Risk Analysis, Vol. 21, No. 1, 2001

Disparity in Quantitative Risk Assessment: A Review of Input Distributions

Bruce S. Binkowitz[1]* and Daniel Wartenberg[1]

Monte Carlo simulations are commonplace in quantitative risk assessments (QRAs). Designed to propagate the variability and uncertainty associated with each individual exposure input parameter in a quantitative risk assessment, Monte Carlo methods statistically combine the individual parameter distributions to yield a single, overall distribution. Critical to such an assessment is the representativeness of each individual input distribution. The authors performed a literature review to collect and compare the distributions used in published QRAs for the parameters of body weight, food consumption, soil ingestion rates, breathing rates, and fluid intake. To provide a basis for comparison, all estimated exposure parameter distributions were evaluated with respect to four properties: consistency, accuracy, precision, and specificity. The results varied depending on the exposure parameter. Even where extensive, well-collected data exist, investigators used a variety of different distributional shapes to approximate these data. Where such data do not exist, investigators have collected their own data, often leading to substantial disparity in parameter estimates and subsequent choice of distribution. The present findings indicate that more attention must be paid to the data underlying these distributional choices. More emphasis should be placed on sensitivity analyses, quantifying the impact of assumptions, and on discussion of sources of variation as part of the presentation of any risk assessment results. If such practices and disclosures are followed, it is believed that Monte Carlo simulations can greatly enhance the accuracy and appropriateness of specific risk assessments. Without such disclosures, researchers will be increasing the size of the risk assessment "black box," a concern already raised by many critics of more traditional risk assessments.

KEY WORDS: Risk assessment; Monte Carlo analysis; exposure factors; distributions

1. INTRODUCTION

Monte Carlo simulations are becoming commonplace in quantitative risk assessments (QRAs). Designed to propagate the variation associated with each individual exposure input parameter in a QRA, Monte Carlo methods statistically combine the individual parameter distributions to yield a single, overall distribution. This distribution is often interpreted as the overall variation of the predicted risk. Critical to such an assessment are the accuracy, precision, and representativeness of these individual distributions. In this article, the distributions used in published QRAs are reviewed and compared.

[1] Environmental and Occupational Health Sciences Institute, Piscataway, NJ.

*Address correspondence to Bruce S. Binkowitz, 11 Tamaques Way, Westfield, NJ 07090; binkowitz@msn.com.

2. THE MOTIVATION FOR USING MONTE CARLO SIMULATIONS IN QUANTITATIVE RISK ASSESSMENT

The established protocol for conducting a QRA is a well-known four-step process. The analyst identifies hazardous materials, quantifies exposure, defines

0272-4332/01/0200-0075$16.00/1

a dose–response relationship, and calculates the cumulative risk for each hazard at the modeled exposure given the dose–response function.[1,2] Traditionally, parameter estimates are derived to represent each of the inputs to the QRA equation. Such estimates often take the form of an arithmetic average, a geometric mean, or a percentile value such as the 95th or the median. The risk analyst chooses which values, such as percentiles or moments of the distribution to use, most often without rigorous justification. These values, or point estimates, frequently are chosen from the tails of the assumed distributions to be "conservative" or "public health protective." When combined algebraically, these values give a single, incremental cancer risk for a lifetime of exposure (ILCR). Typically, only the risk value is reported to policy analysts, decision makers, and the public, without explicit description of the limitations and assumptions of the methodology or the rationale for the cutpoints used. Most risk analysts recognize the limitations and conservatism of their estimates, although they rarely provide estimates of statistical confidence limits or estimation precision. Even so, policy analysts, decision makers, and the public often are appropriately skeptical of the accuracy of final risk estimates, using them most frequently in a relative or priority-setting framework.

As a more rigorous alternative, various authors have suggested using Monte Carlo simulations in QRAs.[3–8] They believe that this approach can provide insight into the distribution underlying the ILCR and provide a useful display. Monte Carlo simulation is a statistical method used to derive a distribution for the algebraic combination of a set of parameters, each of which is treated as a random variable and has its own specified distribution. By repeatedly choosing a set of input values from a prespecified statistical distribution for each variable and combining them into a summary index such as the ILCR, investigators derive an empirical distribution for the summary index, possibly accounting for the dependence among the inputs. Investigators use such methods frequently and in widespread applications. The U.S. Environmental Protection Agency (USEPA) Risk Assessment Forum has published extensively both policy and technical documents regarding probabilistic methods.

When applied to QRA, analysts argue, Monte Carlo simulations can provide an empirical and graphical representation of the degree of variability of an assessment, and can, as a basis of policy considerations, provide risk managers with sufficient data to allow them to select the most prudent quantile of risk.[3–8] By using an empirical distribution of the estimated ILCR, some of these investigators assert that the calculations provide a sensitivity analysis of the full range of possible outcomes that could occur if the modeled exposures were to take place. Thus, they claim that one of the uses of these calculations is to enable risk managers to quantify the degree of conservatism or "worst-caseness" employed in their final decision. Underlying the validity of these estimates are the accuracy, precision, and representativeness of the exposure parameter distributions input into the Monte Carlo simulation. Rarely does any author check the risk assessment to see how robust the results would be to changes in assumptions regarding distributions for model inputs. Further, while it is recognized that there is a fundamental difference between uncertainty and variability,[9] it should also be recognized how rarely practitioners address these differences. Specific discussions of model uncertainty, input uncertainty, and variability along with an assessment (e.g., two-dimensional methods) of their impact on results would help clarify the disparity seen in input distributions. Cullen and Frey[9] detail many methods to not only identify the causes of any disparity, but to also estimate the potential impact of each cause. The current practice in QRA with respect to these distributions is reviewed below.

