a]

Fractile of interest	Suggested sample size from (Ershow and Cantor, 1989, pg. 17)	Calculated range encompassing fractile with 95% confidence	Suggested sample size from (USEPA, 2000a, pg. 3-7)	Calculated range encompassing percentile with 95% confidence
0.50	6	0.30-0.70	26	0.40-0.60
0.75	13	0.63-0.87	51	0.69-0.81
0.90	36	0.85-0.95	128	0.87-0.93
0.95	72	0.92-0.98	256	0.94-0.96
0.99	369	0.985-0.995	1280	0.987-0.993

a. We provide an example of how to read this table. Given a sample size of n=72 (second column, fourth row), equation 22 says that we are 95% confident that the actual 95th percentile of the data falls somewhere between the 92nd and 98th percentile of the cumulative distribution (third column, fourth row).

The ranges calculated using Eq. 22 and reported in Table 1 are for the fractiles (or percentiles where percentile=fractile×100%) that are typically plotted on the ordinate axis (i.e., the y-axis) of a standard cumulative distribution plot. Relating this range to the precision in the value or quantile of interest to the risk manager (i.e., x-axis) requires knowledge of the variance in the data set or the spread of the data. For example, defining two lognormal distributions, **A** and **B**, with arithmetic means of unity but with standard deviations of 0.1 for **A** and 1 for **B** then a fractile range of (0.94, 0.96) (e.g., n = 256 in Table 1) would encompass a quantile range of 0.03 for distributions **A** and 0.4 for distributions **B**. Therefore, it is important for the reviewer to consider both the range encompassing the percentile of interest with a given confidence and the spread of the data (i.e., the standard deviation and the precision in the estimate of the standard deviation).

For estimating the precision about the standard deviation of a distribution as a function of sample size for normally distributed data, the following can be used (Snedecor and Cochran, 1989, pg. 75)

$$(n-1)s^2/\sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{\sigma^2} = \chi^2_{(n-1)} \tag{23}$$

where n is the sample size, s^2 is the estimated variance from the sample, σ^2 is the true variance and $\chi^2_{(n-1)}$ is the value of the chi-square distribution with (n-1) degrees of freedom. Rearranging Eq. 23 for a (1-α)% confidence interval of σ^2 we arrive at

$$\frac{\chi^2_{(1-\alpha/2,n-1)}}{(n-1)} < \frac{s^2}{\sigma^2} < \frac{\chi^2_{(\alpha/2,n-1)}}{(n-1)} \tag{24}$$

which expresses the precision of s^2 as a percentage of the true variance (σ^2) in the population. The two-sided 95% confidence limits for the ratio, s^2/σ^2, is illustrated in Figure 3. For a sample of size n=256, the 95% confidence interval for the variance ratio is (0.83, 1.18). In terms of the standard deviation (i.e., taking the square root of equation 24), the 95% confidence interval for the ratio of s/σ given a sample size n=256 is (0.91, 1.09), or approximately ± 0.09.

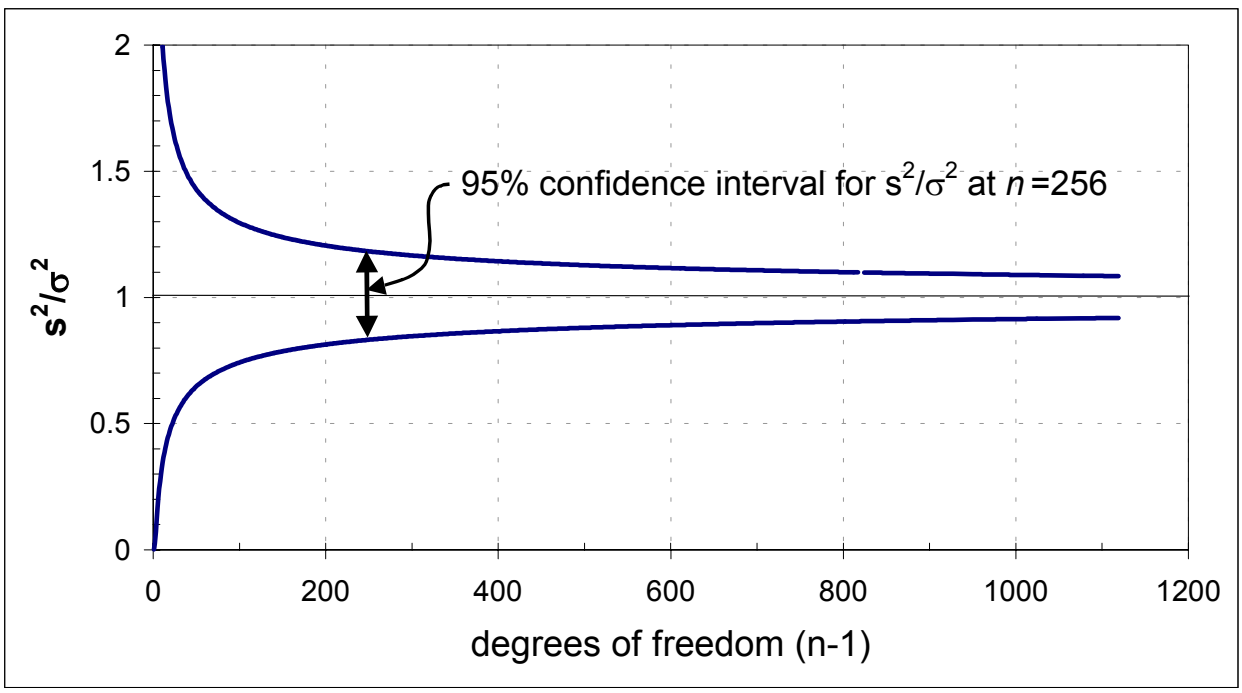


Figure 3: The 95% confidence interval for the ratio of the estimated sample variance to the true (unknown) variance σ^2 illustrating the relationship between sample size and precision as related to the spread of the data.

The precision of the sample standard deviation expressed as a percent of the true standard deviation can be related to a percentile of interest (e.g., the 95th percentile) for normally distributed data using the relationship

$$95^{th} \text{ percentile} = \bar{x} + 1.645s \tag{25}$$

where 1.645 is the *one-tailed* $Z_{0.95.}$ Therefore, according to Eq. 25, there is about a 95% chance that an estimate of the 95th percentile of a normally distributed sample of size n = 256 falls within a range 9% more or less than the actual 95th percentile.

For lognormal data expressed as $\log(x_i)$ the upper tail is given by Eq. 26 except that $\bar{x}$ and s are now the arithmetic mean and standard deviation of the log transformed data. The upper tail for lognormally distributed data can also be calculated by

$$95^{th} \text{ percentile} = GM \times GSD^{1.645} \tag{26}$$

where *GM* and *GSD* are the geometric mean and geometric standard deviation. The *GM* and *GSD* of lognormally distributed data are related to the arithmetic moments by

$$GM = \frac{\mu^2}{\sqrt{\mu^2 + \sigma^2}} \tag{27}$$

$$GSD = \exp\left(\ln\left(1 + \frac{\sigma^2}{\mu^2} \right) \right)^{1/2} \tag{28}$$

As a result, uncertainty in the upper tail (95th percentile) of a lognormally distributed data set, provided with a 95% level of confidence, is also on the order of ±9% of the true value when based on a sample size of n=256.

Thus, the variance ratio illustrated in Figure 3 shows that the relationship between the two-sided 95% confidence limit of the 95th percentile of a sample with sizes n = 50, 100, 250, 500 and 1000 are within approximately 40%, 30%, 18%, 13% and 9% of the true variance, respectively. In terms of the standard deviation, the approximate precision of the estimate of the standard deviation from samples of size n = 50, 100, 250, 500 and 1000 are ± 20%, 14%, 9%, 6% and 4%, respectively. Again, these results are based on the assumption that the sample points were drawn from a normal distribution.

4.1.2 Interpreting Sample Size

The desired level of precision in the tails of the outcome distribution for a PRA will depend on both the given scenario and on the level of precision that the analyst can achieve for other inputs to the assessment. If the sensitivity analysis reveals that a distribution is potentially important and the sample size is small, (e.g., less than about 50) then the risk assessor should have justified the chosen distribution on a theoretical basis (Thompson, 1999) and taken steps to

incorporate uncertainty into the distribution using a two-dimensional (2-D) analysis (Bogen and Spear, 1987; Cullen and Frey, 1999; Morgan and Henrion, 1990). As was demonstrated in the previous section, one cannot expect to attain precision greater that about ±20% for the upper tail (~95th percentile) with samples where *n* is less than 50. Precision would be even less if the analyst was interested in even more extreme events such as the 99th or 99.9th percentiles. In contrast, a much smaller sample size could be used if the analysis is directed only at the 50th percentile or confidence about the mean.

Large samples can best provide an adequate level of information to identify an appropriate distribution when there is some theoretical justification available (Thompson, 1999). In these cases, uncertainty due to sample size will most likely be overshadowed by other sources of uncertainty in the analysis (Hertwich *et al*., 2000). We base our definition of a "large sample" and adequate precision on the premise that uncertainty and/or variability in the exposure concentrations in various media (e.g., air, soil, water, food, …) will often limit the precision of the assessment. Variation in measured or modeled concentrations in exposure media can easily exceed a factor of two or three. Errors due to the chemical analysis alone can be on the order of ±10% (Eiceman *et al*., 1993). Therefore, we suggest that sample sizes in excess of about 250 should provide adequate precision in the mid- to upper-tails (i.e., up to about the 95 percentile of the data). However, a final decision on the necessary level of precision for a given PRA will need to be judged on a case by case basis considering a range of site- and scenario-specific factors. Smaller sample size can be used if theoretical justification for a specific statistical model exists or if the analysis is focused on the central tendency of the outcome.

For distributions based on moderately small samples, e.g., *n* between 50 and 250, the decision to explicitly include uncertainty due to sample size will need to be judged on a case-by-case basis. For these situations, the risk assessor should either explicitly include uncertainty in the analysis or demonstrate that precision at the percentile of interest to the assessment is not significantly influenced by uncertainty due to sample size or uncertainty related to the choice of parametric model.

4.2 Data Quality Attribute

The remaining two attributes of the data (quality and relevance) are more subjective in nature. As a result only general guidance can be provided. Experimental error due to flaws in sampling design or imprecision in the measurement instrumentation are reflected in the quality of data and these errors propagate through to any distribution constructed from that data. Even with high quality data there is often a certain degree of imprecision. For example, analytical methods used to measure chemical residues in environmental samples often have relative standard deviations on the order of 5 - 10% (Eiceman *et al*., 1993) and this imprecision cannot be appreciably reduced by increasing the number of samples.

High quality data for developing distributions will typically be from direct measurements using peer reviewed and accepted methodology. The relative quality of information used in developing distributions for PRA is as follows: direct measurements of the exposure factor of interest > self reported values for the exposure factor of interest > direct measure of surrogate data > self-reported surrogate data. When the actual exposure factor is difficult or impractical to measure, surrogate information (e.g., estimated breathing rate from caloric intake or exposure duration based on population mobility) may be the best information available.

It is important to note that imprecision from lesser quality data is ultimately captured in the distribution as an increase in the range or spread of the input data. However, the reviewer must consider the possibility that lower quality data (self-reported or surrogate) may introduce bias into the data and ultimately into the distribution and this bias is often much more important than the inability to distinguish between uncertainty and variability.

Judging how the data quality attribute relates to the ultimate adequacy of the distribution is highly subjective. If the reviewer suspects that a distribution might be biased because of lower quality data then the results from the sensitivity analysis can provide insight into how important this potential bias might be in the context of the full PRA. If the potential for bias is high and the model is sensitive to the input then the risk assessor should address the issue by artificially inflating uncertainty in the distribution's central tendency (based on expert judgement) or by demonstrating that bias will have little effect on the regulatory decision. Knowing why and how the original data were collected will help the reviewer evaluate both the data quality attribute and the data relevance for a specified task.

4.3 Data Relevance Attribute

Even a significant amount of high quality data may not be adequate if the data is not relevant to the population of interest. Relevant data is data that adequately captures temporal and spatial variability as well as the demographic characteristics of the target population. For example, if a PRA is concerned about a long-term local pollutant source near an Inuit village in Alaska then exposure factor distributions from a probability based national survey, or based on short-term sampling periods (e.g., 2-day average intake rates) may not be relevant. Although direct measures within the population of interest are most relevant, it is often necessary to rely on information from surrogate regions or cohorts drawn from large and complex data sets (USEPA, 1997a) or on short-term data for estimating exposure factors for long-term exposure scenarios.

Judging the relevance of a distribution requires a clearly defined population or cohort, a well-defined exposure scenario, a well-documented source of data and a clear description of the steps, if any, taken by the risk assessor to adjust for data quality or to improve data relevance

(USEPA, 1999a). All of this information should be available in the PRA, particularly for those inputs that are identified as significant during the sensitivity analysis. If, as is most often the case, the data are not actually from the target population, then the PRA should clearly justify the distributions using either available data, expert judgment or a representative sample collected from the target population. If the risk assessor is unable to justify the relevance of an input then it may be necessary to use expert judgement to adjust the distribution towards a more representative value and/or inflate the uncertainty to capture potential differences between the target population and the sampled population. Although a number of methods have been suggested for improving the relevance of 2-day diary data for estimating long term intake rates (Buck *et al*., 1995; Slob, 1993; Slob, 1996; USEPA, 1999a; Wallace *et al*., 1994) little guidance is available for dealing with other types of relevance. Nevertheless, the reviewer need only consider whether any differences between the data used to construct a distribution and the target population of interest will influence the model outcome or the regulatory decision.

5.0 Category Three: Adequacy of the Distribution

Once the reviewer is satisfied with the data and any necessary adjustments, then the next step is to select a distribution that adequately represents this data and adequately incorporates uncertainties from deficiencies in data quantity, quality or relevance. The three attributes in the distribution category focus specifically on how well the final distribution simulates the particular exposure factor within the target population. The attributes include both statistical and graphical goodness-of-fit (GOF) tests along with consideration of cross-validation performance or verification/justification with alternate models and theory. All of this information should be readily available in the risk assessment or from the risk assessor. However, as with the data quality and data relevance attributes, judgment about the adequacy of distribution attributes will be subjective and often strongly dependent on the personal preference of the reviewer and risk assessor towards certain GOF or statistical tests and on their beliefs about the characteristics of the specific exposure factor.

Simple guidelines are provided along with references for further information, where necessary, to help improve the ease and consistency of the review process. Although much has been written on the appropriate methods for selecting, fitting and testing distributions (Conover, 1999; Cullen and Frey, 1999; D'Agostino and Stephens, 1986; USEPA, 1999a), it is not necessary for a reviewer to be proficient with probabilistic methods. Rather, the reviewer only needs to be concerned with judging how well the final input distribution simulates the exposure factor for the target population and the specified risk scenario.

5.1 Statistical Goodness-of-Fit Attribute

A number of GOF tests are available to the risk assessor for evaluating a hypotheses that data come from a particular distribution type for the population or, conversely, that a selected distribution is an adequate representation of the data (Conover, 1999; Cullen and Frey, 1999; D'Agostino and Stephens, 1986; McKone *et al*., 1996; Myers *et al*., 1998; Seiler and Alvarez, 1996; USEPA, 1998; USEPA, 1999a). Some of the more common statistical tests include the Chi-squared, Anderson-Darling, Kolmogorov-Smirnov and various measures of regression and correlation. Although some GOF tests penalize the model for over-parameterization (e.g., the F-test), most of the GOF measures are based solely on how well a particular parametric model fits an empirical distribution of the data.

Results from statistical GOF tests for each distribution, particularly those found to be important in the sensitivity analysis, should be provided in the risk assessment. Assuming the reviewer has already concluded that the data provide an adequate representation of the target population then statistical GOF can provide useful information for judging the adequacy of the distribution. However, a number of experts have pointed out that statistical GOF tests have significant shortcomings and should not be mistaken for purely objective measures of fit (USEPA, 1999a). Therefore, although positive results from the test(s) can increase the reviewer's confidence in the distribution, one should be careful not to base judgement about the adequacy of a distribution solely on the statistical GOF test.

Passing a statistical GOF test should not be taken as proof that a distribution is adequate, nor should failing the test be taken as a fatal flaw in the distribution. In each case (pass or fail), the reviewer should strive to understand why the results of the test were positive or negative and how that might influence the outcome of the PRA. Visual/graphical tests are often more powerful for judging the performance of a distribution – particularly when the assessment is targeting a specific region of the distribution. In short, if a distribution passes a statistical GOF test, the reviewer should take that into consideration but it is still important to consider graphical GOF. If the distribution fails the GOF test the reviewer should strive to understand why it failed (poor fit in a less important region of the distribution or excessively low tolerance for error due to large sample size) then move to the next level of the evaluation.

5.2 Graphical Goodness-of-Fit Attribute

Graphical methods of performance testing are often preferred when evaluating how well a distribution fits a set of data especially if a particular region of the distribution is of interest. Informative graphical methods for visualizing the performance of a distribution include frequency plots or histograms, residual plots, percentage-percentage probability (P-P) plots, quantile-quantile (Q-Q) plots or various forms of linear transformation. A simple layman's Q-Q

plot can be constructed without transforming the values to a linear form by simply drawing *n* "random" values from the selected input distribution, sorting these data by rank and plotting them against the original data (sample size = *n*) also sorted by rank.

Using the graphical GOF, the reviewer only needs to evaluate whether the distribution produces a sample that is compatible with the original data (relevance of the data to the target population is evaluated as an attribute of the data). If the simulated data is found to be biased, the reviewer must consider if that bias is in a critical region of the sample (e.g. the distribution consistently over-estimates values above the 95 percentile of exposure) and, if so, how this bias might influence the regulatory decision. As with statistical GOF, graphical GOF is particularly useful when the reviewer is satisfied with the data (quantity, quality and relevance). In this case, the risk assessor only needs to demonstrate that the distribution used in the assessment adequately simulates the original data.

If, however, the attributes of the data are questionable, then the risk assessor must demonstrate that the range of simulated data (including variability and uncertainty) captures the range of values that are likely to exist in the target population. Demonstrating that a distribution performs adequately in the absence of sufficient data ultimately must rely on professional judgement in the context of the regulatory management objective. If the PRA or the regulatory decision is highly sensitive to the input then it may be necessary to adjust the uncertainty component of the distribution to acknowledge shortcomings with the original data or the GOF.

5.3 Distribution Performance Attribute

This final attribute focuses on performance of the distribution as it relates to supplemental information that is specific to the risk assessment and/or to the target population. Although the performance attribute may draw on various GOF techniques, it is distinguished from the previous GOF attributes in that it goes beyond the original set of data and considers a range of evidence related to the performance of the model. Types of information that might be considered in the performance evaluation include the theoretical basis for a selected distribution, verification using alternate models or assumptions, validation using independent data, cross-validation using samples drawn from the original data and different sources of expert or professional advise. Although the overall reliability of a distribution cannot be fully characterized without testing the model against independent and highly relevant data, a resource that is rarely available, performance evaluations can provide significant insight into the reliability of the selected distribution function beyond the "calibration" data set.