3. METHODS

To identify specific distributions cited for exposure parameters used in Monte Carlo simulations of QRAs, all articles in the *Journal of Exposure Analysis and Environmental Epidemiology* and *Risk Analysis* from 1990–1996, as well as the distributions cited in the American Industrial Health Council (AIHC) *Exposure Factors Sourcebook* and the USEPA's *Exposure Factors Handbook*[10,11] were reviewed. Relevant citations found in each of these sources were followed back and are included in Tables I and II. Nearly half the articles come from *Risk Analysis*, with others scattered about the literature. All distributions used in this article were then identified. A similar proportion of distributions come from *Risk Analysis* and the USEPA's *Exposure Factors Handbook*.[11] Other venues play a relatively small role, representing less than one third of all distributions. Over time, there was a slight increase in publication of articles in the early 1990s, with the largest pulse of distributions coming from the release of the USEPA's *Exposure Factors Handbook* in 1995.[11]

Table I. Source of Distributions Used in Published Quantitative Risk Assessments

Reference source	Frequency by articles	Frequency by distributions used
AIHC Exposure Handbook	1	5
Archives of Environmental Health	1	3
Epidemiology	1	1
GCA Report to USEPA	1	1
Human and Ecological Risk Assessment	1	2
International Archives of Occupational and Environmental Health	1	4
Journal of Exposure Analysis and Environmental Epidemiology	2	2
Regulatory Toxicology and Pharmacology	1	3
Risk Analysis	11	30
Textbook: *Petroleum Contaminated Soils*	1	4
Textbook: *Research Methods in Occupational Epidemiology*	1	1
USEPA *Exposure Factors Handbook*	1	35

Table II. Temporal Patterns of Distributions/Articles Published

Year	Frequency of published articles	Frequency by distributions used
1984	1	1
1986	1	3
1987	2	8
1988	1	2
1989	2	5
1991	1	2
1992	5	9
1993	2	2
1994	5	11
1995	1	35
1996	1	3

From the relevant articles, information about the distributions was summarized and those parameters used in Monte Carlo simulations in at least five articles were selected. These parameters are body weight, food consumption, soil ingestion rates, breathing rates, and fluid intake. Detailed information about the distributions used to estimate these parameters is provided in Tables III–VII. These tables are intended only as references to this article. The reader is encouraged to go back to the cited sources and not rely on the tables of this article as the source of inputs to a QRA. Only the cited sources contain the subtleties investigators should be aware of when preparing a QRA.

To provide a basis of comparison, all estimated exposure parameter distributions were evaluated with respect to four properties: consistency, accuracy, precision, and specificity. These four properties and the list of specific distributions used for each parameter are tabulated in Table VIII.

To assess consistency, the present review includes whether different authors use the same parameter estimates or the same distributions for each cited exposure variable. If the data are not consistent, justification is sought for the data used by each investigator.

For accuracy, the source of the data is noted, recognizing that the underlying data may be disputed within the specific field of research from which it was obtained (e.g., soil ingestion rates). If the data are not accurate, then they are not representative of the data in the specific field of research, and the simulation results will not be able to be interpreted meaningfully.

Precision of parameter estimates or distributions is rarely mentioned. Introduction of the raw measurement data or the characteristics of their distribution will only give the correct result if the variance of the sample set is needed for the risk analysis. It is important that the correct statistics are used to characterize the parameters of interest. If the mean is of interest, the standard error of the mean and *not* the standard deviation of the distribution should be used. Further, if the tail of the distribution is of primary interest, there is rarely a discussion of the associated variability and uncertainty. Morgan and Henrion[12] provide a methodology that yields confidence intervals about a fractile of interest. Precision of an input is composed of the natural variability of a parameter, the sampling variance, and the number of samples taken to establish the estimate. Again, following Morgan and Henrion,[12] suppose determination of the 95th percentile is of interest. A sample of size 100 would give a 95% confidence interval extending from the 91st percentile to the 99th. A sample of size 25 being used to determine the 95th percentile yields a much wider 95% confidence interval, extending from the 84th percentile to the maximum value. Therefore, knowing the size of the population sample and the source of the data are essential in assessing the statistical properties of the distribution. For example, when using a white male body weight distribution with a mean value of 70 kg, rarely is mention made of how large a sample of individuals was used to obtain the parameters for the distributions. A standard devi-