Much of the information needed by the reviewer for the performance evaluation will have been used by the risk assessor in selecting a particular distribution and, as a result, should be available to the risk manager either through dialogue with the risk assessor or directly from the

PRA. However, some of the information and assumptions made during the distribution development process may be difficult to interpret. Although open dialogue is encouraged throughout the process of performing a PRA, evaluation of the performance attribute, more than any of the other attributes will require direct interaction and dialogue between the risk manager and the risk assessor. The dialogue should remain clear, focused and non-technical whenever possible. The reviewer should use all information available and strive to understand *how* the distribution might influence the regulatory decision, *why* the risk assessor selected the distribution and *what* evidence is available for supporting this selection.

If the evidence is sparse, it may be necessary to collect additional data or to subjectively increase the uncertainty for those distributions that are important to the analysis before the reviewer can accept them as adequate for the task. Again, it is important to seek advice from other experts or risk professionals who might be familiar with the specific risk scenario whenever the evidence about the adequacy of a distribution is inconclusive.

6.0 Process for Evaluating Input Distributions in Probabilistic Risk Assessment

In the process of combining the various attributes described above, we note that assigning a score to each attribute is subjective. At this point we see no opportunity to avoid subjectivity on the part of both the risk assessor and the risk manager. Even what seems to be quantitative information, such as statistical goodness-of-fit, can be strongly influenced by choices made during the distribution development process. The only attributes that can truly be considered objective are the quantity and quality of the data and possibly the sensitivity analysis, although the latter two are likely subject to interpretation. Thus, information with varying degrees of subjectivity will need to be synthesized during the evaluation process to maintain consistency and transparency in the review and to focus dialogue with the risk assessor and with other experts. To accomplish this, we provide a qualitative evaluation wheel for tracking the critical attributes of a distribution.

The evaluation wheel is based on the standard radar or radial plot available in most spreadsheet or graphing software. The tool is used here to track and organize the objective and subjective information that will contribute to the reviewer's judgement about the adequacy of each distribution in the PRA. The proposed wheel provides a format that is simple and flexible enough to allow the risk manager to adapt the tool in a way that is suitable to his or her own review style. Referring to Exhibit 1, each probabilistic input distribution in the PRA corresponds to a separate page and each page includes an evaluation wheel. Each attribute of the distribution corresponds to a different ray or spoke on the wheel. The individual rays on the plot are

arbitrarily scaled from either zero to 1 (sensitivity analysis) or low-to-high where "high" is preferable. The wheel is divided into the three major categories (i.e., those related to the SA, the Data and the Distribution). There is a space at the top of the page for recording details about the specific exposure factor, any correlation that might be included and summary statistics of the distribution. A set of checkboxes is also provided to indicate the status of the review. A section for notes is given at the bottom of the page to maintain transparency in the evaluation process. It is envisioned that the sheet containing the evaluation wheel would serve as a cover page summarizing the supporting information for each distribution. This supporting information can either be attached to the evaluation sheet or filed elsewhere depending on the reviewer's preference.

To use the evaluation wheel the reviewer first summarizes the descriptive information for each probabilistic exposure factor (name of exposure factor, PRA or scenario reference, related or correlated inputs, opening date of file, contact information, summary statistics, etc.). Next, each of the distributions in the PRA are scored based on results from the sensitivity analysis using a preliminary mark on the SA ray of the wheel. The reviewer then sorts the sheets in order of importance (based on the sensitivity analysis) and begins with the most important input and systematically works around the evaluation wheel considering, and qualitatively scoring, each of the attributes.

By documenting observations and supporting information for the distribution using the evaluation wheel and working closely with the risk assessor and other risk professionals, the risk manager should be able to arrive at a scientifically defensible and transparent conclusion regarding the adequacy of each exposure factor distribution. In spite of the subjective choices that must be made during the review process, the proposed evaluation procedure can provide consistency without a loss of flexibility. As a result, there is an opportunity to apply a single review framework for all EPA regions and yet provide the regional risk manager the flexibility for dealing with site- and case-specific issues in the PRA process.

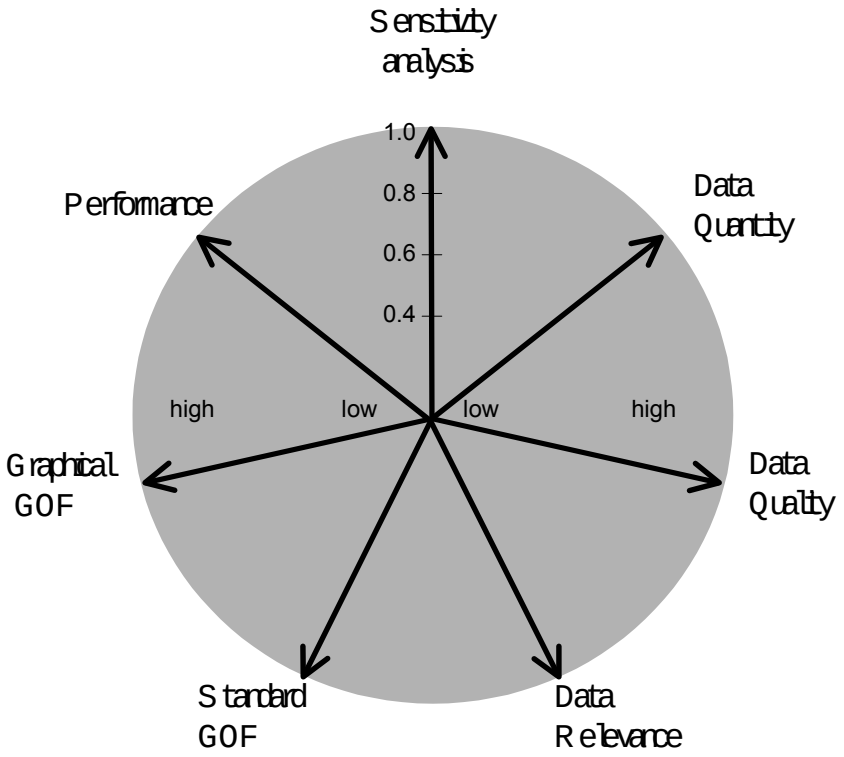
Exhibit 1: Evaluation Wheel and information tracking sheet

Exposure factor name: _____

Correlated with: _____

Summary statistics: _____

Adequate ☐
Inadequate ☐
Open ☐



Notes:

Part 2

On the Development of Recommended Exposure Factor Distributions

<Blank Page>

Part 2
On the Development of Recommended Distributions

As the number of inputs to a probabilistic risk assessment increases, so does the complexity of the process for calculating, communicating and managing risk. As a result, there is increasing effort required of both the risk professionals performing the analysis and the risk manager reviewing it. For deterministic risk assessments, the use of default inputs has improved the ease and the consistency of both performing and reviewing assessments. By analogy, it is expected that similar advantage will be seen in the field of probabilistic risk assessment through the introduction of default distributions. However, the idea of providing specific default distributions for exposure factors used in PRA is inconsistent with the purpose of performing the probabilistic analysis. Probabilistic risk assessments are performed to explicitly treat uncertainty and variability for a given risk scenario and a specific target population. This allows the risk manager to base decisions on both the estimated risk at a given percentile in the population and the confidence associated with that estimate. A single "health protective" default distribution would almost certainly lack relevance to specific target populations and specific Superfund risk scenarios.

Therefore, the primary challenge regarding development of default distributions for PRA will be that of balancing consistency with enough flexibility in the distribution (or the distribution development process) to achieve an adequate level of relevance to the target population. In confronting this challenge, we considered two initial options for developing scientifically defensible default distributions. The first option is to specify a baseline distribution or "prior" for the general U.S. population then provide a standard method for updating the central tendency and variance in the prior to better represent the target population using a Bayesian Monte Carlo approach. The second option is to provide, for a given exposure factor, a set of default distributions that are relevant to a key demographic category as a consistent starting point. The risk assessor then samples from these demographically based defaults to construct a task-specific distribution, with uncertainty and variability, based on readily available information about the size and demographic characteristics of the target population. Although certain aspects of the first option may be necessary, particularly for increasing relevance of short-term data, the recommendations developed in the following sections are based primarily on the second option. We made this choice because of the need for transparency in the distribution development process. Bayesian methods, although increasingly accepted by risk professionals, may be difficult to explain or justify to those stakeholders or risk managers who lack a strong background in Bayesian statistics and probability theory.

Introduction to Exposure Factors

We focus on three common exposure factors: body weight, exposure duration and water intake. This choice does not necessarily imply that these are the most important exposure distributions for risk assessment. We selected these factors to illustrate the process of developing default distributions for probabilistic risk assessment. Each of these exposure factors has available a significant amount of nationally representative data, but the relevance of the available data varies.

- Body weight (*BW*): The data for body weight includes both self-reported and directly measured values from several probability-based nationally representative surveys that include a wide range of demographic factors. As a result, body weight represents the "best case" for developing recommended distributions.

- Exposure Duration (*ED*): Although the various forms of exposure duration are rarely measured directly, several different surrogate measures have been used in the past to approximate "duration". We recognize that the definition of exposure duration depends on the specific risk scenario being considered and, as a result, must be defined on a case-by-case basis. The challenge for developing recommended distributions of exposure duration is that of relating the surrogate measures of duration to the population- or scenario-specific definition used in the risk analysis.

- Water Intake (IU_w): Water intake is estimated from self-reported values collected over two nonconsecutive days for both water and total food intake. The quantity of water derived from food is calculated using recipe data files. The water intake values are relevant, but uncertainty in the reporting process and the time frame in which the data was collected (short-term) make these data useful only as surrogate data for the time frame of interest (long-term average intake). As a result, the primary challenge for developing recommended water intake distributions will be transforming the data into a form that is relevant to long-term average intake across the population of interest.

Sensitivity Analysis of Exposure Factors in a Standard Product/Quotient Risk Model

Before looking at the individual exposure factors, it is helpful to understand the relative importance of each exposure factor in the simple product/quotient models typically used for Superfund risk assessment. Given an initial contaminant concentration in soil (time zero) at a given site, $C_s(0)$ (mass contaminant per mass soil), we can estimate risk, $H(\mathrm{t})$, by summing potential dose and effect over multiple exposure routes (j), environmental media (k) and exposure pathways (i) arriving at

$$H(t) = C_s(0)\left\{\sum_j \sum_k \sum_i \left[Q_{j\left(ADD_{ijk}\right)} \times \left(\frac{ADD_{ijk}}{C_k}\right) \times \Phi\left[C_s(0) \rightarrow C_k, t\right]\right]\right\} \tag{29}$$

where $Qj_{(\mathrm{ADDijk})}$ is the dose response function that relates potential dose (ADD_{ijk}), by route j to the probability (over a specified averaging time) of an adverse outcome for an individual in the population, $\left(ADD_{ijk} / C_k\right)$ is the average daily potential dose (over a specified averaging time) from exposure medium i by route j related to environmental compartment k and $\Phi\left[C_s(0) \rightarrow C_k, t\right]$ is the multimedia dispersion function that converts concentration in the source media (i.e., soil) at time zero, $C_s(0)$, into contaminant concentration in environmental media k at some time t in the future. Ignoring variance in the dose response function and assuming environmental media concentrations are constant over the duration of exposure (ED), the average potential dose for a specific demographic (age, sex, race, …) category (d) within the population is

$$ADD_{ijk}(d) = \frac{C_i}{C_k} \times \left\{\frac{IU_i}{BW}\right\}_d \times \frac{EF_d \times ED_d}{AT_d} \times C_k \tag{30}$$

where C_i/C_k is the ratio of contaminant concentration in exposure media i relative to the concentration in environmental media k, $\{IU_i/BW\}_{\mathrm{d}}$ is the intake/uptake rate of exposure media i relative to body weight for group d, and EF_d, ED_d and AT_d are the exposure frequency, exposure duration and averaging time, respectively, for group d. Averaging over all demographic categories in the population we get

$$ADD = \frac{1}{\left(\sum_{d=1}^{DG} N_d\right)} \sum_{d=1}^{DG} \left\{\frac{UI_i}{BW}\right\}_d \times \frac{C_i}{C_k} \times EF_d \times \frac{ED_d}{AT_d} \times C_k \times N_d \tag{31}$$

were N_d is the number of individuals in demographic group d, DG is the number of demographic categories used to represent the population. Combining all demographic categories and considering only a single exposure media, Eq. 31 reduces to the standard product/quotient exposure model

$$ADD_i = \frac{C_i \times IU_i \times ED \times EF}{BW \times AT} \tag{32}$$

Based on the exact analytical method for propagating variance in product/quotient models given earlier (Eq. 5, Section 3.1.2) and neglecting correlation among the inputs, we approximate the percent contribution to variance ($\%CTV_i$) for each input (n=6) in Eq. 32 as

$$\%CTV_{\mathrm{i}} = \frac{CV_i^2}{CV_w^2} = \frac{\dfrac{\sigma_i^2}{[\mathrm{E}(i)]^2}}{\left\{\sum_{j=1}^{n} \dfrac{\sigma_j^2}{[\mathrm{E}(j)]^2}\right\}} \times 100\% \tag{33}$$

where CV_i is the coefficient of variation for input i and model outcome w. As a first approximation, the CVs for intake (IU), body weight (BW), exposure duration (ED), exposure frequency (EF) and averaging time (AT) are 0.6, 0.2, 1.0, 0.14 and 0.1, respectively (McKone, 1992). Assuming an arbitrary yet somewhat optimistic value of 0.5 for the CV of the concentration term (C_i), the approximate contribution to variance for IU, BW, ED, EF, AT and C_i are 18%, 2%, 66%, 1%, 0% and 12%, respectively. The absolute values of $\%CTV$ will likely change depending on the risk model used, whether or not correlation is included and how much the input distributions are refined during the model development process. However, the relative ranking of the basic model inputs (i.e., $ED > IU$ & $C > BW, EF$ & AT) will likely be conserved unless uncertainty in the inputs is significantly changed and/or variance is partitioned into more homogeneous subgroups within the target population.

Having introduced the three case study exposure factors for the following chapters and evaluated their relative importance with respect to the outcome of the standard product/quotient model, we now move on to the work of developing recommended exposure factor distributions for PRA. The next section (Section 7) will consider body weight followed by exposure duration in Section 8 and finally, water intake in Section 9.

7.0 Body Weight

Although body weight can be important in deterministic risk assessments and in deriving unit risks, it is often one of the least important contributors to variance in the probabilistic risk assessment process. Nevertheless, a large amount of high quality data is readily available for a number of demographic categories within the U.S. population. As a result, body weight is useful for demonstrating the development of recommended PRA distributions. As stated above, we recognize that a single distribution or set of distributions will almost certainly lack the flexibility needed by the risk assessor. However, a reasonable amount of consistency and transparency is beneficial to both the risk assessors performing the analysis and the risk managers responsible for evaluating and using the information that comes out of the analysis. Therefore, our objective is to provide a consistent, scientifically defensible set of "baseline" distributions for the population, as a starting place for the distribution development process, along with a transparent and flexible method for using these distributions to derive a final exposure factor PDF for the

risk assessment. The final distribution should incorporate both uncertainty and variability and be relevant to the target population and specific risk scenario.

The general approach for satisfying our stated objective is to (i) acquire an adequate amount of raw data for body weight, (ii) use standard data analysis techniques to identify the most influential demographic factors and (iii) reduce the raw data to distributions of BW_d for each important demographic category. Finally, we provide a transparent method for combining the demographically based distributions to reconstruct relevant distributions that explicitly account for uncertainty and variability. This includes uncertainty due to differences in size and composition of the population potentially exposed at specific site relative to the populations represented by available distributions and the variability in the exposure factor across the population of interest.

7.1 Source of data

The data sources for body weight include the Center for Disease Control revised growth charts for children (Kuczmarski *et al*., 2000), the Continuing Survey of Food Intakes by Individuals (CSFII) (USDA, 1998) and the most recent National Health and Nutrition Examination Survey (NHANES III) (NCHS, 1996). In addition, the U.S. Census data is used as a source of site specific information for reconstructing distributions from the set of demographically specific distributions. Each information source is summarized below.

7.1.1 CDC Revised Growth Charts for Children from Birth to 20 years

The Center for Disease Control and Prevention / National Center for Health Statistics (Kuczmarski *et al*., 2000) recently released their revised growth charts for children from birth to age 20 years in the United States. Data for the revised charts were from five national health examination surveys collected from 1963 to 1994 and three supplemental data sets (used primarily to evaluate birth weight). Each survey reports measured body weight for individuals in a cross-sectional national probability-based sample of the civilian, non-institutionalized population. The data were pooled to obtain the necessary precision for calculating percentile distributions, and statistical sampling weights (provided for each survey) were applied to the data to make them representative of the U.S. population at the time the surveys were conducted.

Data from the most recent survey (NHANES III) for children ≥ 6 years were excluded from the combined data set to avoid the upward shift in the body weight of children in recent years. Exclusion of recent data for children ≥ 6 years may reduce the relevance of the data set for the current population. However, we agree with, and accept the decision made by the CDC to exclude the recent data for two reasons. First, the current trend toward increased body weight for children in the U.S. population is generally thought to be unhealthy. Efforts are underway to

educate the population and reverse this trend and as such the baseline set of demographically based distributions should be relevant to a "healthy" population. Second, given the function of the *BW* factor in risk models (e.g. decreased *BW* increases "risk"), using body weight data that is biased towards a "healthy" population will error slightly towards increased protection of the overweight population.

Although there appear to be racial differences in growth rates for children, these differences are small and inconsistent. Therefore, only sex/age categories are provided in the CDC growth chart data (Kuczmarski *et al*., 2000). Data were grouped by month from 1 to 11 months, by 3-month intervals from 12 through 23 months and by 6-month intervals from 24 months through 19 years. To reduce the "jagged or irregular" nature of the plotted percentiles, a number of data smoothing procedures were used to develop the final weight-for-age percentile curves for males and females from birth to 20 years. In addition to the empirically smoothed percentile curves, a statistical smoothing procedure was used to derive analytical expressions that allow interpolation between the reported percentiles. The statistical smoothing approach uses a Box-Cox transformation in the form of

$$centile = M\left(1 + LSZ\right)^{1/L} \tag{34}$$

where M and S are the median and standard deviation, L is the power in the Box-Cox transformation and Z is the z-score that corresponds to the percentile of interest. Values for L, M and S were actually derived empirically in the report to provide final curves that are nearly identical to the initially smoothed percentiles. The final values are reported by sex for one-month intervals from birth to 20 years (http://www.cdc.gov/growthcharts).

For convenience, the values of the parameters in equation 34 are provided in Appendix 1. For further details on the data and the empirical and statistical smoothing methods, the reader is referred to the original report on the revised growth charts (Kuczmarski *et al*., 2000). We use these growth charts exclusively for deriving age and sex specific distributions for children and adolescents from birth through 19 years.

7.1.2 Continuing Survey of Food Intakes by Individuals, CSFII

The CSFII was conducted from 1994 through 1996 by the Food Surveys Research Group, Beltsville Human Nutrition Research Center, Agricultural Research Service (USDA, 1998). The survey collected self-reported values of body weight for individuals in a stratified multi-area probability sample of non-institutionalized (non-military, and not in group quarters) households, for the U.S. population. Low-income housing units were over-sampled but weighting factors were provided for the samples so that the final values reflect the composition of the US population based on the 1990 Census. The actual question asked of the interviewee was "About

how much (do you/does NAME) weigh without shoes?" – reported in pounds. The response was converted to kg and listed for each sample person in data file RT25.

A number of other factors were included in the survey. The demographic factors that might contribute to variance in body weight include age, pregnant/lactating status, race (white, black, Asian/Pacific Islander, American Indian/Alaskan native, or other race), origin (Mexican, Puerto Rican, Cuban, or other Hispanic) and region (Northeast, Midwest, South or West). After excluding children and adolescents (birth through 19 years) and excluding sample persons with fields for body weight that included "don't know" or "non-reported", the remaining 9,161 adults (≥ 20 years) included 4,724 men and 4,437 women.

The raw data from the CSFII for all ages were used in a previous report (Maddalena *et al*., 1999) to evaluate and identify the important demographic factors influencing body weight. The raw data (appropriately weighted) for adults was then pooled with the data from the National Health and Nutrition Examination Survey (NHANES III) and the combined data set was used to develop age/sex based distributions for adults.

7.1.3 National Health and Nutrition Examination Survey, NHANES III

The Third National Health and Nutrition Examination Survey was conducted by the National Center for Health Statistics between 1988 and 1994. The general structure of the survey was a stratified multistage probability design representing the total civilian non-institutionalized population, 2 months of age and over, in the 50 States of the U.S. (NCHS, 1994). The study was multifaceted but the overall goal was to evaluate prevalence and trends in selected diseases and risk factors and to estimate population reference distributions of selected health parameters. The study was also designed to investigate disease etiology and the natural history of selected diseases. These goals required precise information on the health status of selected population groups and as a result required the over-sampling of several groups (e.g., young children, the elderly and certain ethnic or racial groups). Sample weighting factors are provided to allow estimates of values relevant to the total U.S. population based on the 1990 census. The body weight values reported in the NHANES III results represent actual measurements taken during a physical examination.

As with the CSFII data, a number of demographic factors are reported for each sample person. After removal of sample persons reporting age<20 years and non-reported body weight values, the 16,977 adults (≥ 20 years) included 7,934 men and 9,043 women.

7.1.3 The Combined NHANES III, CSFII data set

To achieve an adequate sample size for the various demographic categories for adults, we combine the NHANES III and CSFII data. However, prior to combining the samples, the degree

of bias in self-reported values was evaluated by comparing the empirical cumulative distributions (weighted) for adult males and females from each survey (Figure 4). The actual measured body weights from the NHANES III survey are taken as accurate and the deviation from the NHANES III distribution function by the CSFII distribution function is assumed to be due to bias in the reporting/data collection method.

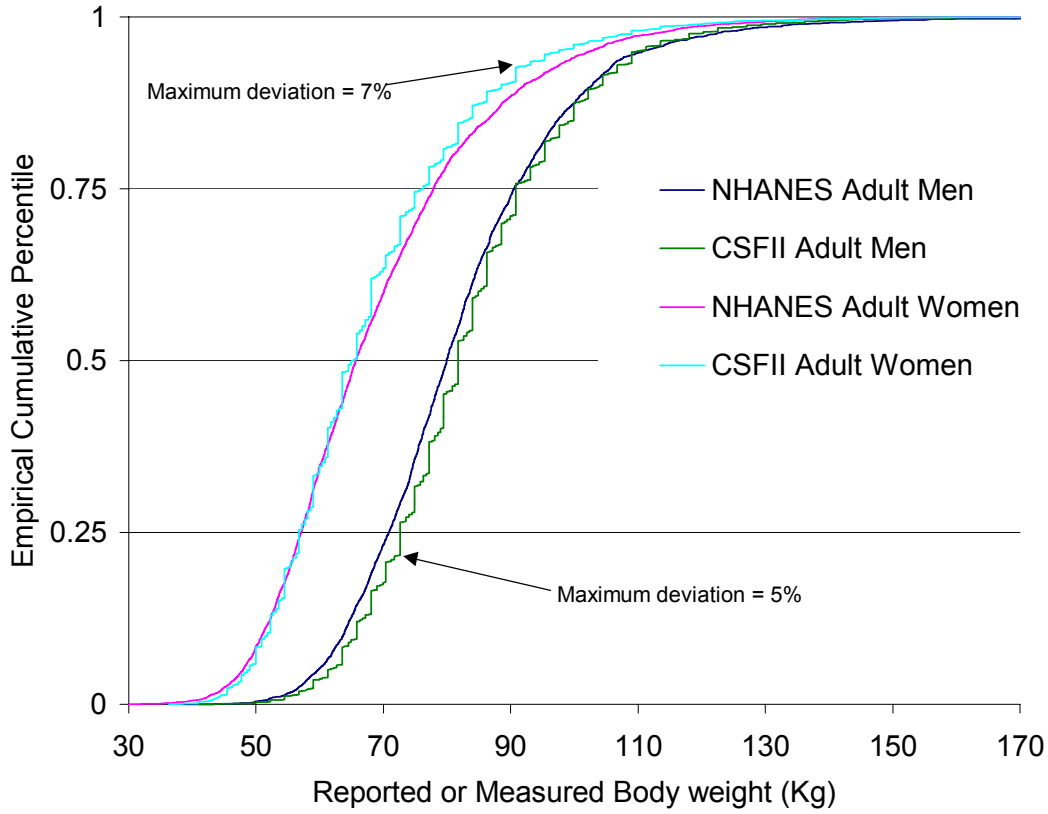


Figure 4 Comparison of empirical cumulative distribution functions from two national surveys where the measured values from NHANES III are shown as smooth curves (women on left, men on right) and the self-reported values from CSFII are the irregular or jagged curves. Figure 4 also illustrates the difference in body weight related to *sex*.

The maximum deviation between the self-reported and measured body weights for women (~7%) occurs at about the 93rd percentile of the data and the maximum deviation for men (~5%) occurs at about the 22nd percentile of the data. In general, we find that women do a good job reporting body weight in the lower half of the distribution but under-report body weight in the upper half. Men do a reasonable job reporting body weight in the upper half of the distribution but tend to over-report body weight in the lower half. However, given the relative

sample size of each data set, the measured data from NHANES III is automatically weighted over the self-reported data from CSFII by a factor of almost 2:1 so the overall effect of bias in the self-reported values is expected to be negligible. Therefore, we combine the results from the two surveys for use in characterizing demographically based (e.g., age and sex) distributions in section 7.3.

7.2 Key demographic factors

Based on results presented in a previous report (Maddalena *et al*., 1999), body weight is primarily dependent upon age, followed by sex and race. Although age dependence seems to continue throughout ones lifetime (Okajima *et al*., 2000), the strongest dependence is seen from birth into the late teens. The dependence of body weight on sex and race is mainly seen in adults. The influence of race was detected in both male and female adults. Even though a relatively sophisticated data analysis package was used previously to identify and evaluate the important demographic factors for body weight (Maddalena *et al*., 1999), the objective of this work is somewhat less stringent in that we are only interested in identifying the main factors influencing body weight. As such, any number of data analysis techniques would suffice. The simplest way to satisfy our current objective was to confirm the results reported earlier by visually comparing the distributions from the demographic categories that were identified as important in the previous work (i.e., age, sex and race).

The influence of age on body weight is obvious even in the absence of graphical analysis. The importance of sex on the distribution of body weight is illustrated in Figure 4, which shows the difference between empirical cumulative distributions for adult males and females. Likewise, the influence of *race* is illustrated in Figure 5 as a comparison of empirical distribution of body weight for adult women. The average reported body weight for adult women classified as *asian/pacific islander, white/other* and *black/native american* were 55 kg, 68 kg and 77 kg, respectively, which, along with the distributions, clearly shows the potential influence of race. The results for adult men (not shown) were similar to those found for adult women in that the average body weight reported by *asian/pacific islanders* was approximately 16 kg less than the average reported for the remainder of the adult male population.

Although race was found to have a measurable influence on body weight (Figure 5) (Maddalena *et al*., 1999), only the CSFII survey collected information specific to “asian, pacific islander” and the data set was not large enough to allow development of separate distributions for each age/sex/race category. Therefore, the demographic categories selected in this study, which provide the basis for deriving task-specific distributions, are constrained to age and sex. If the target population is composed predominantly of individuals of Asian and/or Pacific Islander decent, and the assessment is concerned with adults, then the risk assessor may need to adjust the

central tendency of the final distribution or re-evaluate the raw data to develop relevant distributions for the specific task.

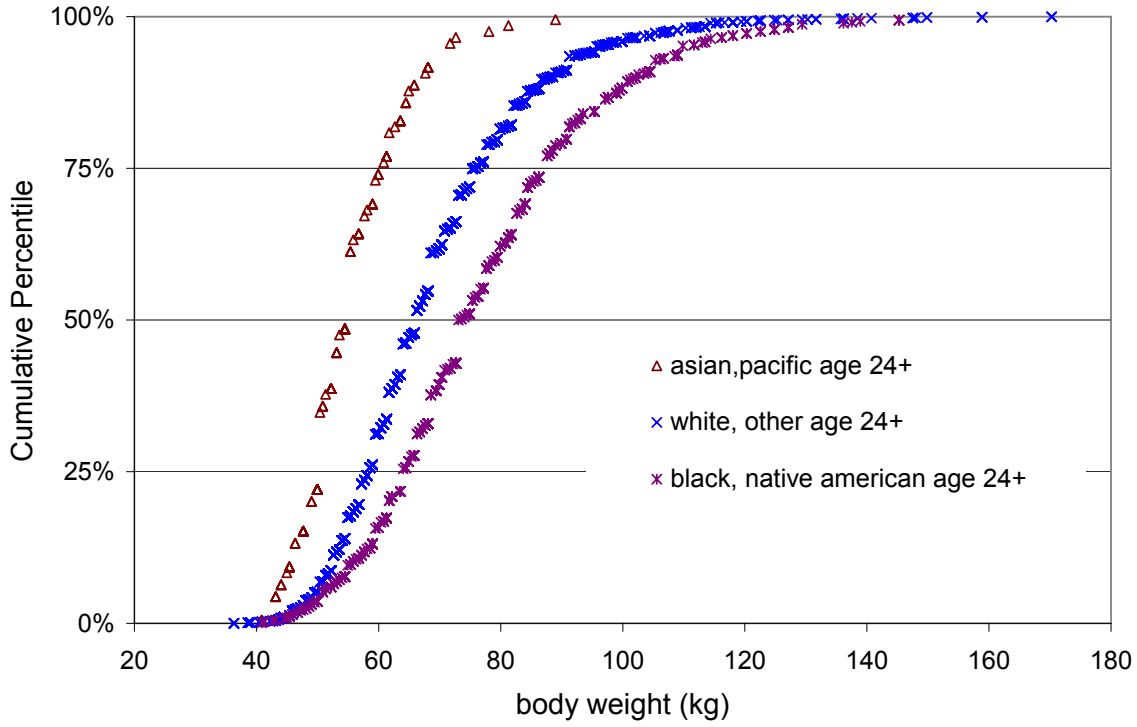


Figure 5: An illustration of difference between the empirical cumulative distribution of reported body weight from CSFII for females identified with different racial groups. Similar racial differences in reported body weight were seen for adult men (Maddalena *et al*., 1999).

7.2.1 U.S. Census Data Summary Tables

Summary tables of the 1990 U.S. census data provide information on the demographic composition of site specific populations at different geographically relevant scales. The information is collected every 10 years although the 2000 data was not yet generally accessible when this report was finalized. Information from the 1990 census is easily accessed from the U.S. Census Bureau's LOOKUP Servers (interactive data retrieval via HTTP) at http://homer.ssd.census.gov/doc/lookup_doc.html. Summary tables are provided at a number of geographic scales including State, County, Place, Urbanized Area and Metropolitan Statistical

Area. These options provide the flexibility to acquire population information on a scale that is relevant to a particular risk scenario.

The summary tables provide 100% counts for a specified location. The results are reported by sex in 14 age categories from birth through 19 years along with 17 additional age categories for adults (age ≥ 20). These categories form the basis for development of demographically based distributions (e.g., subgroups of the population are defined to match the age/sex categories reported by the Census Bureau).

7.3 Constructing Age/Sex Baseline Distributions

In this section, we use the information described above to develop distributions that are relevant to a specified set of age/sex categories. The set of categories is selected to coincide with the available information from the U.S. Census summary tables. These tables include by sex 14 age categories for birth through 19 years and 17 additional age categories for adults (age ≥ 20).

For children from birth through 19 years, we begin with the statistically smoothed growth curves from the CDC (Kuczmarski *et al*., 2000) and transform them into probability distributions that match the age/sex categories reported in the census summary tables. For adults, we begin with raw data from the combined CSFII, NHANES III. The data is grouped in categories that match the census tables then specific percentiles are estimated from the raw data. These percentiles are smoothed and individual parametric distributions are fit to each age/sex category. The approach is similar in principle to the methods used by the CDC in developing percentile growth curves for children, (Kuczmarski *et al*., 2000) except that a parametric model is fit to the percentiles rather than a Box-Cox transformation. The approach is also similar to recent work characterizing distributions of body weight using data from NHANES II (Burmaster and Crouch, 1997).

7.3.1 Baseline distribution for children (birth through 19 years)

Distributions of body weight by sex are developed for the following age categories: birth to 1, 1 through 2, 3 through 4, 5, 6, 7 through 9, 10 through 11, 12 through 13 and each year from 14 through 19 years of age. Equation 34 and the parameters listed in Appendix 1 were used to simulate 500 equally weighted values for body weight for each 1-month age interval using the following combined Excel – Crystal Ball function

$$percentile = M*(1+L*S*\text{NORMSINV}(\text{CB.Uniform}(0,1)))\text{^}(1/L) \qquad (35)$$

where "CB.Uniform(0,1)" is used to generate random values from zero to unity, "NORMSINV(CB.Uniform(0,1))" transforms the random value into a "z-score", and *L, M* and *S* are as previously defined (see equation 34). Simulated data for the 1-month intervals were then

combined according to the desired age/sex categories. For example, the 500 random body weights for each 1-month interval included in the age category defined as one through two year olds (total of 24 1-month intervals) resulted in a total of 12,000 random body weights for the category.

Table 2: Best fit parameters for each age/sex category for children to be used with Eq. 37

	Males				Females			
Age category	θ	$\mu_{LN(x)}$	$\sigma_{LN(x)}$	SS_{xy}	θ	$\mu_{LN(x)}$	$\sigma_{LN(x)}$	SS_{xy}
1 and 2	-1.799	2.664	0.124	5.42E-05	0.115	2.470	0.154	4.87E-05
3 and 4	6.576	2.267	0.239	5.55E-05	7.829	2.072	0.306	8.65E-05
5	9.092	2.343	0.258	8.52E-05	10.599	2.139	0.331	5.20E-04
6	10.674	2.420	0.288	6.64E-05	11.705	2.277	0.348	1.53E-04
7 through 9	13.512	2.610	0.379	7.38E-05	12.065	2.726	0.379	1.96E-05
10 and 11	16.391	2.972	0.373	8.26E-05	15.138	3.090	0.365	1.39E-04
12 and 13	14.605	3.437	0.307	7.71E-05	19.456	3.274	0.350	1.31E-04
14	24.018	3.392	0.345	1.01E-04	26.961	3.178	0.398	2.33E-04
15	20.764	3.640	0.281	1.35E-04	31.306	3.083	0.437	1.76E-04
16	30.238	3.482	0.331	1.37E-04	33.681	3.038	0.436	3.41E-04
17	33.911	3.462	0.344	8.25E-05	37.051	2.908	0.499	7.01E-04
18	36.659	3.450	0.356	6.40E-05	36.574	3.006	0.461	4.48E-04
19	34.431	3.573	0.317	1.16E-04	35.318	3.119	0.437	3.02E-04

The simulated body weights were then used to derive a baseline distribution for each age/sex category. Exploratory data analysis using the curve fitting functionality in the Crystal Ball software failed to identify a parametric model that adequately fit the data for each category. Upon further evaluation, we determined that the 3-parameter lognormal model provided good agreement across all age categories. The percentiles of the 3-parameter lognormal are estimated using the following Excel function

$$percentile = \text{LOGNORMDIST}((x_i - \theta), \mu_{LN(x)}, \sigma_{LN(x)}) \quad (36)$$

where θ is the location parameter, $\mu_{LN(x)}$ is the mean of $\ln(x_i)$ and $\sigma_{LN(x)}$ is the standard deviation of $\ln(x_i)$. The 3-parameter lognormal was fit to the data by minimizing the sum of the square differences (SS_{xy}) between the empirical percentile of the data and the percentile estimated using Eq 36. The "fit" was performed on a total of 21 different points include the 1st, 5th, 10th, 15th, …, 85th, 90th, 95th and 99th percentiles of the data. The resulting parameters are reported in Table 2 along with the sum of the squares for the fit.

The parameters listed in Table 2 can be used to generate random body weight values for each demographic (age/sex) category using the following Excel function

$$x_i = (\text{LOGINV}(\text{CB.Uniform}(0,1), \mu_{LN(x)}, \sigma_{LN(x)})) + \theta \quad (37)$$

where the "LOGINV" function transforms the random value generated by "CB.Uniform(0,1)" into a lognormal distribution with the defined parameters. Thus, the 3-parameter model can be readily adapted for use in available probabilistic risk software.

7.3.2 Baseline distributions for adults (≥ 20 years)

Distributions of body weight for adult males and females (age ≥ 20 years) are developed for the following age categories: 20, 21, 22 through 24, 5 year intervals from 25 through 59, 60 through 61, 62 through 64, 5 year intervals from 65 through 84 and 85 and above. This provides a total of 17 categories by sex in accordance with the categories presented in the census summary tables. Values reported in the combined CSFII, NHANES III surveys were first grouped according to the above age categories. Next, the percentiles (weighted) for each age/sex category were calculated. The raw percentiles for males and females are reported in Appendix 2. Prior to constructing distributions, the raw percentiles were empirically smoothed using

$$y = a + bx^3 + \frac{c}{\sqrt{x}} \tag{38}$$

and

$$y = a + bx^2 + cx^2\sqrt{x} \tag{39}$$

for male percentiles and female percentiles, respectively. The smoothed percentiles for each age category are also provided in Appendix 2. The smoothed curves for the 5th, 50th and 95th percentiles of body weight for adult males are plotted along with the raw percentiles in Figure 6. For reference, Figure 6 also includes weight versus age curves for children (birth through 19 years) based on the CDC *LMS* model described earlier. The slight discontinuity at the transition from the growth chart curves to the adult curves, particularly for the 95th percentile curve, is likely due to the exclusion of the NHANES III data from the growth charts (e.g., removal of the effect of overweight children in recent surveys).