Table III. Body Weight (kilograms) and Estimated Parameters

Subject	Distribution	*Mean*	*SD*	Min	5%tile	Median	95%tile	Max	Source	References and footnotes
Adult (male)	Lognormal	(77.8)	(29.4)		(58.0)	(76.7)	(101.4)		NHANES II [20]	Brainard and Burmaster[21][a], n = 9,983
Adult (male)	Empirical			51		75.9		107	NHANES II[20]	AIHC[10] p. 4.3[g]
Adult (male)	Empirical	78.1	13.5						NHANES II[20]	USEPA Handbook[11][b], n = 5916
Adult (female)	Lognormal	(66.0)	(13.3)		(46.1)	(64.7)	(89.0)		NHANES II[20]	Brainard and Burmaster[21][a], n = 10339
Adult (female)	Empirical			44		57.1		103	NHANES II[20]	AIHC 1994[10] p. 4.3[g]
Adult (female)	Empirical	65.4	14.6						NHANES II[20]	USEPA Handbook[11][b] n = 6,588
Adult (both)	Lognormal	58	14		(38.1)	(56.4)	(83.4)		ICRP[37] and USEPA EFH[39]	McKone and Bogen[22]
Adult (both)	Lognormal	71	(14.2)		(50.3)	(69.6)	(96.4)		Najjar and Roland[23]	McKone[24][c]
Adult (both)	Empirical			44		68.7		107	NHANES II[20]	AIHC[10] p. 4.3[g]
Adult (both, 20–59)	Lognormal	(70.8)	(14.3)		(49.9)	(69.4)	(96.4)		Stephens and Craig[25]	Richardson and Allen[26][h]
Senior (both, 60+)	Lognormal	(70.4)	(15.7)		(47.9)	(68.7)	(98.7)		Stephens and Craig[25]	Richardson and Allen[26][h]
Adult (Japanese)	Point est.	55.2							Kerr[27]	Bogen and McKone[28][d]
Adult (Japanese)	Point est.	50							Kerr[27]	Bogen and McKone[28][d]
Child 8–18	Normal	47	8.3		(33.3)	(47)	(60.7)		NHANES II[20]	GCA[29]
Child	Triangle	(15.9)	(4.0)	6.5	(9.4)	(15.7)	(22.8)	26.1	NHANES II[20]	Smith[42][f], gave "best" = 15
Child	Uniform	(33.75)	(7.2)	21.2	(22.5)	(33.75)	(45.0)	46.3	USEPA Methods for Assessing Exposure to Chem Substances, Vol. 2, 1985[62]	Finley & Paustenbach[30][e]
Child 0–15	Lognormal	29	(7.0)		(19.1)	(28.2)	(41.7)		Najjar and Roland[23]	McKone[24][c]
Toddler (3–4)	Lognormal	(18.0)	(2.7)		(13.9)	(17.8)	(22.8)		Stephens and Craig[25]	Richardson and Allen[26][h]
Child (5–11)	Lognormal	(27.8)	(7.3)		(17.5)	(26.8)	(41.2)		Stephens and Craig[25]	Richardson and Allen[26][h]
Teen (12–19)	Lognormal	(60.0)	(13.4)		(40.8)	(58.6)	(84.1)		Stephens and Craig[25]	Richardson and Allen[26][h]
Child (0.5–5)	Lognormal	15.8	2.0	10.6	12.7	15.7	19.4	22.8	NHANES II[20]	Oregon DEQ Guidance[31][i]
Juvenile (6–17)	Lognormal	44.1	10.2	21.6	29.5	42.9	62.6	86.5	NHANES II[20]	Oregon DEG Guidance[31][i]

[a] Brainard and Burmaster[21] fit a bivariate distribution, however, the marginal distribution for weight is used here. The arithmetic mean and standard deviation are estimated from the mean and standard deviation in log-lbs. The median is the geometric mean, converted to kg. The percentiles are computed on the log scale, and backtransformed.

[b] *Exposure Factors Handbook*.[11]

[c] Body weight data is from Najjar and Roland.[23]

[d] Cites the "anatomical model of Kerr[27] for reference Japanese adults." (p. 514). Kerr cites his own paper in 1976 where a mathematical model of a phantom with a total body mass of 55 kg was designed to approximate a present-day Japanese adult. The 50-kg estimate is for atomic bomb survivors 20 years or older in 1945.

[e] The reported values were converted from pounds to kilograms for consistency in this table. Source: USEPA *Methods for Assessing Exposure to Chemical Substances*, Vol. 2, 1985.[62]

[f] Smith[42] references USEPA *Exposure Factors Handbook* EPA/600/8-89/043.[39] Child body weight approximated by triangle distribution, but no justification given.

[g] AIHC[10] references the USEPA *Exposure Factors Handbook* EPA/600/8-89/043,[39] which references NHANES II.[20]

[h] The arithmetic mean and standard deviation are estimated from the mean and standard deviation in log-kg. The median is the geometric mean. The percentiles are computed on the log scale, and backtransformed.

[i] Oregon DEQ[31] assumes the NHANES II[20] data for 3-year-olds is representative of all children 6 months to 5 years. For juveniles ages 6 to 17 years, Oregon DEQ uses the NHANES II[20] data for 12-year-olds.

Table IV. Food Consumption and Estimated Parameters

Food source	Distribution	Units	*Mean*	*SD*	Min	5%tile	Median	95%tile	Max	Source	References and footnotes
Eggs***[a]	Empirical	g/day	26.9							USDA NFCS 1977–78[32]	USEPA Handbook[11][b]
Eggs: 3-day avg.	Empirical	g/day	47	33		15	40	109	728	Pao et al.[33]	USEPA Handbook[11][b]
Eggs: 1-week avg.	Empirical	g/kg-day	0.68 (47.6)[c]		0	0.0079 (0.55)[c]	0.045 (3.2)[c]	2.01 (140)[c]	18.8 (1,316)[c]	USDA NFCS 1987/88[34]	USEPA Handbook[11][b]
Total beef	Normal	g/day	75	56		(−17)	(75)	(167)		Finley & Paustenbach[35] analysis of USDA NFCS 1977–78[32]	AIHC[10] p. 6.31
Meat: 3-day avg.	Empirical	g/day	85	67		11	69	211	1,792	Pao et al.[33][d]	USEPA Handbook[11]
Meat: 1-week avg.	Empirical	g/kg-day	4.69 (328.3)[c]		0	0.67 (46.9)[c]	3.22 (225)[c]	13.2 (924)[c]	136 (9,520)[c]	USDA NFCS 1987/88[34][e]	USEPA Handbook[11], n = 9,685
Meat	Empirical	g/day	47							USDA NFCS 1987/88[34]	USEPA Handbook[11]
Meat (males)	Empirical	g/day	65							USDA NFCS 1987/88[34]	USEPA Handbook[11]
Meat (females)	Empirical	g/day	42							USDA NFCS 1987/88[34]	USEPA Handbook[11]

[a] Eggs** means the USDA *Nationwide Food Consumption Survey 1977–78*[32] data have been reorganized by the USEPA Office of Radiation Programs, and food items have been classified according to the characteristics of radionuclide transport.

[b] Egg estimates from USEPA Handbook[11] come from a survey of size 37,874 or less. The specific size included for each food is not given. The standard error of the mean is given as 0.5 g/day.

[c] These estimates are based on a 70-kg-body weight adult male.

[d] Pao *et al.*[33] as cited in USEPA[11] also gives the percent of individuals in the population who consumed across the 3-day period. Meat = beef and pork and lamb and veal.

[e] Meat: 1-week average represents intake rates for consumers of the food item/group for family members from those surveyed households who reported eating the food item/group. Data are given by age group, season, urbanization, race, and geographical region.

ation may be given when, in fact, it is the standard error of the mean value (the standard deviation divided by the square root of the sample size) that is of interest. The information to be input into a QRA is categorized as "precise" if the data leading to these inputs arise from a large data set; also checked is if the standard deviation of a distribution is generally less than half the mean. If there is naturally a lot of variation in the data, one will have to get many more samples before one is confident in the estimate of the QRA input. We also utilize Haas's[13] definition of the relative standard deviation to categorize data as "imprecise" if the standard deviation is generally as large as the mean when the sources of the distributions do not cite the sample size. It should be noted that any estimate of the dispersion of a parameter value will be less precise than the estimate of the parameter value itself.[14,15]

Population specificity is also considered for each exposure input parameter distribution. That is, some Monte Carlo analyses use a general distribution that includes all possible people exposed. While this may be an accurate reflection of the population, there also may be substantial heterogeneity within the population. If a particular part of that population is the focus of the risk of interest, this segment of the population should be explained. Most investigators are working toward a specified fractile, with little justification other than "policy" and little interpretation of who makes up that part of the population, for example, what are the unique characteristics of the upper 5%? Other risk assessors take a particular exposure parameter and estimate its distribution for each of a set of subpopulations. Examples include age (adults versus children), health status (working or not working), geography (rural versus urban), and type of intake (all fluids versus all water versus tap water). Also noted is that the distributions may vary by region of application—diets likely vary between New England, New Orleans, and California, for instance. This geographic specificity is not addressed in this article.

Finally, the consistency of the distributional shape is catalogued. Some authors fit a parametric distribution to a data set, while others used the empirical distribution itself, and still others cite a distribution from the literature without regard to its relevance for the population under study. Some investigators make arbitrary choices of distributions, such as using a uniform or triangular distribution for situations in which only limited knowledge or data are available. Previous research has shown that distributional shape can have a substantial influence on estimates of risk.[13,16,17] Only the range of distributions used is listed.

Table V. Breathing Rates (m^3/day) and Estimated Parameters

	Distribution	*Mean*	*SD*	Min	5%tile	Median	95%tile	Max	Source	References and footnotes
BR at home	Point est.	12							Bogen & McKone[28] plus T&E's[36] own assumptions	Thompson and Evans[36]
BR (unknown specific activity)	Point est.	13.3								USEPA Handbook[11]
BR	Point est.	10							ICRP[37]	Checkoway *et al.*[38]*a*
BR/unit BW (m^3/kg-day)*c*	Lognormal	0.40	0.50		(0.05)	(0.25)	(1.28)		ICRP[37]	McKone and Bogen[22]*b*
	Multiply by 70 kg	(28)	(35)		(3.5)	(17.5)	(89.6)			
Adult BR	Triangle	(15.1)	(3.1)	8	(10.2)	(14.9)	(20.6)	23.2	USEPA Handbook[39]	Stern[40] gives "best" = 14*d*
Lifetime daily rate calculated from food-energy intake	Point est.	14 (M) 10 (F)							Layton[41]	USEPA Handbook[11]*e*
In the shower	Triangle	(14.2)	(5.7)	2.304	(5.7)	(13.6)	(24.4)	29.376	Various sources, see footnote*c*	USEPA[39] as cited in Smith[42]*f* gives "best" as 11.0016
BR child (0.5–5y)	*g*	5	0.5	3.6	4.2	5	5.85	6.7	NHANES II[20], Stahl[43]	Oregon DEQ Guidance[31]*g*
BR juvenile (6–17)	*g*	11.2	2.1	6.4	8.2	11	14.9	19.4	NHANES II[20], Stahl[43]	Oregon DEQ Guidance[31]*g*
BR adult	*g*	17.7	2.4	11.57	14.07	17.6	22.0	26.8	NHANES II[20], Stahl[43]	Oregon DEQ Guidance[31]*g*

Note: BR = breathing rate; BW = body weight.

a The "average breathing rate of air is assumed to be a constant, 10 m^3/day," and is given as "typical for light activity" (p. 125).

b The median column contains the geometric mean.

c These estimates are based on a 70-kg-body weight adult male.

d From Stern[40]: "The EFH suggests an adult daily average inhalation volume of 14 m^3/day based on 67% light activity and 33% rest. The suggested worst case adult average inhalation volume of 23.2 m^3/day is based on 5% heavy activity, 10% moderate rest, 60% light activity, and 25% at rest. An adult minimum inhalation volume can be estimated as 8 m^3/day by assuming a female 100% at rest" (p. 460). Triangular distributions were employed when more limited probabilistic data existed and upper and lower percentile values had to be estimated.

e Layton[41] is from USEPA pp. 3–6, and also gives data by age range.

f Converted from meters cubed per minute. Smith[42] cites GCA Corp.[29] This citation states that the inhalation values are arrived utilizing many sources of data.

g Oregon DEQ[31] uses Stahl[43] allometric equation ($n = 691$, $r^2 = 0.98$) to account for the unspecified correlation between body weight and inhalation rate. The body weight distributions are as given in Table III. The distributions are simulated from the equation: breathing rate = $0.5458 \times$ body weight$^{0.80}$. The standard error of the mean from the simulations are given as 0.01 for 0.5–5, and 0.02 for 6–17 and adult.