The cross-sectional data illustrated in Figure 6 shows a slight temporal trend towards increasing body weight up until about 50 years of age followed by a decreasing trend. This trend is similar to one seen in data from the NHANES II survey (Burmaster and Crouch, 1997). Although the data is extremely limited in scope, there is recent evidence that the longitudinal long-term temporal variation in individual body weights follow the general pattern found in the cross-sectional data (Okajima *et al*., 2000). Although the trend is subtle, it should be considered to maximize the relevance of the distribution to the age composition of the target population. Therefore, the next step is to develop distributions for each of the age/sex categories for adults.

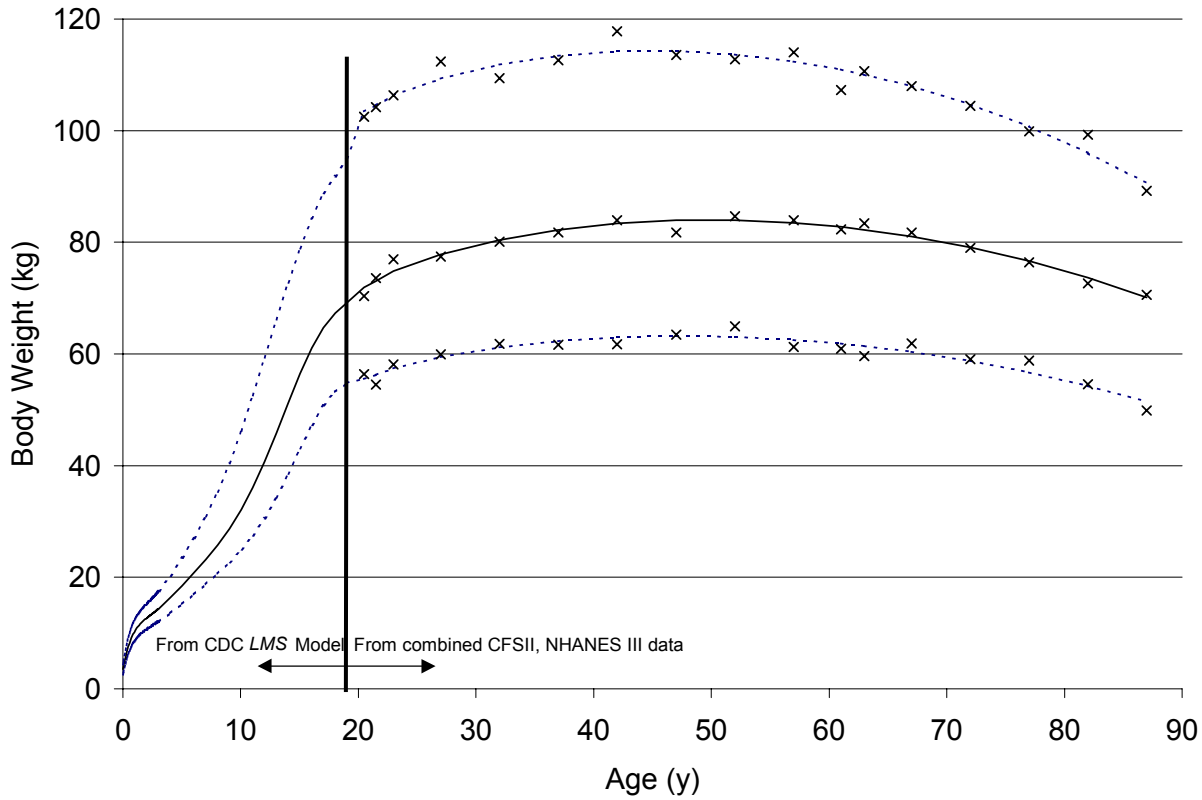


Figure 6: Illustration of the 5th (lower dash), 50th (solid line) and 95th (upper dash) percentile of body weight for adult males as a function of age from birth to 85 years and above. The curves to the left of the solid line are from the CDC *LMS* model and the curves to the right of the solid line are the smoothed percentiles from the combined CSFII, NHANES III data. The data marks represent the empirical percentiles for each age category.

As with the children's data, an exploratory data analysis using the curve fitting functionality in the Crystal Ball software failed to identify a parametric model that performed well across all categories. Again, upon further analysis, the 3-parameter lognormal had the flexibility to perform well with each age/sex category. Therefore, the same method previously described (Section 7.3.1) was used here to fit the 3-parameter lognormal model to each age/sex category. The results are presented in Table 3 along with the sum of the square differences (SS_{xy}) between the smoothed percentile of the data and the percentile estimated using Eq 36.

Table 3: Best fit parameters for each age/sex category for adults based on CSFII/NHANES III

Age category	Males				Females			
	θ	$\mu_{LN(x)}$	$\sigma_{LN(x)}$	SS_{xy}	θ	$\mu_{LN(x)}$	$\sigma_{LN(x)}$	SS_{xy}
20	34.826	3.618	0.349	1.28E-03	38.624	3.037	0.607	7.73E-04
21	34.312	3.659	0.339	1.09E-03	38.595	3.057	0.601	7.20E-04
22 – 24	33.291	3.731	0.321	8.43E-04	38.512	3.096	0.587	6.22E-04
25 – 29	31.490	3.839	0.297	6.46E-04	38.217	3.176	0.560	4.62E-04
30 – 34	28.158	3.961	0.270	6.48E-04	37.569	3.278	0.526	3.24E-04
35 – 39	25.200	4.047	0.251	7.30E-04	36.545	3.378	0.492	2.38E-04
40 – 44	21.761	4.124	0.234	8.03E-04	35.089	3.475	0.458	1.89E-04
45 – 49	17.817	4.194	0.219	8.30E-04	33.097	3.570	0.425	1.62E-04
50 – 54	13.245	4.261	0.204	8.01E-04	30.533	3.660	0.392	1.46E-04
55 – 59	7.013	4.338	0.187	7.18E-04	27.184	3.750	0.358	1.33E-04
60 – 61	3.297	4.376	0.178	6.28E-04	23.855	3.822	0.330	1.23E-04
62 – 64	0.003	4.410	0.171	5.79E-04	21.834	3.861	0.315	1.18E-04
65 – 69	-4.863	4.453	0.161	4.99E-04	17.321	3.937	0.286	1.12E-04
70 – 74	-6.477	4.448	0.157	5.35E-04	9.679	4.048	0.245	1.24E-04
75 – 79	-28.000	4.649	0.124	6.48E-04	0.558	4.158	0.206	2.07E-04
80 – 84	-46.770	4.789	0.103	1.27E-03	-22.435	4.428	0.144	4.18E-04
85 – above	-61.807	4.879	0.089	2.88E-03	-49.448	4.668	0.099	1.18E-03

7.4 Relating the Demographically Based Distributions to Site-specific Distributions

Information about the demographic composition of the target population can be accessed through the U.S. Census interactive data retrieval service located at http://homer.ssd.census.gov/doc/lookup_doc.html or by direct survey of the target population. This census information is used to design a sampling strategy for reconstructing the distribution of body weights for the target population using the demographically based distributions developed in the previous sections.

The recommended approach begins with the sample numbers taken from the census data for the specific age/sex categories. These sample numbers may need to be updated to account for changes in the population that may have occurred since the last census was taken. Census data for 2000 was not generally available for this study, but population projections by state are available from the U.S. Census Bureau. These projections were used to improve the relevance of the 1990 data with respect to the current population. In general, the population/demographic numbers should be selected to maximize relevance of the data to the target population in the region of influence around a particular site. For a discussion of how to approximate the region of influence or the distance that a chemical might travel in the environment, see Bennett et al. (1998).

Once the representative number of individuals has been determined for the target population, the appropriate number of random values is drawn from each age/sex distributions to create a single realization of the distribution of the exposure factor of interest for the specified

population (or cohort). These samples for the demographic categories are combined to address the specific needs of the PRA (e.g., children ages 7-13 or adult women). Finally, an appropriate parametric model is identified and fit to the data (see Part 1 of this report for discussion and references). For the body weight parameter, typically a 2-parameter or 3-parameter lognormal will be appropriate for the data. When the target population in the region of influence is relatively small, the sampling and fitting process should be repeated several times to evaluate uncertainty about the choice of parametric model and to estimate the level of uncertainty in the model parameter values.

Given the tabulated parameters for the age/sex specific distributions provided in Tables 2 and 3, the development of task-specific default distributions can be automated in a highly transparent manner using any number of spreadsheet, mathematical or statistical software packages. These baseline distributions and the method for reconstructing body weight PDFs provide a first example of how risk professionals might be able to improve the consistency of the probabilistic risk assessments without sacrificing the flexibility that is necessary for relating the process to specific target populations and risk scenarios.

8.0 Exposure Duration

Exposure duration (ED) can be an influential input to the risk assessment process. However, depending on the risk scenario, exposure duration can be defined in a number of different ways. Some of the more common definitions include time spent in a given occupation; at a particular location, (e.g., time spent at current residence) (USEPA, 1997a); or participating in a particular activity (e.g., duration of time as a sport fisherman (Price *et al*., 1998)). In general, exposure duration is the time interval during which exposure occurs, either continuously or intermittently, at a given exposure concentration and intake/uptake rate (USEPA, 1997a, page 1-12). Based on this general definition and for the purposes of this report, we treat exposure duration as the amount of time spent in an "exposure district," the spatial range potentially impacted by a particular risk agent (i.e., a contaminant at a Superfund site). Given this definition, a useful first approximation for exposure duration might be the amount of time an individual remains in his/her current residence – a value previously referred to as "total residence time" (Israeli and Nelson, 1992), "residential occupancy period" (Johnson and Capel, 1992) and "residence duration" (Sedman *et al*., 1998).

The problem with simply using occupancy period as a surrogate for exposure duration is that changing residences does not necessarily mean that the exposed individual has moved out of an exposure district. Statistics from the National Association of Realtors, reported in the Exposure Factors Handbook (USEPA, 1997a, Table 15-171) indicate that approximately half of

home buyers purchase homes within 10 miles of their previous residence. Therefore, to estimate exposure duration one should know both the total occupancy period for individuals in the target population and the likelihood that at the end of each occupancy period the individual moves out harms way with respect to a given risk scenario.

Neither the total occupancy period nor absolute distance to a new residence has been reported on a scale that is relevant to the national population. Rather, surrogate data such as mobility, mortality, current residence time and/or tax records have been used to estimate occupancy period for various demographic categories in the U.S. population (Israeli and Nelson, 1992; Johnson and Capel, 1992; Price *et al*., 1992; Sedman *et al*., 1998). Information on distance of move is even more sparse. Currently, the best readily available source of information related to distance of move for the national population is from migration data that reports type of move at the scale of state, county, metropolitan area, central city or suburb. Although nationally representative, the migration data lack site-specific information. We expect this lack of site specific data to increase the uncertainty in distributions of exposure duration. Nevertheless, without information about both mobility and distance of move, any estimate of exposure duration would lack relevance to the exposed population.

Therefore, our objective for exposure duration is to provide a consistent, scientifically defensible "baseline" distribution for the population along with a transparent and flexible method for using this information to derive a final exposure factor PDF. The final distribution should incorporate both uncertainty and variability and be relevant to the target population and specific risk scenario.

Our general approach is to (1) identify and acquire the information to estimate occupancy period, (2) use data analysis techniques to identify important demographic factors and (3) review existing methods for estimating occupancy period in the context of uncertainty about site-specific demographics and distance of move. A relatively simple and transparent method is needed for reconstructing relevant distributions that explicitly account for uncertainty and variability due to the size and demographic composition of the population in the target region and the general lack of information related to distance of move.

8.1 Sources of data

The U.S. Bureau of the Census currently conducts national housing surveys every other year. These surveys provide comprehensive housing statistics for the U.S. Department of Housing and Urban Development (HUD) and include information on housing (apartments, single-family homes, and mobile homes, tenure), attributes of housing units (locale, number of rooms, square footage, etc.), and data on household members (age, race, sex, income, education, etc.). The current residence time and demographic information reported in the survey data from

1995 (USDOC, 1995) were used in a previous report (Maddalena *et al*., 1999) to evaluate and identify key demographic factors in the population that influence residence time.

The U.S. Bureau of Census also provides information on population mobility that is collected using the Annual Demographic Supplement to the Current Population Survey (USDOC, 1999). Mobility status is classified on the basis of a comparison between the place of residence of each individual at the time of the March survey and the place of residence the previous March. The information about whether the sample person is living in the same house at the end of the survey period as the beginning is used to identify movers and non-movers. Movers are further classified as to whether they were living in the same or different county, state, region, or were movers from abroad. Movers are also categorized by whether they moved within or between central cities, suburbs, and non-metropolitan areas of the United States (USDOC, 2000). Outside of a direct sample of the target population, the Current Population Survey provides the most relevant information available on a national scale for estimating the annual probability and approximate distance of a move.

8.2 Identification of key demographic factors

Current residence time (CRT) (USDOC, 1995) was used as a surrogate for exposure duration to identify important demographic factors for occupancy period (Maddalena *et al*., 1999). The results indicate that exposure duration is primarily dependent upon age and tenure (rent or own). Farm status also seems to influence the reported length of current residence, since farm families tend to remain in their current residence longer than the general population. However, some of this difference is explained by the tenure and urban/rural status that is typically associated with farms. There also seems to be a slight regional influence. People residing in the Northeast region of the country report somewhat longer residence times than the general population. These results were similar to those found in an earlier report that used current residence time to estimate total residence time (Israeli and Nelson, 1992).

The age dependence of reported current residence time, as illustrated in Figure 7, shows a bimodal relationship that peaks at around age 18 then drops to a minimum at around age 30 followed by a steady increase to a maximum current residence time of about 27 years. If we assume that the current residence time reported for children (birth to around 20 years) is approximately that of their parents, (i.e., young adults) then we can ignore reported values for children in Figure 7 and consider only adults. Without the influence of reported residence time for children, the relationship between residence time and age is approximately linear from age 24, increasing at a rate of approximately 0.46 per year of age.

In contrast to data on time at current residence, migration data indicate that the median occupancy period remains relatively constant across all age categories in owner occupied homes.

The trend towards increased residence time for adults is mainly due to renters where occupancy period doubles between ages 24-29 and ages 65 and above (Hansen, 1998). Furthermore, among the general population, renters are approximately 4 times more likely to move in a given year than owners are. Within the general population, those that do move are approximately 3 times more likely to remain in the same county than they are to move to a different county (Figure 8). It is not possible to know, *a priori*, the distribution of ages within a given household. Therefore, we conclude that the most important information that is readily available for understanding exposure duration are the composition of the housing stock in the target region (i.e., rental versus owner occupied) and the type of move (local versus long distance).

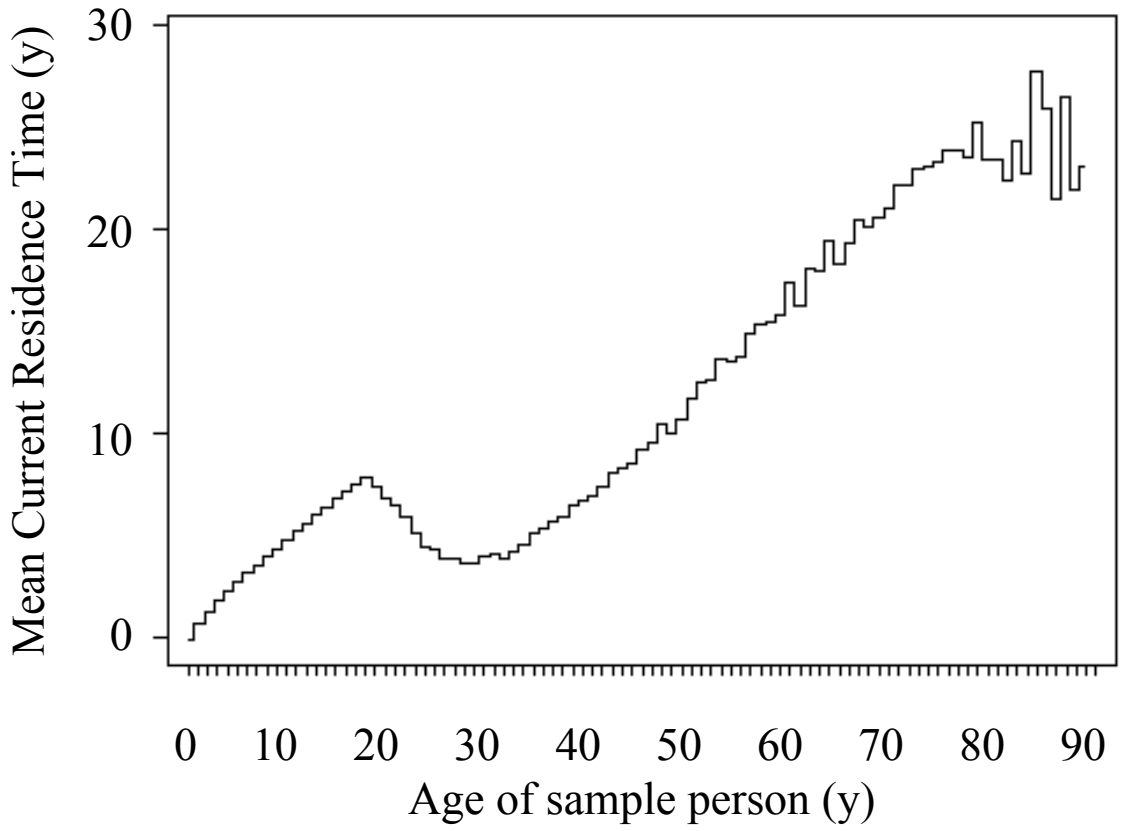


Figure 7: The relationship between the average reported current residence time for the national population and the age of the sample person showing a bimodal distribution that arises from correlation between residence time of children and their parents.

Based on the availability and quality of information related to mobility, housing stock composition and migration, we can recommend a balanced approach for estimating exposure duration but first we provide a brief overview of previously published methods as a background.