4. RESULTS

4.1. Body Weight

Historically, one of the most commonly used parameters in QRA is body weight. Typically, adult (male) body weight is represented by a point estimate of 70 kg. Body weight is also a commonly recorded characteristic of populations. Usually, it is recorded with other population characteristics, enabling stratification across subgroups, such as by gender and age. One large database on body weight is the set of anthropometric measurements for the U.S. population that were collected as part of the second National Health and Nutrition Examination Survey (NHANES II).[20] This survey assessed a probability sample of approximately 28,000 persons, aged 6 months to 74 years from the civilian, noninstitutionalized population of the United States from 1976 through 1980. The sur-

vey was designed to oversample certain subgroups thought to be at high risk of malnutrition, such as the elderly, preschool children, and low-income families. The handling of this oversampling when deriving estimates for use in a QRA is rarely discussed. Risk assessors should mention whether they made adjustments based on oversampling, and if so, how was the information incorporated into the risk assessment.

Table III summarizes body weight parameters that have been published. Point and parameter estimates for adult body weight appear moderately consistent, although those for males are more consistent than those for females, which are more consistent than estimates for pooled populations. For children, neither the parameter estimates nor the distributions used to summarize body weight are consistent. Estimates for the Japanese population differ from those for U.S. populations and are not replicated. Another issue regarding consistency is that data from NHANES III are now available, and it will be interesting to see whether parameter estimates and distributions derived from this more recent survey differ from those based on NHANES II.

The point and parameter estimates for adults based on data from NHANES II[20] are both precise—because they are based on a large sample—and accurate—because they are based on statistically based survey designs and population adjustments. For children, on the other hand, data sources differ, and the accuracy and precision are difficult to assess. When investigators do not utilize NHANES information (see Table III), they rarely include information regarding the size of the sample that gave rise to the inputs, nor the standard error. For these, the precision is impossible to assess.

Because the estimates used in the Monte Carlo simulations based on the NHANES II[20] data are total population summaries rather than specific to particular subpopulations that have different characteristics, they are not specific. The data for children also are not specific. The implications of this lack of specificity would depend on the particular application.

For adults, only lognormal and empirical distributions are used. The distributions also can be assessed in terms of the models to which the data are fit and the associated goodness of fit. Brainard and Burmaster[21] fit the NHANES II[20] body weight and height data to a bivariate, lognormal distribution noting that the tails of the distribution are based on sparse, truncated, partially censored data. Since, in a Monte Carlo simulation, often it is the tails that are of greatest concern, this distribution may not be a good representation of the extreme percentiles of the actual data.

For children, lognormal, normal, triangle, and uniform distributions were used. The symmetry, skewness, and tail characteristics of these distributions vary widely.

4.2. Food Consumption

Food consumption data are often used in risk assessment to quantify the intake of certain contaminants. Data on the distributions of food consumption among adults are shown in Table IV. The data are not consistent. The mean estimates for egg consumption vary by a factor of two, while the medians vary by over an order of magnitude. The mean meat consumption values each are within a factor of two, except for the 1-week meat estimate, which, when projected for an average adult male body weight, is approximately five times other estimates. There is some consistency within subsets of categories. For meat consumption, where meat consumption is defined as a combination of consuming pork, lamb, veal, and beef, the empirical 95th percentile is reported as 211 g per day. As should be the case, the total beef consumption, which is given as normally distributed, projects a 95% value of 167 g per day, appropriately smaller as beef is a subset of meat.

The data used come from large population surveys, which suggests that the estimates should be accurate reflections of the population surveyed. The accuracy of the estimates, however, also needs to be considered in light of the different averaging periods. Estimates based on 3-day or 1-week averages are likely not to be accurate when extrapolated over several years or decades. For estimates of long-term risk, this could be particularly troublesome. Another distribution results in negative values. That is, for beef consumption, the normal distribution implies an intake of 0 g at the ninth percentile, and less (i.e., negative values) in the tails. The interpretation of this lower 9% of the beef consumption distribution is ambiguous and poorly reflects upon both the accuracy and specificity.

The data are moderately precise as they come from population studies of varying sizes. The distributions provided are not specific to particular subpopulations. Given the wide variations in diet within and across age, gender, and ethnicity groups, it is suspect that there is substantial inaccuracy with respect to subpopulations. Further, application to subpopulations could be biased, although the direction of that bias would depend on the subpopulation. Some subpopulations, such as vegetarians, have markedly different diets than the average U.S. resident. This type of variation is not addressed by use of these distributions.

Table VI. Fluid Intake: Tap Water and Drinking Water Intake (ml/day) and Estimated Parameters