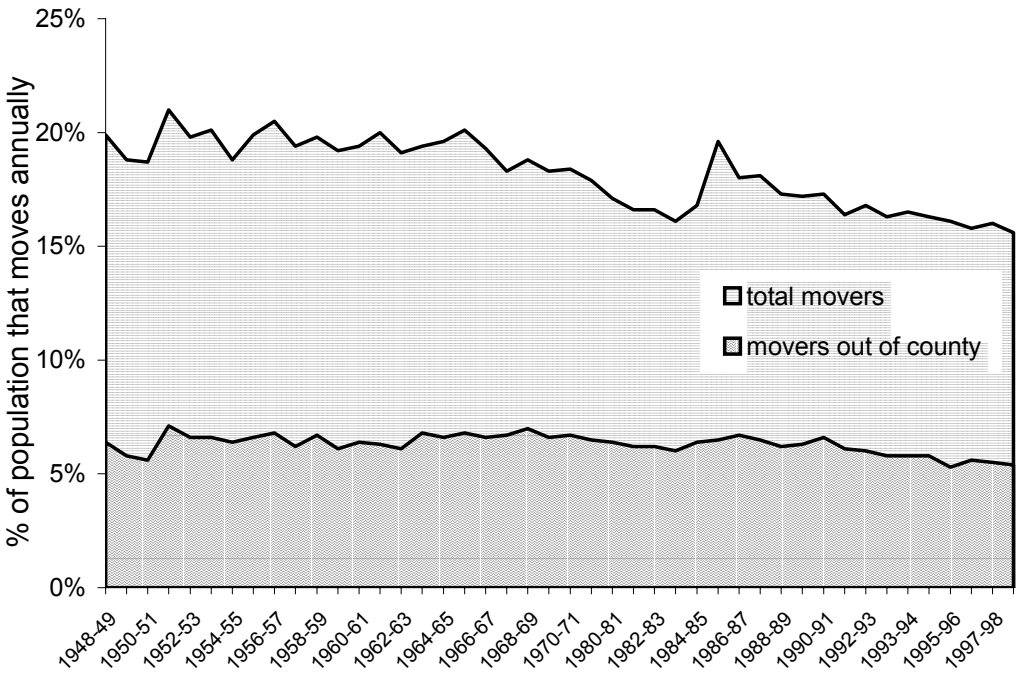


Figure 8: Annual geographic mobility data showing percent of population that reported moving during the year that proceeded each Annual Demographic Supplement to the Current Population Survey. The average percent movers for the period from 1989 to 1998 for the "total movers" category (top line), is 16.2% ($\sigma = 0.6$) of which renters and owners moved at a rate of 32% and 8% per year, respectively. For movers out of the county, the average over the same period is 5.9% ($\sigma = 0.4$) of which renters and owners moved at a rate of 11% and 3%, respectively.

8.3 Overview of Existing Methods

Israeli and Nelson, (1992) used weighted data from the 1985 and 1987 AHS-N surveys to estimate expected total residence time for the following groups: all households, renters, owners, urban households, rural households, farms (subset of rural households), and households in four geographic regions (northeast, midwest, south and west). They used a semi-analytical approach in which they related the fraction of households that moved into their current residence *t* years before the survey to the fraction of households just moving in at the time of the survey that will be found in the same residence *t* years from now. Values of the fraction of households that moved into their current residence *t* years before the surveys were calculated from American Housing Survey data and fit with a five-parameter survival function. Three of the five fitted parameters were then used to estimate the expected total residence time.

Johnson and Capel (1992) used a Monte Carlo approach to develop distributions of residential occupancy time by sex and age. In their simulations, they used population and mobility data from the U.S. Bureau of the Census and mortality data from the National Center for Health Statistics. The data that they used represent the 1987 U.S. population. The first step in their process was to determine the number of persons in each demographic group of interest, (e.g., the number of males or females of a given age). Next, they developed mobility tables and mortality tables for the different demographic groups. Mobility tables give the probability that a person in the demographic group did not move during the previous year (see Figure 8). Mortality tables give the probability that a person in the demographic group will die during the upcoming year. They then applied a Monte Carlo algorithm that generates current residence time using the mobility tables and future residence time using both the mobility and mortality tables. The occupancy period is then estimated as one plus the sum of the current and future residence times. Johnson and Capel estimated occupancy period for 500,000 simulated persons.

Finley *et al*. (1994) summarized and evaluated the work of both Israeli and Nelson and Johnson and Capel in their review of distributions of exposure factors. In addition, they derived a formula for and calculated the residential occupancy periods of children born in the given household based on moving rates from the U.S. Bureau of the Census. Finley et al. recommend using the estimates of Israeli and Nelson for exposure assessments that depend on housing unit characteristics such as geographic location, etc., and using the distributions of Johnson and Capel for exposure assessments for individuals of specified ages.

Price et al. (1992) describe a simulation approach using mortality and mobility data to estimate what they term the Fraction of Life Exposed, and is defined as the exposure duration divided by the lifetime of the simulated individual. The model generates a hypothetical person based on current distributions of age and sex in the U.S. population then simulates a lifetime of exposure for that individual based on age specific mobility and mortality data.

Price et al, (1998) also describe a somewhat different computational approach for calculating the total duration (*TD*) of a behavior based on the reported duration (*RD*) and the starting age (*SA*) of the behavior. Their approach is based on relating the probability of *RD* to the probability of *TD* given the value of *SA*. Price *et al*. (1998) applied their method to surveys of anglers rather than to housing data. However, the approach should be applicable to housing data given information about reported current residence time and starting age. In the context of housing data, *TD* would be the total time an individual lives in a residence, *RD* would be the reported residence time, and *SA* would be the age of the individual when he/she first occupied the residence.

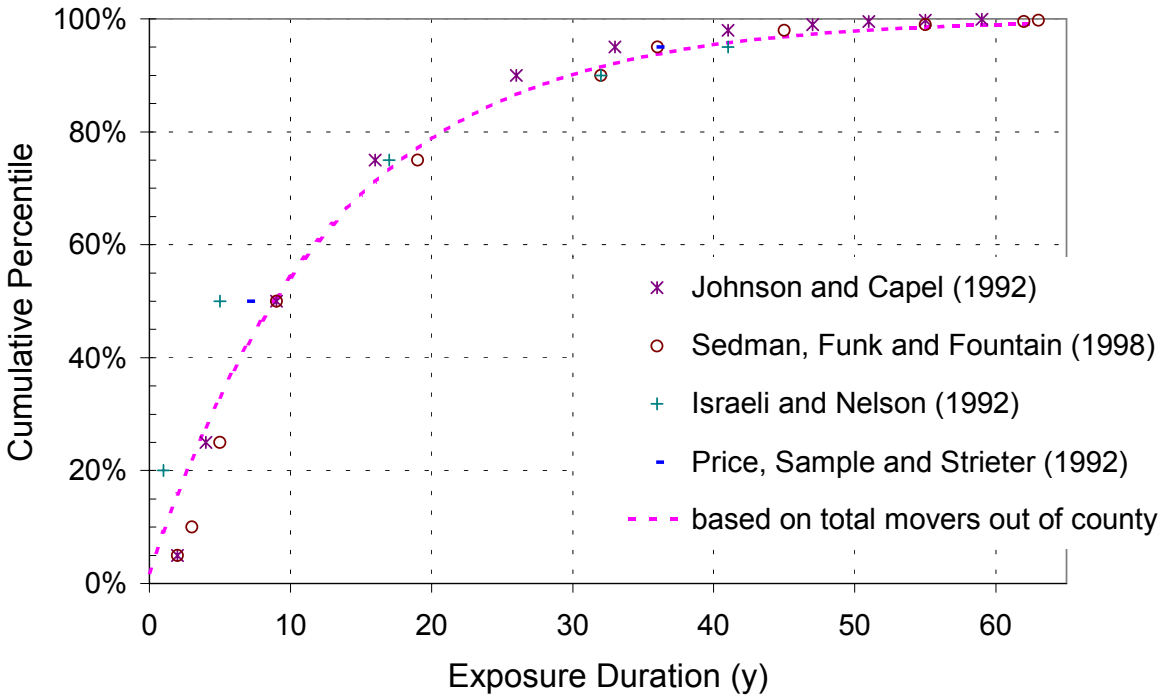


Figure 9: Comparison of the estimates for occupancy period (total time that an individual lives in current residence) from different methods. The dash line is derived from frequency (used here as probability) of moving to a different county during a given year (Hanson, 1997; USDOC, 1999).

Finally, Sedman et al. (1998) used public records on property title transfers to develop surrogates for occupancy periods. The specific public records used were the tax records for single family residences from the Tax Assessors office. The advantage of the method used by

Sedman et al. was that the records provide a closed interval for total residence duration that is not available in the other methods. The limitation is that the method can only be used for owner-occupied housing and it assumes that the individuals live in the house during the entire period of ownership and that they subsequently move upon sale of the property.

The various percentiles reported in Israeli and Nelson, (1992), Johnson and Capel (1992), Price et al. (1992) and Sedman et al., (1998) are graphically compared in Figure 9.

As illustrated in Figure 9, all of the existing methods give comparable results, particularly when one considers the inherent uncertainties regarding the demographic makeup of the target population. The dash line in Figure 9 is derived solely from information about the annual percentage of the U.S. population that moves to a different county. This simple approach agrees reasonably well with the values predicted by the other four methods. Therefore, given the general absence of both site-specific mobility data and information about the total distance of a move, we cannot justify the complexity of any of the previously described models. Rather, we use readily available information on the tenure-based mobility of the national population and the approximate distance of move (out of metropolitan area, county, state,…) as surrogate data for exposure duration (i.e., duration of time spent in the target region). A method for deriving exposure duration distributions from this information and accounting for site-specific characteristics is given in the following section.

8.4 Recommendation for Estimating Distributions of Exposure Duration

In estimating exposure duration, we assume no prior knowledge of individual ages within the target population for reasons given in Section 8.2. Rather, we estimate *ED* for a given household in the target region using tenure-based mobility data. An exposure duration value for a given home is

$$ED_i = \sum_{i=1}^{n_i} Y_i \tag{40}$$

where

$$Y_i = \begin{cases} 1 \text{ year} & \text{if } R < P_d \\ 0 & \text{if } R \geq P_d \end{cases} \tag{41}$$

and R is a random number between 0 and 1, P_d is the demographically-based probability of moving (i.e. the probability of moving out of county for the U.S. population) and n_i is the total number of random draws, until $R \geq P_d$ (i.e., the household moves), that satisfy the constraint $R < P_d$. This process is repeated for a given home until a full averaging time (*AT*) profile is

established. Further, the number of homes within the target region and the composition of the housing stock (rental versus owner occupied) are used to construct a set of random ED_{ij} values over a specified AT for j different homes having any combination of P_d values (e.g., d equals renter or owner). The process is easily applied using available statistical, mathematical or spreadsheet software packages. The approach can be applied to simulate occupancy period over a specified averaging time for a given number of homes based on the tenure composition across the housing stock, i.e., the number of owner occupied and renter occupied homes within the target region.

Information about the composition of the housing stock within the target region can be accessed through the U.S. Census interactive data retrieval service at http://homer.ssd.census.gov/doc/lookup_doc.html or by direct survey of the target population. Information about the mobility of different demographic subsets of the population can be accessed at http://www.census.gov by looking up "Migration" in the "Subjects A to Z" index. In applying this method to different geographical regions of the U.S., we found that the two-parameter exponential distribution (Johnson *et al*., 1994a) consistently provided the best fit to different sets of simulated exposure duration values. But this is expected because the model is simply a mixture of demographically based moving rates. For target regions where the number of houses is small, we recommend that the sampling and fitting process be repeated several times to estimate uncertainty about the model parameterization. In addition, the inputs to the process (e.g., demographically based probability of moving) should be defined as distributions to explicitly account for uncertainty and lack of site-specificity.

Although not as robust as the method for estimating body weight, we believe that the recommended method for deriving exposure duration distributions is adequate given the availability, and quality of surrogate data. That is, the complexity of the approach is appropriate given the state of the science and general lack of site-specific data.

9.0 Water Intake

For many exposure scenarios, water intake is an important input to the risk assessment process. Reported valures for water intake rate are from the same nationally representative survey that includes self-reported values of body weight (USDA, 1998). The data on water intake are collected using 24-hour recall about the amount of food, beverage and plain water that the individual surveyed has consumed the previous day. Food and beverage values are subsequently converted to water intake using food recipe files (USEPA, 2000a). Although the 24-hour recall sampling method and the food-to-water conversions likely lead to imprecision

and/or bias in the data, an equally if not more important limitation of the data, in the context of PRA, is the fact that only two nonconsecutive days are sampled for each individual.

These "short-term" recall samples have been used to represent the distribution of individual long-term average intake for the US population (Burmaster, 1998; Ershow *et al*., 1991; Ershow and Cantor, 1989; Roseberry and Burmaster, 1992; USEPA, 2000a). Although short-term data provides a good estimate of the population mean it greatly inflates the variance in the distribution of average daily intakes for individuals. As a consequence the distribution of long-term daily average intake across the population is also inflated. This results in significant over-estimates of exposure for the high end of the population and significant under-estimates for the low end (Buck *et al*., 1995; Slob, 1993; Slob, 1996; USEPA, 1997a; Wallace *et al*., 1994). Variance associated with exposure factor distributions is critical in the context of probabilistic analyses where both the range and likelihood of an outcome are of interest. Distributions that are derived using short-term survey data not only lack relevance for the tails of the distributions, they are wrong, but wrong in way that is predictable.

Without adjusting the data to compensate for this variance inflation effect, there is little value in understanding the demographic characteristics of the population and using higher-order statistical models that provide more precise fits to the data. Therefore, the limitations associated with short-term survey data must be addressed before the data can be used to construct "recommended" exposure factor distributions for chronic (long-term) exposure scenarios. Method development and evaluation of existing methods for improving the relevance of short-term data is beyond the scope of this study. As a result, we are not able to provide either a set of demographically based distributions or a method for constructing site-specific PDFs for water intake. However, we do provide in the following sections, a brief illustration of the extent to which the distribution of long-term average water intake is effected by sample size and by assumptions about the variance structure within the data. We conclude the section with recommendations for future studies that begin to address the limitations in the data.

9.1 The Implications of using Short-term Data to Estimate Long-term Averages

The problem of using small, short-term samples to estimate long-term averages comes up frequently in environmental statistics, and has previously been considered for food and soil ingestion data (Slob, 1993; Slob, 1996; Stanek *et al*., 1998; Wallace *et al*., 1994). Previous researchers have used parametric modeling to relate the short- and long-term exposure distributions. A common assumption made for both convenience and for substantive reasons is that either individual daily observations or annual consumption values are lognormally distributed. But failures to meet these assumptions have been noted and alternative assumptions presented (Wallace et al. 1994). Other assumptions used in existing methods include the lack of measurement error and specific variance structure across the population (i.e., constant CVs,

constant standard deviation or independence of day-to-day variation within and/or between individuals).

We expand on the model of Buck et al. (1995) to illustrate how short-term survey data might influence the final selection of a PDF. In this approach, we assume that the amount of water that a person might report consuming on a given day is represented mathematically as

$$X_{ij} = \mu_i + \tau_{ij} + \tau_{ik} \qquad (42)$$

where X_{ij} is the amount person i reports consuming on day j, μ_i is the true long-term average intake for person i, τ_{ij} is the deviation for person i from μ_i on day j due to day-to-day variability and τ_{ik} is the deviation due to measurement or reporting error.

The long-term average intake for the ith person (μ_i in equation 42) is a random variable from distribution function G which is the true distribution of per capita intakes across the population of interest (e.g., the distribution needed for a PRA concerned with a chronic exposures at the population level). For simplicity we assume that distribution G is lognormal with mean, μ_p and variance, σ^2_p where the subscript p indicates population.

Deviation from the true long-term average intake for the ith person on day j (τ_{ij} in equation 42) is a random variable from distribution function H with mean = 0 (assumes no bias) and variance = σ^2_i for individual i. Neglecting reporting/measurement error (τ_{ik}) in Eq. 42, we pose three possible scenarios for day-to-day variability of individual intakes.

(1) Day-to-day variability, σ^2_i, is the same for all individuals in the population.

(2) Day-to-day variability, σ^2_i, is different for each individual in the population, e.g., σ^2_i is a random variable drawn from distribution function F.

(3) The coefficient of variation for day-to-day variability (σ_i/μ_i) is the same for all individuals.

To evaluate the influence of the three scenarios on the distribution of long-term average, we specify parameters for distribution function G (lognormal) that are near what we expect to see for an exposure factor such as water intake (μ_p = 1.5 liter per day and σ_p = 0.9). Next we draw a set of 250 individual means (μ_i) from distribution G where each mean represents the true long-term average intake for an individual in the population.

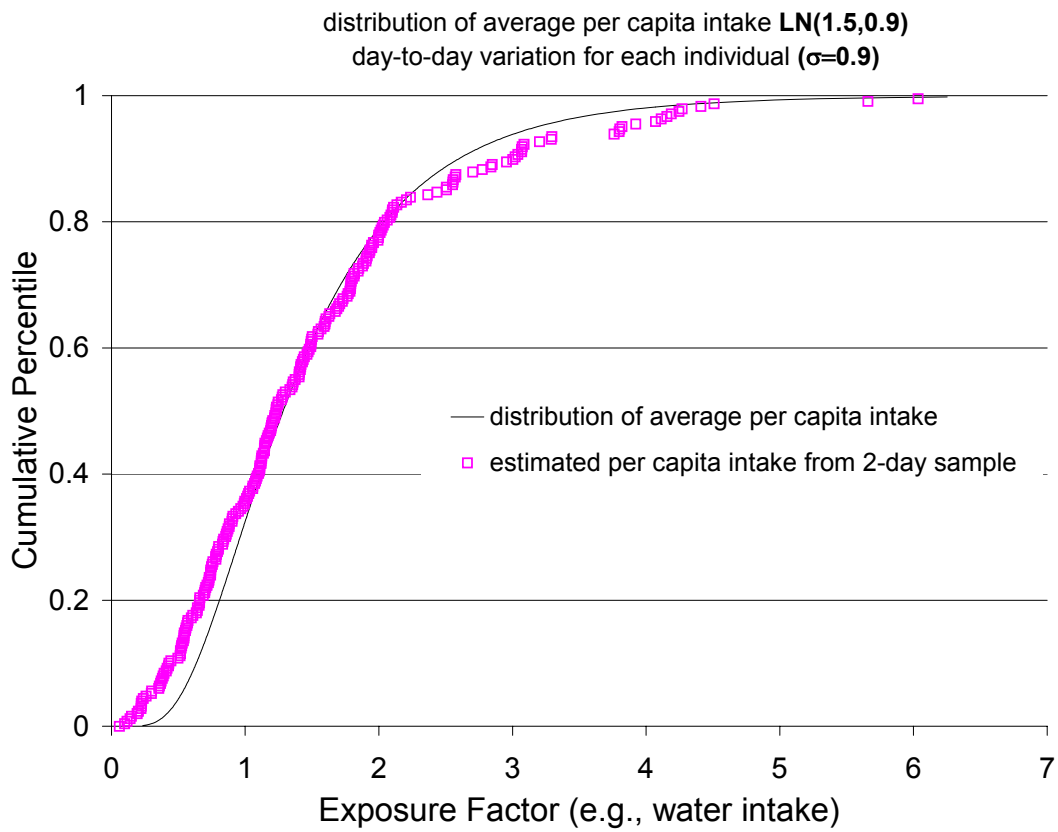


Figure 10: Results for Case 1. Comparison of the "true" distribution of average values (solid line) with an empirical distribution constructed using 250 hypothetical individuals where two samples are drawn for each individual and used to estimate the average intake for that person. The variance structure for individuals in the population is assumed to be constant from day-to-day for all members of the population; $\sigma = 0.9$.

For case 1 we use the 250 values along with a constant standard deviation ($\sigma_p = 0.9$) to simulated two sets of X_{ij} (day 1 and day 2) using equation 42 for each sample person. The resulting empirical distribution function constructed using the average of these two simulated days is shown in Figure 10 along with the actual distribution function of average per capita intake across the population.

For the second case we again use the set of 250 individual means generated from LogNorm(1.5,0.9), but the day-to-day variation, τ_{ij} in Eq. 42, is now assumed to be a random variable drawn from a lognormal distribution with mean = 1.35 and $\sigma = 1.00$. Thus, the values used to construct the empirical distribution of short-term average intakes are from the

distribution LogNorm{LogNorm(1.5,0.9), LogNorm(1.34,1.00)}. The result for Case 2 is illustrated in Figure 11.

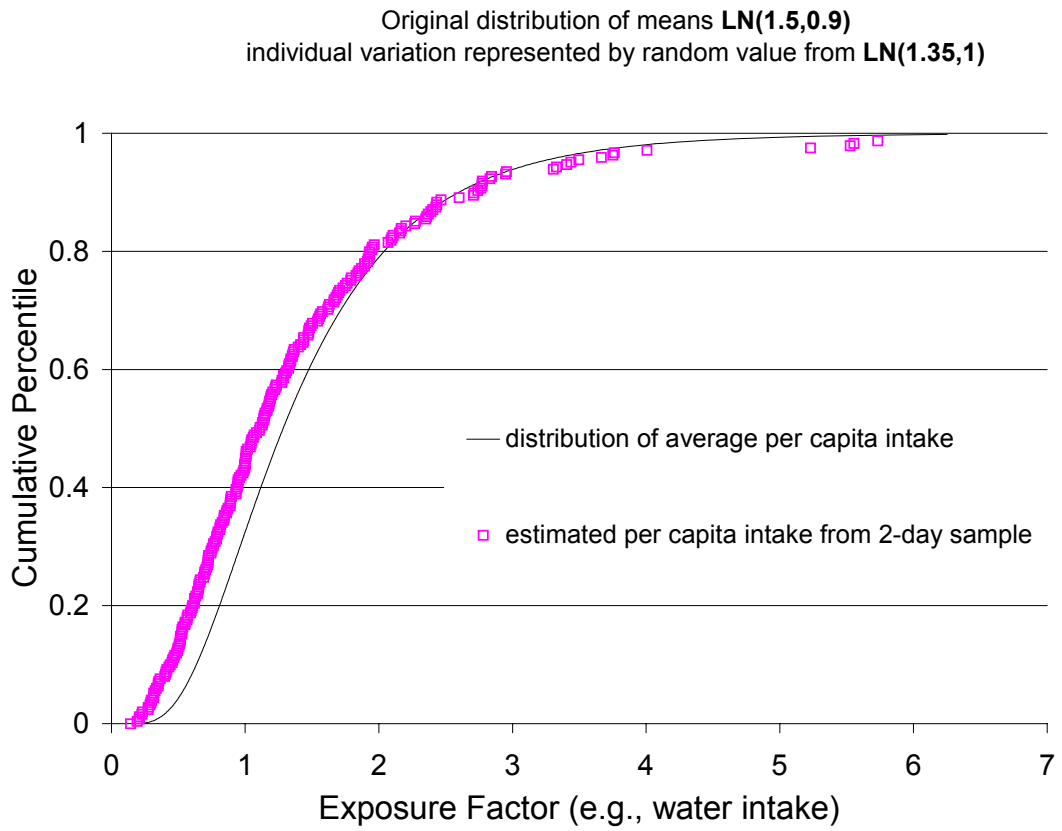


Figure 11: Results for Case 2. Comparison of the "true" distribution of average values (solid line) with an empirical distribution constructed using short-term data (2-day) where the day-to-day variance structure in the population is assumed to be distributed randomly with lognormal distribution; LogNorm(1.35,1.00).

Case 3 uses the same individual means but the day-to-day variation is represented by a constant coefficient of variation (CV = 0.5) for all individuals. The result for this case is shown in Figure 12.

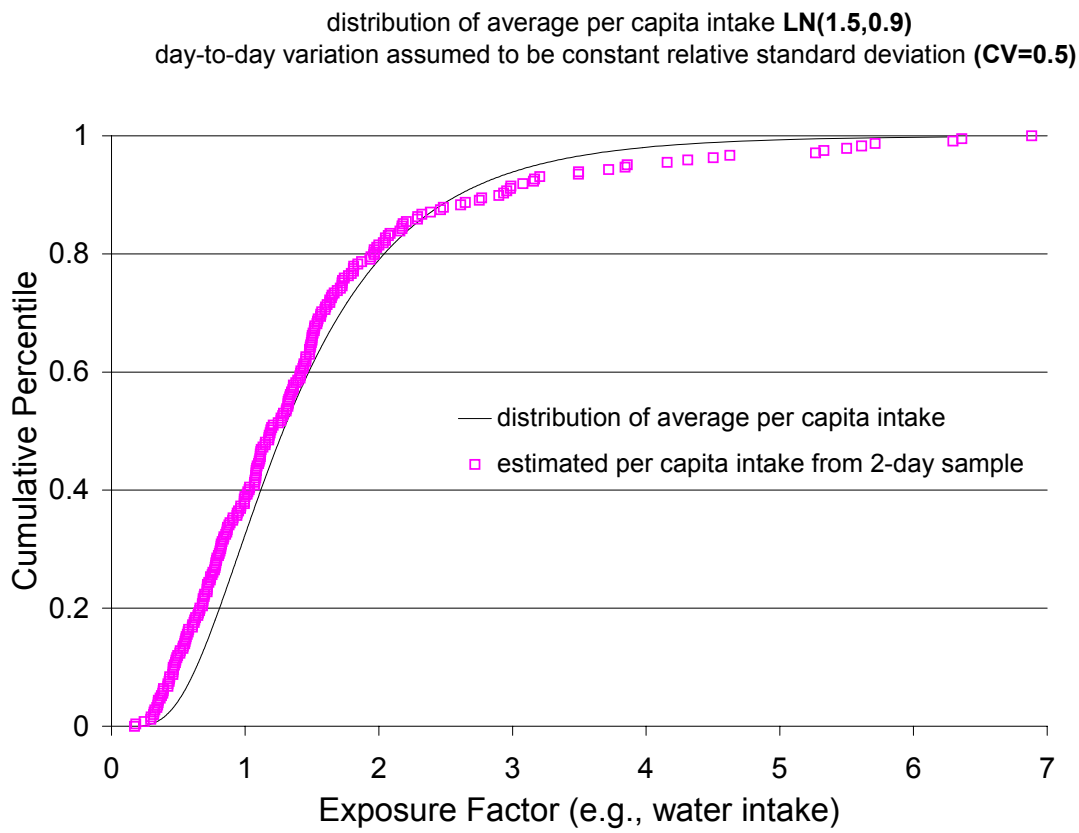


Figure 12: Results from Case 3. Comparison of the "true" distribution of average values with an empirical distribution constructed using short-term data (2-day) where the day-to-day variance structure in the population is assumed to be represented by a constant relative standard deviation, i.e., day-to-day variation for each member of the population has the same coefficient of variation; CV(0.5).

In each case, regardless of our assumption about the structure of the day-to-day variance in intakes for individuals in the population, the empirical distribution of the data from the 2-day average intakes clearly deviates from the "true" distribution of average intakes for the population. The implication for constructing exposure factor distributions is that regardless of how sophisticated our statistical model is, or how well we "fit" the model to the data, the relevance of the resulting PDF to the actual (all-be-it unknown) exposure factor distribution is limited, i.e., the resulting PDF is wrong. It is clear that until we can develop a scientifically defensible method for updating the relevance of short-term data, the exposure factor PDFs that are based on this data will be unreliable.

9.2 A Final Recommendation for Water Intake

Although lognormal distributions have been shown to provide a reasonable fit to water intake data, (Burmaster, 1998; Roseberry and Burmaster, 1992) a recent report suggests a more complex 5-parameter generalized F-distribution may be appropriate (USEPA, 2000b). Our own screening analysis of a range of demographically based sets of water intake values found that mixture models work best for fitting large data sets (results not shown). However, without addressing the issue of variance inflation due to the survey design, it will not be possible to identify the most appropriate parametric model or to recommend a transparent approach for reconstructing task-specific default distributions.

To address this limitation in the data we suggest that studies be targeted at improving the way short-term survey data is used to derive distributions of long-term average intakes. The problem described and demonstrated above is also relevant to a number of other exposure factors (e.g., food intakes, activity diaries, etc.) and as such warrants close consideration. We suggest that the general task of improving data relevance be focused on two main efforts using water intake as the case study but where the results would be relevant to the general case of using short-term data to estimate long-term exposure factors. We recommend the following:

1. Use existing water intake data to construct an initial distribution by assuming each person's daily intake can be drawn from a lognormal distribution as illustrated in the case studies, with person-dependent geometric mean (GM) and a partially pooled estimate of geometric standard deviation (GSD). The personal GM (and GSD) for each person cannot simply be extracted from the two available measurements per person because that will result in an inflated variability. Instead, Bayesian modeling can be used to estimate annual mean consumption from a combination of the data and relevant factors such as sex, age, region of the country, etc. Following the standard techniques of posterior predictive updates, the estimates of GMs and GSDs could be used to generate simulated data, which could then be compared to the observed data in order to investigate model violations, with particular emphasis on the upper tail of the distribution. This task would essentially extend previous lognormal modeling of intake data (Slob, 1993; Slob, 1996; Stanek *et al*., 1998; Wallace *et al*., 1994) by (i) including demographic factors (i.e., age, sex, body weight, etc.) that are expected to influence water intake, and by (ii) using posterior predictive checks (Gelman *et al*., 1995) to quantify model violations. The outcome of this research thrust could provide more accurate estimated water intake distributions for various demographic categories and provide quantitative error estimates for each quantile of the distribution for the population or specific cohorts within the population.

2. It is possible that the first recommended task will yield fully satisfactory estimates of water intake distributions by sex, age, region, etc., with only insignificant model violations, and thus provide accurate prediction of the high tail of the intake distribution. More likely, however, some violations will remain. For example, the use of a lognormal distribution, with its extended high tail, may be inappropriate. A truncated lognormal, a gamma or log-gamma distribution, or other distribution, might be found to better agree with the data, particularly at the high end of the distribution. Fitting such models is substantially more complicated than fitting lognormal models, which is one reason they are not frequently used. Still, we must use models that fit the data, not the other way around, and it may be that these alternative models will prove necessary. Thus, the second recommendation is to evaluate the use of non-lognormal models, as needed, to address substantial problems with the lognormal assumptions, particularly at the high end of the intake distribution. Although conceptually straightforward, fitting modified distributions (such as truncated lognormal distributions) would require a significant level of computational effort. Once the short-term data are appropriately transformed to represent long-term statistics in the population then the general approach described in Section 7 for body weight could be developed to reconstruct task-specific distribution for target populations at various geographic locations and scales.

<Blank Page>

References

Beck, M. B., and Chen, J. (2000). "Assuring the Quality of Models Designed for Predictive Tasks". *In* Sensitivity Analysis (A. Saltelli, K. Chan and E. M. Scott, eds.), pp. 475. John Wiley & Sons, LTD, New York.

Bennett, D. H., McKone, T. E., Matthies, M., and Kastenberg, W. E. (1998). "General Formulation of Characteristic Travel Distance for Semivolatile Organic Chemicals in a Multimedia Environment". *Environmental Science and Technology* **32**(24):4023-4030.

Bevington, P. R. (1969). Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York.

Binkowitz, B. S., and Wartenberg, D. (2001). "Disparity in Quantitative Risk Assessment: A Review of Input Distributions". *Risk Analysis* **21**(1):75-90.

Bogen, K. T., and Spear, R. C. (1987). "Integrating Uncertainty and Variability in Environmental Risk Assessment". *Risk Analysis* **7**:427-436.

Buck, R. J., Hammerstrom, K. A., and Ryan, P. B. (1995). "Estimating Long-term Exposures from Short-term Measurements". *Journal of Exposure Analysis and Environmental Epidemiology* **5**(3):359-373.

Burmaster, D. E. (1998). "Lognormal Distributions for Total Water Intake and Tap Water Intake by Pregnant and Lactating Women in the United States". *Risk Analysis* **18**(2):215-219.

Burmaster, D. E., and Anderson, P. D. (1994). "Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessments". *Risk Analysis* **14**(4):477-481.

Burmaster, D. E., and Crouch, E. A. C. (1997). "Lognormal Distributions for Body Weight as a Function of Age For Males and Females in the United States, 1976-1980". *Risk Analysis* **17**(4):499-505.

Campolongo, F., Saltelli, A., Sorensen, T., and Tarantola, S. (2000). "Hitchhiker's Guide to Sensitivity Analysis". *In* Sensitivity Analysis (A. Saltelli, K. Chan and E. M. Scott, eds.). John Wiley & Sons, LTD, New York.

Chen, J., and Beck, M. B. (1998). “Quality Assurance of Multi-media Model for Predictive Screening Tasks,” Rep. No. EPA/600/R-98/106. National Exposure Research Laboratory, Athens, Georgia.

Conover, W. J. (1999). Practical Nonparametric Statistics, 3rd/Ed. John Wiley & Sons, New York.

Cullen, A. C., and Frey, H. C. (1999). Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs, Plenum Press, New York.

D'Agostino, R. B., and Stephens, M. A., eds. (1986). Goodness of Fit Techniques, Vol. 68, pp. 560. Marcel Dekker, Inc., New York.

Draper, N. R., and Smith, H. (1981). Applied Regression Analysis, Wiley, New York.

Eiceman, G. A., Urquhart, N. S., and O'Connor, G. A. (1993). "Logistic and Economic Principles in Gas Chromatography-Mass Spectrometry Use for Plant Uptake Investigations". *Journal of Environmental Quality* **22**(January-March):167-173.

Ershow, A. G., Brown, L. M., and Cantor, K. P. (1991). "Intake of Tapwater and Total Water by Pregnant and Lactating Women". *American Journal of Public Health* **81**(3):328-334.

Ershow, A. G., and Cantor, K. P. (1989). "Total Water and Tapwater Intake in the United States: Population-Based Estimates of Quantities and Sources," Prepared under the National Cancer Institute Order #263-MD-810264. Life Sciences Research Office Federation of American Societies for Experimental Biology, Bethesda, MD, May 1989.

Evans, M., Hastings, N., and Peacock, B. (1993). Statistical Distributions, Second Edition/Ed. John Wiley & Sons, Inc., New York.

Finley, B., Proctor, D., Harrington, P., Paustenbach, N., and Price, P. (1994). "Recommended Distributions for Exposure Factors Frequently Used in Health Risk Assessment". *Risk Analysis* **14**(4):533-553.

Freund, R. J., and Minton, P. D. (1979). Regression Methods: A Tool for Data Analysis, Vol. 30, M. Dekker, New York.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). Bayesian Data Analysis, Chapman and Hall, New York.

Hansen, K. A. (1998). "Seasonality of Moves and Duration of Residence," Rep. No. P70-66. U.S. Bureau of the Census, Current Population Reports, Washington DC, October 1998.

Hanson, K. (1997). "Geographical Mobility: March 1995 to March 1996,". US Department of Commerce, Economics and Statistics Administration, November 1997.

Helton, J. C., Iman, R. L., Johnson, J. D., and Leigh, C. D. (1989). "Uncertainty and Sensitivity Analysis of a Dry Containment Test Problem for the MAEROS Aerosol Model". *Nuclear Science and Engineering* **102**:22-42.

Hertwich, E. G., McKone, T. E., and Pease, W. S. (2000). "A Systematic Uncertainty Analysis of an Evaluative Fate and Exposure Model". *Risk Analysis* **20**(4):439-454.

Hornberger, G. M., and Spear, R. C. (1980). "Eutrophication in Peel Inlet, I, The Problem: Defining Behavior and a Mathematical Model for the Phosphorus Scenario". *Water Research* **14**:29-42.

IAEA. (1989). "Evaluating the Reliability of Predictions made using Environmental Transport Models. Safety Series 100,". International Atomic Energy Agency, Vienna, Austria.

Iman, R. L., and Conover, W. J. (1980). "Small Sample Sensitivity Analysis Techniques for Computer Models, with and Application to Risk Assessment". *Communications in Statistics* **A9**(17):1749-1842.

Iman, R. L., and Helton, J. C. (1988). "An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models". *Risk Analysis* **8**(1):71-90.

Israeli, M., and Nelson, C. B. (1992). "Distribution and Expected Time of Residence for U.S. Household". *Risk Analysis* **12**(1):65-72.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994a). Continuous Univariate Distributions, Second/Ed., Vol. 2, John Wiley & Sons Inc., New York.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994b). Continuous Univariate Distributions, Second/Ed., Vol. 1, John Wiley & Sons Inc., New York.

Johnson, T., and Capel, J. (1992). "A Monte Carlo Approach to Simulating Residential Occupancy Periods and Its Application to the General U.S. Population," Rep. No. EPA 450/3-92-011. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Emissions Standards Division, Research Triangle Park, NC, August 1992.

Kuczmarski, R. J., Ogden, C. L., Grummer-Strawn, L. M., and others. (2000). "CDC Growth Charts: United States, Advanced data from vital and health statistics," Rep. No. 314. National Center for Health Statistics, Hyattsville, Maryland, June 8, 2000.

LSRO. (1995). "Third Report on Nutrition Monitoring in the United States: Volume 1.," Life Sciences Research Office, Federation of American Societies of Experimental Biology, Washington D.C..

Maddalena, R. L., McKone, T. E., Bodnar, A., and Jacobson, J. (1999). "Development and Evaluation of Probability Density Functions for a Set of Human Exposure Factors," Rep. No. LBNL 43604. Environmental Energy Technologies Division; Lawrence Berkeley National Laboratory, Berkeley, June, 1999.

Maddalena, R. L., McKone, T. E., Hsieh, D. P. H., and Geng, S. (2001). "Influential Input Classification in Probabilistic Multimedia Models". *Stochastic Environmental Research and Risk Assessment* **15**(1):1-17.

McKone, T. E. (1992). "CalTOX, A Multimedia Total-Exposure Model for Hazardous-Waste Sites Part III: The Multiple-Pathway Exposure Model," Rep. No. UCRL-CR-111456PtIII. Prepared for: The Office of Scientific Affairs, Department of Toxic Substances Control, California Environmental Protection Agency, Sacramento, November 1992.

McKone, T. E. (1994). "Uncertainty and Variability in Human Exposures to Soil Contaminants Through Home-Grown Food - A Monte Carlo Assessment". *Risk Analysis* **14**(4):449-463.

McKone, T. E., and Bogen, K. T. (1992). "Uncertainties in Health-Risk Assessment: An Integrated Case Study Based on Tetrachloroethylene in California Groundwater". *Regulatory Toxicology and Pharmacology* **15**:86-103.

McKone, T. E., Kiefer, K., Currie, R., Geng, S., and Hsieh, D. P. H. (1996). "Representing Uncertainty in Risk Assessment Part 1: Constructing Distributions," Office of Environmental Health Hazard Assessment. California Environmental Protection Agency, Berkeley, California, January 1996.

McKone, T. E., Maddalena, R. L., Currie, R., Kiefer, K., and Hsieh, D. P. H. (1995). "Representing Uncertainty in Risk Assessment: Part 4: Ranking Uncertain Parameters,". Risk Science Program, University of California, Davis, Prepared for the Office of Environmental Health Hazard Assessment, California Environmental Protection Agency, Berkeley.

Morgan, M. G., and Henrion, M. (1990). Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis, Cambridge University Press, New York.