Fluid	Distribution	*Mean*	*SD*	Min	5%tile	Median	95%tile	Max	Source	References and footnotes
Tap water										
Infants (age ≤1)	Triangle	(413)	(213)	0	(109)	(382)	(802)	1,000	Ershow & Cantor[44]	Goodrum[45]*a*, best = 240
Infants (age ≤1)	Empirical	302	258	0	0	240	775	1,102	Ershow & Cantor[44]	USEPA Handbook[11]*b*, sem = 13
Infants (age ≤)	Lognormal	323	(219)		(97)	267	(734)		Ershow & Cantor[44]	Roseberry & Burmaster[46]*c*, *n* = 403
6–11 months	Point est.	200							NFCS 77/78[32], NHANES II[20]	USEPA Handbook[11]*b*
2 years	Point est.	500							NFCS 77/78[32], NHANES II[20]	USEPA Handbook[11]*b*
Children (ages 1–7)	Triangle	(897)	(364)	190	(377)	(846)	(1,567)	1,900	Ershow & Cantor[44]	Goodrum[45]*a*, best = 600
Children (ages 1–10)	Empirical	736	410	56	192	665	1,516	1,954	Ershow & Cantor[44]	USEPA Handbook[11]*b*, sem = 5
Children (ages 1–10)	Lognormal	701	(372)		(273)	620	(1,406)		Ershow & Cantor[44]	Roseberry & Burmaster[46]*c*, *n* = 5,605
Teens (ages 11–19)	Lognormal	907	(522)		(326)	786	(1,895)		Ershow & Cantor[44]	Roseberry & Burmaster[46]*c*, *n* = 5,801
Teens (ages 11–19)	Empirical	965	562	67	240	867	2,026	2,748	Ershow & Cantor[44]	USEPA Handbook[11]*d*, sem = 7
14–16 years	Point est.	720							NFCS 77/78[32], NHANES II[20]	USEPA Handbook[11]*b*
25–30 years	Point est.	1,040							NFCS 77/78[32], NHANES II[20]	USEPA Handbook[11]*b*
Adults (20–64)	Lognormal	1,265	(657)		(502)	1,122	(2,508)		Ershow & Cantor[44]	Roseberry & Burmaster[46]*c*, *n* = 11,731
Adults (20–64)	Empirical	1,366	728	148	416	1,252	2,707	3,780	Ershow & Cantor[44]	USEPA Handbook[11]*d*, sem = 7
60–65 years	Point est.	1,260							NFCS 77/78[32] & NHANES II[20]	USEPA Handbook[11]*b*
Adults (ages 65+)	Empirical	1,459	643	299	598	1,367	2,636	3,338	Ershow & Cantor[44]	USEPA Handbook[11]*d*, sem = 13
Adults (ages 65+)	Lognormal	1,341	(676)		(547)	1,198	(2,620)		Ershow & Cantor[44]	Roseberry & Burmaster[46]*c*, *n* = 2,541
Adult tap water (F)	Point est.	1,360							Cantor *et al.*[47]	USEPA Handbook[11]*e*, n ≈ 5,000
Adult tap water (M)	Point est.	1,400							Cantor *et al.*[47]	USEPA Handbook[11]*e*, n ≈ 5,000
Adult tap water	Empirical	20%tile ≈ 800	42%tile ≈ 1,120		62%tile ≈ 1,440		82%tile ≈ 1,950		Cantor *et al.*[47]	USEPA Handbook[11]*e*, n ≈ 5,000
All	Empirical	1,193	702	80	286	1,081	2,477	3,415	Ershow & Cantor[44]	USEPA Handbook[11]*d*, *n* = 26,081, sem = 4
All	Lognormal	1,108	(631)		(403)	963	(2,303)		Ershow & Cantor[44]	Roseberry & Burmaster[46]*c*, *n* = 26,081
Fluid intake	Lognormal	1,500	700		(654)	(1,359)	(2,821)		ICRP[37]	McKone & Bogen[22]*f*

Total water intake										
Infants (age <1)	Empirical	1,148	332	510	631	1,120	1,727	2,060	Ershow & Cantor[44]	USEPA Handbook[11]d, sem = 17
Infants (age <1)	Lognormal	1,120	(333)		(665)	1,074	(1,733)		Ershow & Cantor[44]	Roseberry & Burmaster[46]c, n = 403
Children (ages 1–10)	Empirical	1,559	507	617	838	1,497	2,507	3,013	Ershow & Cantor[44]	USEPA Handbook[11]d, sem = 7
Children (ages 1–10)	Lognormal	1,394	(488)		(752)	1,316	(2,301)		Ershow & Cantor[44]	Roseberry & Burmaster[46]c, n = 5,605
Teens (ages 11–19)	Lognormal	1,901	(680)		(1,012)	1,790	(3,168)		Ershow & Cantor[44]	Roseberry & Burmaster[46]c, n = 5,801
Teens (ages 11–19)	Empirical	1,989	719	717	1,025	1,874	3,336	4,251	Ershow & Cantor[44]	USEPA Handbook[11]d, sem = 9
Adults (20–64)	Empirical	2,243	839	821	1,133	2,109	3,793	5,081	Ershow & Cantor[44]	USEPA Handbook[11]d, sem = 8
Adults (20–64)	Lognormal	2,086	(869)		(997)	1,926	(3,718)		Ershow & Cantor[44]	Roseberry & Burmaster[46]c, n = 11,731
Adult total fluid	Point est.	2,000							Cantor *et al.*[47]	USEPA Handbook[11]e
Adult water	Empirical			400				2,000	USEPA Handbook[39]	AIHC[10]
Adults (ages 65+)	Empirical	2,199	728	860	1,196	2,109	3,482	4,370	Ershow & Cantor[44]	USEPA Handbook[11]d, sem = 14
Adults (ages 65+)	Lognormal	2,096	(780)		(1,087)	1,965	(3,552)		Ershow & Cantor[44]	Roseberry & Burmaster[46]c, n = 2,541
All	Empirical	2,072	803	728	1,012	1,950	3,550	4,655	Ershow & Cantor[44]	USEPA Handbook[11]d, sem = 5
All	Lognormal	1,937	(818)		(917)	1,785	(3,474)		Ershow & Cantor[44]	Roseberry & Burmaster[46]c, n = 26,081
Water	Triangle		20% = 800		best = 1,400	90% = 1,900			USEPA Handbook[39]	Stern[40]g
Bottled water (pregnant women)	Point est.	45%							Own data	Swan *et al.*[48]h, n = 5,236

[a] Goodrum, Diamond, Hassett, and Johnson[45] say the triangular PDF (probability distribution function) was used "not because it was assumed the 'true' distribution has a triangular shape, but rather because of the uncertainty in the percentile estimation." Values converted from l/day to ml/day.

[b] Tap water includes tap water, coffee, tea, reconstituted juices (75% tap water),[32] and reconstituted soups (50% tap water). These data are from Pennington (1983), which is based on the U.S. FDA Total Diet Study, which in turn is based on USDA[32] and NHANES II.[20]

[c] Roseberry and Burmaster[46] do not do any formal goodness of fit to check lognormal assumption but graphical methods indicate a reasonable fit. Estimates based on the Travis and Land[63] article, which assumes lognormality. Roseberry and Burmaster[46] do give the formula they use, and any percentile value can be estimated. In addition, these data are from the Ershow and Cantor[44] paper, where a statistical analysis is done of USDA.[32] Roseberry and Burmaster[46] fit to the "binned" data of Ershow and Cantor,[44] and not the actual values.

[d] Ershow and Cantor[44] use tap water and total fluid intake estimates recorded only over a 3-day period based on short-term recall from USDA,[32] likely overestimating the true long-term variation in the population. Total tap water intake is defined as all water from the household tap consumed directly as a beverage or used to prepare foods and beverages. Total fluid is defined as tap water plus water intrinsic to foods and beverages—this is defined as "total water" in NRC.[1] For these empirical distributions, the minimum column is actually the 1st percentile, and the maximum column is actually the 99th percentile.

[e] USEPA[11] cites Cantor *et al.*[47] using 5,258 controls from a case-control study as a surrogate for the general population. A survey asked about intake in a typical week during the winter prior to the interview. Converted from l/day.

[f] Median is the geometric mean. McKone and Bogen[22] say this is the lifetime-averaged arithmetic mean and standard deviation of fluid intake within the U.S. population, but fluid intake is from their section on "ingestion of tapwater" and is therefore placed in the tap water section of this table.

[g] Stern[40] using five studies from USEPA.[39] The overall average of 1.4 is from all five studies. The 90th percentile is an average of the three (of five) studies reporting a 90th percentile. One study (of the five) reported a 20th percentile of 0.8. The 5 studies were: Cantor *et al.*[47] a survey of 9,000 people; Gilleis and Paulin 1983 survey of n = 109; Pennington 1983, used FDA's Total Diet Study; USEPA[39] using USDA[32]; and NAS 1977. NAS 1977 was made up of eight different studies.

[h] Specifically, percent of pregnant women using bottled water.

Table VII. Soil Ingestion Rates in (mg/day) and Estimated Parameters

Group	Distribution	*Mean*	*SD*	Min	5%tile	Median	95%tile	Max	Source	References and footnotes
All ages	Lognormal	(42.9)	(40.7)		(8.4)	(31.2)	(116.3)		Binder *et al.*[49]	Thompson, Burmaster, Crouch[50][i]
Adults		61							Hawley[51][c]	USEPA Handbook[11]
Age > 11 yrs	Empirical	25						100	Mean is from Hawley[51] and max is from Kimbrough *et al.*[52]	LaGoy[18]
> 11 yrs + freq	Empirical	50						100	Mean is from Hawley[51] and max is from Kimbrough *et al.*[52]	LaGoy[18][h]
Adult	Lognormal	54.7	32.8		(18.9)	(46.9)	(116.7)		Calabrese *et al.*[53]	Stern[54][b], median is geometric mean
Child 6–11	Empirical	50						250	Binder *et al.*[49], Clausing *et al.*[55] and own assumptions	LaGoy[18]
Child (1–4)	Cumulative	21			−70	16	110	1,391	from Calabrese *et al.*[56]	Finley *et al.*[57][f], $n = 64$
Child 0–1 yrs	Empirical	50						250	Binder *et al.*[49], Clausing *et al.*[55], and own assumptions	LaGoy[18]
1–3 yrs *Al*[a]	Lognormal	97	169	11		45		1,201	Binder *et al.*[49]	Thompson & Burmaster[58][d]
1–3 yrs *Al*	Empirical	181	203	25		121	584	1,324	Own data	Binder *et al.*[49], also Calabrese *et al.*[56], $n = 59$
1–4 yrs *Al*	Empirical	187	850	1	12	59	243	6,858	Own data	Calabrese *et al.*[53][e], $n = 64$
1–3 yrs *Si*	Lognormal	85	95	10		60		642	Binder *et al.*[49]	Thompson & Burmaster[58][d]
1–3 yrs *Si*	Empirical	184	175	31		136	578	799	Own data	Binder *et al.*[49], also Calabrese *et al.*[56], $n = 59$
1–4 yrs *Si*	Empirical	211	692	6	21	81	402	5,582	Own data	Calabrese *et al.*[53][e], $n = 64$
1–3 yrs *Min*	Empirical	108	121				386		Own data	Binder *et al.*[49], $n = 59$
1–3 yrs *Ti*	Empirical	1,834	3,091	4		618	9,590	17,076	Own data	Binder *et al.*[47], $n = 59$
1–4 yrs *Ti*	Empirical	577	1,220	10	12	192	2,353	6,911	Own data	Calabrese *et al.*[53][e], $n = 64$
Child 1–6 yrs	Empirical	100						500	Binder *et al.*[49], own assumptions, and Lepow *et al.*[59]	LaGoy[18]
Age = 2.5 yrs	Point est.	165							Hawley[51][c]	USEPA Handbook[11]
Nursery school *Al*	Normal	232	263		(−200)	232	(665)		Own data	Clausing *et al.*[55][g], $n \approx 300$
Nursery school *Min*	Normal	105	67		(−5)	105	(215)		Own data	Clausing *et al.*[55][g], $n \approx 300$
Age = 6 yrs		24							Hawley[51][c]	USEPA Handbook 1995[11]
Hospital 2–4 yrs *Al*	Normal	56	24		(16)	56	(95)		Own data	Clausing *et al.*[55][g], $n \approx 300$
Hospital 2–4 yrs *Min*	Normal	49	22		(13)	59	(85)		Own data	Clausing *et al.*[55][g], $n \approx 300$
Child (0.5–5 yrs)	Triangle	175	81	25	62.5	162.8	324.8	400	Own judgment and USEPA Handbook[60]	Oregon DEQ[31][j]
Nonchild (>6 yrs)	Triangle	58.3	15.6	25	34.7	56.7	86.2	100	Own judgment and USEPA Handbook[60]	Oregon DEQ[31][j]
Child (2–7 yrs) *Al*	Empirical	38.9				25.3			Own data	Davis[61][k], sem = 14.4
Child (2–7 yrs) *Si*	Empirical	82.4				59.4			Own data	Davis[61][k], sem = 12.2
Child (2–7 yrs) *Ti*	Empirical	245.5				81.3			Own data	Davis[61][k], sem = 119.7