Myers, L., Lashley, J., and Whitmore, R. (1998). “Development of Statistical Distributions for Exposure Factors, Final Report,”. Research Triangle Institute, Research Triangle Park, North Carolina, March 18, 1998.

NAS (1994). “Science and Judgment in Risk Assessment,”. National Research Council, Commission on Life Sciences, Board on Environmental Studies and Toxicology, Committee on Risk Assessment of Hazardous Air Pollutants, Washington D.C., 1994.

NCHS (1994). “Plan and operation of the Third National Health and Nutrition Examination Survey, 1988-1994,” Rep. No. 1(32). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. Vital Health Stat, Washington, DC, July 1994.

NCHS (1996). “National Health and Nutrition Examination Survey, III (NHANES III) 1988-94 reference manuals and reports (CD-ROM).,”. National Center for Health Statistics, Department of Health and Human Services, Centers for Disease Control and Prevention, Hyattsville, MD, Revised October 1997.

Okajima, Y., Togo, M., Kitagawa, G., and Nishikawa, S. (2000). "Time Series Analysis of Monthly Body Weight and Blood Pressure of One Man from 29 to 65 years". *American Journal of Human Biology* **12**:526-541.

Price, P. S., Sample, J., and Strieter, R. (1992). "Determination of Less-Than- Lifetime Exposures to Point Source Emissions". *Risk Analysis* **12**(3):367-382.

Price, P. S., Scott, P. K., Wilson, N. D., and Paustenbach, D. J. (1998). "An Empirical Approach for Deriving Information on Total Duration of Exposure from Information on Historical Exposure". *Risk Analysis* **18**(5):611-619.

Roseberry, A. M., and Burmaster, D. E. (1992). "Lognormal Distributions for Water Intake by Children". *Risk Analysis* **12(1)**:99-104.

Saltelli, A., Chan, K., and Scott, M. E., eds. (2000). Sensitivity Analysis, pp. 475. John Wiley & Sons, LTD, New York.

Scott, E. M., Saltelli, A., and Sorensen, T. (2000). "Practical Experience in Applying Sensitivity and Uncertainty Analysis". *In* Sensitivity Analysis (A. Saltelli, K. Chan and E. M. Scott, eds.), pp. 475. John Wiley & Sons, LTD, New York.

Sedman, R., Funk, L. M., and Fountain, R. (1998). "Distribution of Residence Duration in Owner Occupied Housing". *Journal of Exposure Analysis and Environmental Epidemiology* **8**(1):51-57.

Seiler, F. A., and Alvarez, J. L. (1996). "On the Selection of Distributions for Stochastic Variables". *Risk Analysis* **16**(1):5-18.

Slob, W. (1993). "Modeling Long-Term Exposure of the Whole Population of Chemicals in Food". *Risk Analysis* **13**(5):525-530.

Slob, W. (1996). "A Comparison of Two Statistical Approaches to Estimate Long-Term Exposure Distributions from Short-Term Measurements". *Risk Analysis* **16**(2):195-200.

Smith, R. L. (1994). "Use of Monte Carlo Simulation for Human Exposure Assessment at a Superfund Site". *Risk Analysis* **14**(4):433-439.

Snedecor, G. W., and Cochran, W. G. (1989). Statistical Methods, Eighth/Ed. Iowa State University Press / AMES, Ames, Iowa.

Spear, R. C., Grieb, T. M., and Shang, N. (1994). "Parameter Uncertainty and Interaction in Complex Environmental Models". *Water Resources Research* **30**(11):3159-3169.

Spear, R. C., and Hornberger, G. M. (1980). "Eutrophication in Peel Inlet, II, Identification of Critical Uncertainties via Generalized Sensitivity Analysis". *Water Research* **14**:43-49.

Stanek, E. J., Calabrese, E. J., and Xu, L. (1998). "A Caution for Monte Carlo Risk Assessment of Long Term Exposures based on Short-Term Exposure Study Data". *Human and Ecological Risk Assessment* **4**(2):409-422.

Thompson, K. M. (1999). "Developing Univariate Distributions from Data for Risk Analysis". *Human and Ecological Risk Assessment* **5**(4):755-783.

Thomspon, K. M., Burmaster, D. E., and Crouch, E. A. C. (1992). "Monte Carlo Techniques for Quantitative Uncertainty Analysis in Public Health Risk Assessment". *Risk Analysis* **12**(1):53-62.

USDA (1998). "1994-96 Continuing Survey of Food Intakes by Individuals and 1994-96 Diet and Health Knowledge Survey," Rep. No. PB98-500457. Department of Agriculture, Agricultural Research Service, Springfield, VA.

USDOC (1995). "American Housing Survey (AHS) for the United States in 1995." Rep. No. Current Housing Reports H150/95RV. US Department of Commerce Economics and Statistics Administration, Bureau of Census, Washington, DC, .

USDOC (1999). "Annual Geographical Mobility Rates, By Type of Movement: 1947-1999 Table A-1.", Vol. 2000. U.S. Census Bureau.

USDOC (2000). "Current Population Survey Technical Paper 63: Design and Methodology," Rep. No. TP63. U.S. Department of Labor, Bureau of Labor and Statistics; U.S. Department of Commerce Economics and Statistics Administration, U.S. Census Bureau, March 2000.

USEPA (1997a). "Exposure Factors Handbook (Final) Volumes I, II, III," Rep. No. EPA/600/P-95/002Fa. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Washington D.C., August 1997.

USEPA (1997b). "Guiding Principles for Monte Carlo Analysis," Rep. No. EPA/630/R-97/001. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC, March 1997.

USEPA (1998). "Guidance for Submission of Probabilistic Exposure Assessments to the Office of Pesticide Programs,". U.S. Environmental Protection Agency, Office of Pesticide Programs, November 1998.

USEPA (1999a). "Report of the Workshop on Selecting Input Distribution for Probabilistic Risk Assessment," Rep. No. EPA/630/R-98/004. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington, D.C. 20460, January 1999.

USEPA (1999b). "Risk Assessment Guidance for Superfund (RAGS): Volume 3A Process for Conducting Probabilistic Risk Assessment. DRAFT," Rep. No. EPA 000-0-99-000. Office of Emergency and Remedial Response, Washington D.C., December 30, 1999.

USEPA (2000a). “Estimated Per Capita Water Ingestion in the United States Based on Data Collected by the United States Department of Agriculture's 1994-1996 Continuing Survey of Food Intakes by Individuals,” Rep. No. EPA-822-R-00-008. U.S. Environmental Protection Agency, Office of Water, April 2000.

USEPA (2000b). “Options for Development of Parametric Probability Distributions for Exposure Factors,” Rep. No. EPA/600/R-00/058. U.S. Environmental Protection Agency, National Center for Environmental Assessment, Office of Research and Development, Washington D.C., July 2000.

Wallace, L. A., Duan, N., and Ziegenfus, R. (1994). "Can Long-Term Exposure Distributions be Predicted from Short-Term Measurements?". *Risk Analysis* **14**(1):75-85.

Appendix 1: Estimation of Body Weight Percentiles for Children

Percentiles for body weight per sex and age category for children from birth through 19 years are calculated using the standard normal distribution (z-score) for the desired percentile and Eq. 34 along with the *L, M, S* parameters listed below from reference (Kuczmarski *et al*., 2000). The values are listed below for convenience. The reader should refer to the original report for further details. The location is provided in the original report (Kuczmarski *et al*., 2000) for downloading an electronic copy of the variables.

Age in Months	Male			Female		
	L	*M*	*S*	*L*	*M*	*S*
0	1.815151	3.530203	0.152385	1.509188	3.399186	0.142107
0.5	1.547523	4.003106	0.146025	1.357944	3.797528	0.138076
1.5	1.068796	4.879525	0.136479	1.105538	4.544777	0.131734
2.5	0.695974	5.672889	0.129678	0.902597	5.230584	0.126893
3.5	0.419815	6.391392	0.124717	0.734121	5.859961	0.123025
4.5	0.219867	7.041836	0.12104	0.590235	6.437588	0.119841
5.5	0.077506	7.630425	0.118271	0.464392	6.96785	0.117167
6.5	-0.02191	8.162951	0.116154	0.352164	7.454854	0.114894
7.5	-0.08944	8.644832	0.11451	0.250498	7.902436	0.11295
8.5	-0.13341	9.08112	0.113217	0.157248	8.314178	0.111285
9.5	-0.1601	9.4765	0.112186	0.070886	8.693418	0.109864
10.5	-0.1743	9.835308	0.111355	-0.00968	9.043262	0.108661
11.5	-0.17972	10.16154	0.110676	-0.08526	9.366594	0.107656
12.5	-0.17925	10.45885	0.110119	-0.15641	9.666089	0.106835
13.5	-0.17518	10.73063	0.109657	-0.22356	9.944226	0.106183
14.5	-0.16932	10.97992	0.109274	-0.28701	10.20329	0.105691
15.5	-0.16311	11.20956	0.108956	-0.347	10.44541	0.10535
16.5	-0.15771	11.42207	0.108695	-0.40369	10.67251	0.10515
17.5	-0.15402	11.61978	0.108483	-0.45722	10.88639	0.105084
18.5	-0.15276	11.80478	0.108317	-0.5077	11.08868	0.105144
19.5	-0.15447	11.97897	0.108194	-0.55524	11.2809	0.105323
20.5	-0.15952	12.14404	0.108111	-0.59992	11.4644	0.105613
21.5	-0.16818	12.30154	0.108067	-0.64185	11.64043	0.106007
22.5	-0.18057	12.45283	0.108062	-0.68114	11.81014	0.106498
23.5	-0.1967	12.59913	0.108095	-0.71788	11.97454	0.107078
24.5	-0.2165	12.74154	0.108166	-0.75221	12.13456	0.10774
25.5	-0.23979	12.88102	0.108275	-0.78423	12.29102	0.108477
26.5	-0.26632	13.01842	0.108421	-0.8141	12.44469	0.109281
27.5	-0.29575	13.1545	0.108605	-0.84194	12.59622	0.110144
28.5	-0.32773	13.2899	0.108826	-0.86789	12.74621	0.111061
29.5	-0.36182	13.42519	0.109083	-0.8921	12.89517	0.112023
30.5	-0.39757	13.56088	0.109378	-0.91472	13.04357	0.113023
31.5	-0.43452	13.69738	0.109708	-0.93588	13.19181	0.114056
32.5	-0.47219	13.83505	0.110073	-0.95572	13.34023	0.115115
33.5	-0.51012	13.97418	0.110473	-0.97438	13.48913	0.116193

34.5	-0.54789	14.11503	0.110907	-0.99198	13.63877	0.117286
35.5	-0.58507	14.2578	0.111375	-1.00864	13.78937	0.118387
36.5	-0.62132	14.40263	0.111875	-1.02447	13.94108	0.119492
37.5	-0.6563	14.54965	0.112406	-1.03957	14.09407	0.120596
38.5	-0.68974	14.69893	0.112967	-1.05404	14.24844	0.121695
39.5	-0.72141	14.85054	0.113558	-1.06795	14.40429	0.122785
40.5	-0.75118	15.00449	0.114177	-1.08137	14.56168	0.123863
41.5	-0.7789	15.16078	0.114822	-1.09438	14.72064	0.124927
42.5	-0.80452	15.3194	0.115493	-1.10702	14.88121	0.125973
43.5	-0.828	15.4803	0.116188	-1.11934	15.04341	0.127
44.5	-0.84938	15.64343	0.116904	-1.13137	15.20721	0.128006
45.5	-0.8687	15.80873	0.117641	-1.14314	15.37263	0.12899
46.5	-0.88603	15.9761	0.118397	-1.15466	15.53962	0.129951
47.5	-0.90151	16.14548	0.119169	-1.16596	15.70817	0.130889
48.5	-0.91524	16.31677	0.119955	-1.17703	15.87824	0.131802
49.5	-0.92738	16.48986	0.120755	-1.18787	16.04978	0.132692
50.5	-0.93807	16.66468	0.121565	-1.19848	16.22277	0.133559
51.5	-0.94748	16.8411	0.122385	-1.20885	16.39715	0.134403
52.5	-0.95577	17.01904	0.123212	-1.21897	16.57289	0.135226
53.5	-0.9631	17.19839	0.124044	-1.2288	16.74994	0.136028
54.5	-0.96963	17.37906	0.124879	-1.23833	16.92827	0.136811
55.5	-0.97553	17.56096	0.125716	-1.24754	17.10783	0.137576
56.5	-0.98094	17.744	0.126554	-1.25639	17.28859	0.138324
57.5	-0.98601	17.92809	0.12739	-1.26486	17.47052	0.139058
58.5	-0.99087	18.11316	0.128224	-1.27293	17.65361	0.139779
59.5	-0.99564	18.29912	0.129054	-1.28055	17.83782	0.14049
60.5	-1.00045	18.48592	0.129879	-1.28769	18.02314	0.141191
61.5	-1.0054	18.6735	0.130698	-1.29433	18.20956	0.141885
62.5	-1.01058	18.8618	0.13151	-1.30044	18.39709	0.142574
63.5	-1.01606	19.05077	0.132315	-1.30599	18.58571	0.14326
64.5	-1.02193	19.24037	0.133111	-1.31095	18.77545	0.143944
65.5	-1.02824	19.43058	0.133898	-1.31529	18.96631	0.144629
66.5	-1.03504	19.62136	0.134676	-1.31899	19.15831	0.145317
67.5	-1.04237	19.8127	0.135444	-1.32204	19.35149	0.146009
68.5	-1.05025	20.00459	0.136203	-1.3244	19.54588	0.146707
69.5	-1.05871	20.19703	0.136952	-1.32606	19.74151	0.147412
70.5	-1.06773	20.39002	0.137691	-1.32702	19.93843	0.148127
71.5	-1.07732	20.58357	0.138422	-1.32726	20.1367	0.148852
72.5	-1.08747	20.7777	0.139143	-1.32676	20.33636	0.14959
73.5	-1.09815	20.97243	0.139855	-1.32554	20.53748	0.15034
74.5	-1.10933	21.16779	0.14056	-1.32358	20.74013	0.151105
75.5	-1.12097	21.36383	0.141256	-1.32089	20.94438	0.151885
76.5	-1.13302	21.56058	0.141947	-1.31747	21.1503	0.152682
77.5	-1.14543	21.75811	0.142631	-1.31333	21.35797	0.153495
78.5	-1.15813	21.95645	0.14331	-1.30849	21.56748	0.154326
79.5	-1.17106	22.15567	0.143985	-1.30295	21.77891	0.155174
80.5	-1.18414	22.35584	0.144657	-1.29673	21.99235	0.156041
81.5	-1.19731	22.55702	0.145327	-1.28986	22.20789	0.156927
82.5	-1.21048	22.7593	0.145996	-1.28236	22.42562	0.157831
83.5	-1.22357	22.96273	0.146666	-1.27424	22.64564	0.158753
84.5	-1.2365	23.16742	0.147337	-1.26555	22.86804	0.159693

85.5	-1.24919	23.37343	0.148012	-1.2563	23.09293	0.160651
86.5	-1.26156	23.58086	0.14869	-1.24653	23.32039	0.161627
87.5	-1.27352	23.78979	0.149374	-1.23627	23.55052	0.162619
88.5	-1.28501	24.00031	0.150065	-1.22555	23.78342	0.163628
89.5	-1.29595	24.21251	0.150764	-1.21441	24.01918	0.164651
90.5	-1.30627	24.42648	0.151472	-1.20288	24.25789	0.165689
91.5	-1.3159	24.64231	0.15219	-1.19101	24.49965	0.16674
92.5	-1.32478	24.8601	0.15292	-1.17882	24.74454	0.167802
93.5	-1.33286	25.07992	0.153663	-1.16635	24.99264	0.168876
94.5	-1.34008	25.30189	0.154419	-1.15365	25.24403	0.169959
95.5	-1.34641	25.52607	0.155189	-1.14075	25.4988	0.17105
96.5	-1.35181	25.75257	0.155974	-1.12768	25.75702	0.172147
97.5	-1.35625	25.98146	0.156775	-1.11449	26.01874	0.173249
98.5	-1.35971	26.21284	0.157592	-1.1012	26.28404	0.174355
99.5	-1.36217	26.44679	0.158425	-1.08786	26.55298	0.175462
100.5	-1.36361	26.68339	0.159275	-1.0745	26.82559	0.176568
101.5	-1.36404	26.92273	0.160142	-1.06115	27.10193	0.177673
102.5	-1.36346	27.16489	0.161026	-1.04785	27.38203	0.178774
103.5	-1.36187	27.40995	0.161926	-1.03462	27.66593	0.17987
104.5	-1.35928	27.65797	0.162842	-1.0215	27.95365	0.180958
105.5	-1.35572	27.90904	0.163775	-1.00852	28.24521	0.182037
106.5	-1.3512	28.16324	0.164722	-0.99571	28.5406	0.183105
107.5	-1.34575	28.42064	0.165684	-0.98309	28.83984	0.18416
108.5	-1.33941	28.6813	0.166659	-0.97069	29.14291	0.185201
109.5	-1.33219	28.9453	0.167647	-0.95853	29.4498	0.186225
110.5	-1.32414	29.21271	0.168646	-0.94664	29.76048	0.187231
111.5	-1.31529	29.48359	0.169655	-0.93504	30.07493	0.188218
112.5	-1.30569	29.758	0.170673	-0.92376	30.39308	0.189183
113.5	-1.29537	30.03602	0.171698	-0.9128	30.7149	0.190124
114.5	-1.28437	30.3177	0.172729	-0.9022	31.04032	0.191041
115.5	-1.27275	30.60311	0.173763	-0.89196	31.36928	0.191932
116.5	-1.26054	30.8923	0.174799	-0.88211	31.70168	0.192796
117.5	-1.24778	31.18533	0.175836	-0.87266	32.03745	0.19363
118.5	-1.23453	31.48225	0.176871	-0.86363	32.37649	0.194434
119.5	-1.22082	31.78312	0.177903	-0.85503	32.71868	0.195207
120.5	-1.20669	32.08799	0.178929	-0.84687	33.06392	0.195947
121.5	-1.19219	32.3969	0.179947	-0.83917	33.41208	0.196653
122.5	-1.17736	32.70991	0.180955	-0.83193	33.76303	0.197325
123.5	-1.16224	33.02704	0.181951	-0.82518	34.11663	0.197961
124.5	-1.14688	33.34835	0.182934	-0.81891	34.47272	0.198561
125.5	-1.1313	33.67387	0.183899	-0.81314	34.83116	0.199123
126.5	-1.11554	34.00363	0.184847	-0.80787	35.19176	0.199648
127.5	-1.09965	34.33766	0.185774	-0.80312	35.55437	0.200134
128.5	-1.08365	34.67599	0.186678	-0.7989	35.9188	0.200581
129.5	-1.06759	35.01864	0.187558	-0.7952	36.28486	0.200988
130.5	-1.05148	35.36562	0.188411	-0.79205	36.65236	0.201356
131.5	-1.03537	35.71695	0.189236	-0.78944	37.02111	0.201684
132.5	-1.01928	36.07263	0.19003	-0.78737	37.39089	0.201971
133.5	-1.00324	36.43266	0.190791	-0.78587	37.76149	0.202218
134.5	-0.98727	36.79704	0.191518	-0.78493	38.1327	0.202425
135.5	-0.97141	37.16577	0.19221	-0.78456	38.5043	0.202591

136.5	-0.95567	37.53881	0.192864	-0.78476	38.87605	0.202717
137.5	-0.94008	37.91616	0.193479	-0.78554	39.24775	0.202803
138.5	-0.92467	38.29777	0.194053	-0.7869	39.61914	0.202848
139.5	-0.90945	38.68361	0.194586	-0.78886	39.99	0.202854
140.5	-0.89445	39.07364	0.195077	-0.7914	40.36009	0.20282
141.5	-0.87968	39.46781	0.195523	-0.79455	40.72918	0.202747
142.5	-0.86516	39.86604	0.195925	-0.79829	41.09701	0.202636
143.5	-0.85092	40.26828	0.196281	-0.80264	41.46336	0.202486
144.5	-0.83696	40.67444	0.196592	-0.8076	41.82798	0.202299
145.5	-0.82332	41.08443	0.196855	-0.81317	42.19063	0.202074
146.5	-0.80999	41.49817	0.197072	-0.81936	42.55108	0.201814
147.5	-0.79701	41.91555	0.197241	-0.82616	42.90909	0.201517
148.5	-0.78439	42.33644	0.197363	-0.83359	43.26442	0.201185
149.5	-0.77214	42.76073	0.197437	-0.84163	43.61683	0.200819
150.5	-0.76027	43.18828	0.197464	-0.85031	43.96612	0.200419
151.5	-0.74882	43.61896	0.197445	-0.85961	44.31204	0.199987
152.5	-0.73778	44.05259	0.197378	-0.86953	44.65437	0.199522
153.5	-0.72718	44.48903	0.197266	-0.88009	44.99291	0.199027
154.5	-0.71704	44.92809	0.197109	-0.89127	45.32745	0.198501
155.5	-0.70736	45.3696	0.196907	-0.90308	45.65777	0.197946
156.5	-0.69817	45.81336	0.196662	-0.91551	45.98369	0.197363
157.5	-0.68948	46.25917	0.196375	-0.92857	46.30501	0.196753
158.5	-0.6813	46.70681	0.196046	-0.94225	46.62155	0.196116
159.5	-0.67367	47.15606	0.195677	-0.95654	46.93314	0.195455
160.5	-0.66659	47.60669	0.195269	-0.97144	47.23962	0.194769
161.5	-0.66007	48.05847	0.194825	-0.98695	47.54083	0.194061
162.5	-0.65414	48.51113	0.194344	-1.00305	47.83661	0.19333
163.5	-0.64882	48.96443	0.19383	-1.01974	48.12685	0.19258
164.5	-0.64412	49.4181	0.193283	-1.03701	48.41141	0.191809
165.5	-0.64006	49.87187	0.192706	-1.05485	48.69018	0.191021
166.5	-0.63665	50.32546	0.1921	-1.07323	48.96305	0.190216
167.5	-0.63392	50.77859	0.191467	-1.09216	49.22993	0.189395
168.5	-0.63188	51.23096	0.190808	-1.11161	49.49075	0.18856
169.5	-0.63054	51.68229	0.190127	-1.13155	49.74544	0.187712
170.5	-0.62992	52.13226	0.189425	-1.15198	49.99394	0.186852
171.5	-0.63004	52.58059	0.188703	-1.17287	50.23621	0.185983
172.5	-0.63091	53.02696	0.187964	-1.19418	50.47222	0.185104
173.5	-0.63253	53.47107	0.187209	-1.21591	50.70196	0.184219
174.5	-0.63492	53.91261	0.186442	-1.23801	50.92541	0.183328
175.5	-0.63808	54.35128	0.185663	-1.26045	51.14259	0.182432
176.5	-0.64203	54.78677	0.184874	-1.28319	51.35353	0.181534
177.5	-0.64676	55.21878	0.184079	-1.30621	51.55825	0.180635
178.5	-0.65226	55.64701	0.183277	-1.32946	51.75681	0.179736
179.5	-0.65855	56.07116	0.182472	-1.3529	51.94926	0.17884
180.5	-0.66561	56.49096	0.181666	-1.37648	52.13568	0.177947
181.5	-0.67343	56.90611	0.180859	-1.40015	52.31616	0.177059
182.5	-0.68199	57.31634	0.180054	-1.42388	52.4908	0.176179
183.5	-0.69127	57.72139	0.179253	-1.44759	52.6597	0.175307
184.5	-0.70126	58.121	0.178457	-1.47125	52.82299	0.174446
185.5	-0.71192	58.51492	0.177668	-1.49479	52.98079	0.173597
186.5	-0.72322	58.90293	0.176887	-1.51816	53.13327	0.172761

187.5	-0.73512	59.2848	0.176116	-1.54129	53.28056	0.171941
188.5	-0.74758	59.66033	0.175357	-1.56412	53.42284	0.171137
189.5	-0.76055	60.02932	0.17461	-1.5866	53.56028	0.170352
190.5	-0.77398	60.39159	0.173877	-1.60866	53.69307	0.169588
191.5	-0.78782	60.74699	0.17316	-1.63023	53.82138	0.168844
192.5	-0.80199	61.09537	0.172459	-1.65125	53.94544	0.168125
193.5	-0.81645	61.4366	0.171776	-1.67165	54.06543	0.167429
194.5	-0.83111	61.77057	0.171111	-1.69138	54.18158	0.16676
195.5	-0.84591	62.09719	0.170466	-1.71036	54.29411	0.166118
196.5	-0.86079	62.41639	0.169841	-1.72854	54.40324	0.165504
197.5	-0.87565	62.72809	0.169237	-1.74586	54.50921	0.164921
198.5	-0.89044	63.03228	0.168655	-1.76224	54.61224	0.164368
199.5	-0.90506	63.32892	0.168095	-1.77764	54.71257	0.163847
200.5	-0.91946	63.61802	0.167558	-1.79201	54.81044	0.163359
201.5	-0.93354	63.89959	0.167044	-1.80528	54.9061	0.162905
202.5	-0.94725	64.17367	0.166553	-1.81742	54.99978	0.162486
203.5	-0.96051	64.44032	0.166085	-1.82837	55.09172	0.162101
204.5	-0.97324	64.69961	0.16564	-1.83809	55.18217	0.161753
205.5	-0.9854	64.95165	0.165218	-1.84655	55.27135	0.16144
206.5	-0.9969	65.19653	0.164819	-1.85372	55.35951	0.161164
207.5	-1.00771	65.4344	0.164442	-1.85957	55.44686	0.160924
208.5	-1.01776	65.6654	0.164087	-1.86407	55.53362	0.160721
209.5	-1.027	65.8897	0.163753	-1.86721	55.62001	0.160554
210.5	-1.0354	66.10749	0.163439	-1.86898	55.70624	0.160423
211.5	-1.04292	66.31897	0.163144	-1.86937	55.79248	0.160329
212.5	-1.04951	66.52437	0.162867	-1.86839	55.87892	0.160269
213.5	-1.05516	66.7239	0.162608	-1.86603	55.96573	0.160245
214.5	-1.05984	66.91784	0.162365	-1.86233	56.05305	0.160254
215.5	-1.06353	67.10642	0.162137	-1.85729	56.141	0.160296
216.5	-1.06622	67.28993	0.161923	-1.85095	56.2297	0.16037
217.5	-1.06791	67.46863	0.161721	-1.84333	56.31922	0.160474
218.5	-1.06859	67.64281	0.161532	-1.8345	56.40963	0.160607
219.5	-1.06826	67.81277	0.161352	-1.82448	56.50096	0.160768
220.5	-1.06693	67.97877	0.161183	-1.81334	56.5932	0.160955
221.5	-1.06462	68.14111	0.161022	-1.80115	56.68633	0.161166
222.5	-1.06134	68.30005	0.16087	-1.78798	56.78026	0.161399
223.5	-1.05712	68.45585	0.160726	-1.7739	56.8749	0.161652
224.5	-1.05199	68.60872	0.16059	-1.75901	56.9701	0.161923
225.5	-1.04599	68.75889	0.160462	-1.74339	57.06565	0.162209
226.5	-1.03917	68.90653	0.160343	-1.72716	57.16132	0.162509
227.5	-1.03158	69.05176	0.160234	-1.71041	57.2568	0.162819
228.5	-1.02329	69.19467	0.160138	-1.69327	57.35176	0.163138
229.5	-1.01439	69.33527	0.160056	-1.67585	57.44578	0.163463
230.5	-1.00495	69.47351	0.159992	-1.6583	57.5384	0.163791
231.5	-0.9951	69.60926	0.15995	-1.64075	57.6291	0.16412
232.5	-0.98496	69.74228	0.159934	-1.62333	57.71728	0.164447
233.5	-0.97466	69.87224	0.159951	-1.60621	57.80227	0.164771
234.5	-0.96438	69.99869	0.160007	-1.58953	57.88334	0.165088
235.5	-0.95427	70.12104	0.160112	-1.57347	57.95967	0.165398
236.5	-0.94455	70.23857	0.160274	-1.55818	58.0304	0.165698
237.5	-0.93541	70.3504	0.160505	-1.54385	58.09453	0.165985

238.5	-0.92706	70.45546	0.160819	-1.53064	58.15104	0.16626
239.5	-0.91972	70.55252	0.16123	-1.51875	58.19877	0.16652
240	-0.91649	70.59761	0.161477	-1.51336	58.21897	0.166645

Appendix 2: Body Weight Percentile Data for Adults

This appendix includes actual values of raw percentile in Tables A2.1 and A2.2 for women and men, respectively. The values of the smoothed percentiles are provided in Tables A2.3 and A2.4. For details on the source of data and the smoothing process see Part 2 of this report, Section 7.1 and 7.3, respectively.

Table A2.1: Empirical percentiles of body weight (kg) for adult women

Combined CSFII and NHANES III Data

Age	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	0.99
20	44	47	49	50	50	52	53	55	55	58	60	61	62	64	64	68	70	73	81	92	116
21	40	47	48	50	51	52	54	56	57	58	59	61	64	66	67	73	77	83	90	95	116
22-24	42	47	50	51	53	56	57	58	59	60	61	63	65	66	68	71	75	78	82	99	121
25-29	42	47	50	52	54	54	56	57	59	60	61	63	65	67	70	74	77	81	87	100	117
30-34	43	48	51	52	54	55	57	59	61	62	64	66	68	70	73	77	80	84	92	103	128
35-39	44	49	52	54	56	58	59	61	62	64	66	68	70	72	75	79	84	89	94	106	125
40-44	43	50	53	55	57	59	60	62	64	65	67	69	72	74	76	79	82	87	92	104	134
45-49	44	50	54	56	58	59	62	64	65	67	69	70	73	75	78	81	84	89	95	104	122
50-54	44	52	54	57	59	61	63	65	66	68	70	73	75	77	80	83	87	91	96	109	132
55-59	46	51	55	57	59	61	64	65	67	68	70	72	75	77	79	82	86	90	93	103	124
60-61	43	50	53	55	59	60	62	64	66	68	68	70	73	74	76	80	83	86	93	100	115
62-64	44	50	53	58	60	61	63	64	66	68	69	71	73	76	77	80	83	87	93	100	121
65-69	43	50	52	54	57	59	60	62	64	66	68	69	72	74	75	77	80	84	88	95	116
70-74	43	49	52	54	57	59	60	61	63	64	66	68	70	73	75	77	80	84	89	99	117
75-79	37	44	49	52	54	56	58	59	61	63	65	67	68	70	73	74	76	79	83	88	108
80-84	36	43	47	48	52	54	56	57	59	60	62	64	65	66	68	70	73	75	77	84	97
≥ 85	33	41	45	47	49	51	52	53	54	55	57	59	60	61	64	65	68	70	72	78	94

The age categories are listed down the leftmost column and the percentiles are listed across the top of the table. To read the table, start at a selected age category then scan across the table to the column under the percentile of interest. For example, the 50th percentile bodyweight based on the empirical cumulative distribution for 22-to-24 year-old females is 61 kg.

Table A2.2: Empirical percentiles of body weight (kg) for adult men

Combined CSFII and NHANES III Data

Age	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	0.99
20	51	56	59	62	63	64	66	67	68	68	70	74	77	79	80	82	84	88	95	102	138
21	50	54	58	62	64	66	68	70	70	71	74	74	75	77	79	82	84	86	93	104	120
22-24	54	58	61	64	66	67	69	71	73	75	77	79	81	82	84	86	89	92	100	106	127
25-29	54	60	63	66	68	70	72	73	75	76	77	79	82	84	85	88	92	95	102	112	134
30-34	54	62	64	66	68	72	74	75	77	79	80	82	84	86	88	91	93	98	102	109	132
35-39	55	62	66	69	71	73	74	76	78	79	82	83	85	87	89	92	96	100	104	113	150
40-44	54	62	66	69	73	75	77	79	81	82	84	86	88	90	93	95	99	104	107	118	136
45-49	58	63	67	70	72	75	76	78	79	80	82	83	84	86	89	92	95	99	103	114	141
50-54	56	65	68	72	75	77	79	81	82	84	85	87	89	91	92	95	97	102	105	113	142
55-59	55	61	67	71	73	75	77	79	80	82	84	86	87	89	91	94	97	99	105	114	128
60-61	52	61	65	68	70	73	75	77	79	81	82	84	86	88	90	93	94	97	101	107	121
62-64	54	60	64	69	72	75	76	79	79	82	83	84	86	89	91	93	95	100	104	111	125
65-69	53	62	65	68	71	73	75	77	78	80	82	84	86	89	90	92	95	98	102	108	124
70-74	52	59	63	66	68	70	72	74	75	77	79	82	83	84	86	88	91	95	99	104	120
75-79	52	59	62	64	66	68	70	72	73	75	76	78	80	82	83	85	86	91	95	100	113
80-84	46	55	59	62	64	66	67	68	70	71	73	75	76	77	79	81	84	86	93	99	111
≥ 85	44	50	55	57	59	61	64	65	67	69	71	72	73	74	75	77	79	81	84	89	101

The age categories are listed down the leftmost column and the percentiles are listed across the top of the table. To read the table, start at a selected age category then scan across the table to the column under the percentile of interest. For example, the 50th percentile bodyweight based on the empirical cumulative distribution for 22-to-24 year-old males is 77 kg.

Table A2.3: The smoothed percentiles of body weight (kg) for adult women.

Data is from the combined CSFII and NHANES III Data.
The smoothing function was $y=a+bx^2+cx^{2.5}$.

Age	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	0.99
20	42	46	48	50	51	52	54	55	57	58	60	61	63	65	67	70	74	78	84	96	118
21	42	46	49	50	51	53	54	56	57	58	60	61	63	65	67	71	74	79	85	97	118
22-24	42	47	49	51	52	53	55	56	58	59	61	62	64	66	68	72	75	80	86	97	119
25-29	43	48	50	52	53	55	56	58	59	60	62	64	66	68	70	74	77	81	88	99	121
30-34	43	48	51	53	55	56	58	59	61	62	64	66	68	70	72	76	79	84	90	101	123
35-39	44	49	52	54	56	58	59	61	62	64	66	68	70	72	74	78	81	86	92	103	125
40-44	45	50	53	55	57	59	61	62	64	66	67	69	71	74	76	79	83	87	93	104	126
45-49	45	51	54	56	58	60	62	63	65	67	68	70	73	75	77	81	84	88	94	105	126
50-54	45	51	54	57	59	60	62	64	66	68	69	71	74	76	78	81	85	89	95	105	126
55-59	44	51	54	57	59	61	62	64	66	68	70	72	74	76	78	81	85	89	94	104	125
60-61	44	50	54	56	59	60	62	64	66	68	69	71	74	76	78	81	84	88	94	103	123
62-64	43	50	53	56	59	60	62	64	66	67	69	71	73	76	78	81	84	88	93	102	122
65-69	42	49	53	55	58	60	62	63	65	67	68	70	72	75	77	79	83	86	91	99	119
70-74	41	48	51	54	56	58	60	62	63	65	67	69	71	73	75	77	80	83	88	95	114
75-79	39	46	50	52	55	56	58	60	61	63	64	66	68	70	72	74	77	80	83	90	108
80-84	36	44	47	49	52	54	55	57	58	60	61	63	64	66	68	70	72	75	78	84	100
≥ 85	33	41	44	46	48	50	52	53	54	56	57	59	60	61	63	64	66	69	71	76	91

The age categories are listed down the leftmost column and the percentiles are listed across the top of the table. To read the table, start at a selected age category then scan across the table to the column under the percentile of interest. For example, the 50th percentile bodyweight for 22-to-24 year-old females is 61 kg.

Table A2.3: The smoothed percentiles of body weight (kg) for adult men.

Data is from the combined CSFII and NHANES III Data.
The smoothing function was $y=a+bx^3+c/x^{0.5}$.

Age	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	0.99
20	51	56	59	62	63	64	66	67	68	68	70	74	77	79	80	82	84	88	95	102	138
21	50	54	58	62	64	66	68	70	70	71	74	74	75	77	79	82	84	86	93	104	120
22-24	54	58	61	64	66	67	69	71	73	75	77	79	81	82	84	86	89	92	100	106	127
25-29	54	60	63	66	68	70	72	73	75	76	77	79	82	84	85	88	92	95	102	112	134
30-34	54	62	64	66	68	72	74	75	77	79	80	82	84	86	88	91	93	98	102	109	132
35-39	55	62	66	69	71	73	74	76	78	79	82	83	85	87	89	92	96	100	104	113	150
40-44	54	62	66	69	73	75	77	79	81	82	84	86	88	90	93	95	99	104	107	118	136
45-49	58	63	67	70	72	75	76	78	79	80	82	83	84	86	89	92	95	99	103	114	141
50-54	56	65	68	72	75	77	79	81	82	84	85	87	89	91	92	95	97	102	105	113	142
55-59	55	61	67	71	73	75	77	79	80	82	84	86	87	89	91	94	97	99	105	114	128
60-61	52	61	65	68	70	73	75	77	79	81	82	84	86	88	90	93	94	97	101	107	121
62-64	54	60	64	69	72	75	76	79	79	82	83	84	86	89	91	93	95	100	104	111	125
65-69	53	62	65	68	71	73	75	77	78	80	82	84	86	89	90	92	95	98	102	108	124
70-74	52	59	63	66	68	70	72	74	75	77	79	82	83	84	86	88	91	95	99	104	120
75-79	52	59	62	64	66	68	70	72	73	75	76	78	80	82	83	85	86	91	95	100	113
80-84	46	55	59	62	64	66	67	68	70	71	73	75	76	77	79	81	84	86	93	99	111
≥ 85	44	50	55	57	59	61	64	65	67	69	71	72	73	74	75	77	79	81	84	89	101

The age categories are listed down the leftmost column and the percentiles are listed across the top of the table. To read the table, start at a selected age category then scan across the table to the column under the percentile of interest. For example, the 50th percentile bodyweight for 22-to-24 year-old males is 77 kg.