Notes to Table VII

[a] Italicized values are the tracer elements *Al* = aluminum, *Si* = silicon, *Ti* = titanium, *Min* = minimum across the tracers.

[b] Stern[54] lognormal assumed because "small sample size and intraindividual uncertainty argue against attempting to rigorously define the distribution of adult soil ingestion from these data. The distribution . . . is therefore modeled assuming lognormality" (p. 203).

[c] Hawley[51] based estimates on ingestion of both soil and dust.

[d] Thompson and Burmaster[58] fit the data from Binder *et al.*[49] Data in Thompson and Burmaster for underlying normal distribution do not agree with given lognormal values.

[e] Calabrese *et al.*[53] data collected over a short period of time = 2 weeks, likely overestimating long-term average soil ingestion rate. Numbers in table are from a study,[56] using average fecal output over 8 days, and ignoring food intake.

[f] Finley *et al.*[57] construct a cumulative distribution representing daily soil ingestion for children using the mean (21 mg/day), median (16 mg/day), 5th percentile (−70 mg/day), 10th (−35 mg/day), 90th (67 mg/day), 95th (110 mg/day), and maximum (1,391 mg/day possible outlier) values reported by Calabrese *et al.*[56] Finley *et al.*[57] truncate the distribution at the 36th percentile (0 mg/day) to avoid negative values, and consider the distribution to be "amply" conservative for children 1–4 years.

[g] Clausing *et al.*[55] say that the data are "almost normally distributed."

[h] LaGoy[18] increases the average intake for frequent hand-mouth contact (e.g., smokers).

[i] Thompson *et al.*[50] use the data from Binder *et al.*[49] using actual stool weights instead of assuming stool weights of 15 g/day as Binder *et al.* did. Thompson *et al.* cite LaGoy[18] and Thompson and Burmaster.[58]

[j] Oregon DEQ[31] uses the USEPA[60] plus "judgment" to define the boundaries of the soil ingestion. The soil ingestion for juveniles (6 to 17 years) is "conservatively" (DEQ's word) assumed to be the same as adults.

[k] Davis *et al.*[61]

Since all but one of the distributions are empirical, one can say little about the use of distributions. For the one assumed distribution—the normal distribution of beef consumption—it is noted that the specific foods have skewed distributions with long right-tail distributions. The normal distribution likely underestimates these effects.

4.3. Breathing Rates

Breathing rates sometimes are used in risk assessments to quantify the intake of airborne contaminants. Data on the distributions of breathing rates among adults are shown in Table V. Data are presented for a variety of different situations, with more than one estimate provided for only one situation. Nonetheless, there is moderate consistency among means, with one study being a notable outlier. The medians show closer agreement. Few other parameters are reported. A triangular distribution (skewed to the right) is used to describe nonspecific adult breathing rate and adult breathing rate in the shower. Breathing rate at home, and lifetime daily rate as calculated from food-energy intake are provided as point estimates. Breathing rate per unit body weight is provided as lognormally distributed. When the implied breathing rate from this distribution is calculated assuming a 70-kg male, the estimate is more than twice the point estimates reported by others.

The accuracy of the data varies. Thompson and Evans[36] based the point estimate of 12 m^3/day on the physiologically based pharmacokinetic (PBPK) model developed by Bogen and McKone.[28] McKone and Bogen[22] calculate the breathing rate per unit body weight using data from the International Commission on Radiological Protection (ICRP).[37] Smith[42] based his breathing rate data on minimum, maximum, and most-likely values obtained from USEPA publications[39] and, with little justification, approximated them by the triangular distribution. Stern[40] also based the distribution he used on USEPA data,[39] utilizing the USEPA handbook data that bases inhalation rate activity (e.g., percentage of time at rest, and involvement in light, moderate, and heavy activity), and also used the triangle distribution. The parameter estimates used by Stern,[40] however, differ from those used by Smith.[42] Layton[41] calculates inhalation rate from food-energy intake and provides only a point estimate.

Breathing rates present a classic case where there is a lack of information on the precision of these estimates, as the sample sizes, standard errors, or both from which the breathing rates are extrapolated

are not provided. Without the precision, the quality of these distributions as inputs to a Monte Carlo–based risk assessment is left in question. The distributions are specific to certain activities (i.e., at home, in the shower), but not to subpopulations.

4.4. Fluid Intake

Total fluid intake estimates are often based on widely distributed and purchased beverages and, if used to estimate exposure to particular substances in the fluid, may overestimate exposure if the substance is only present in local sources. Fluid intake is usually recorded with other population characteristics, enabling stratification across subgroups, such as gender and age. In addition, geographical estimates have been attempted, as well as climate-specific studies. An enumeration of all such site- and population-specific studies is beyond the scope of this article. Further, most Monte Carlo simulations reviewed used fluid and water intake data that were stratified only on age, if at all. Therefore the focus is on the other four properties defined on non-population-specific estimations.

For point estimates, the USEPA uses the quantity of 2 l/day for adults and 1 l/day for infants (defined as 10 kg or less) as point estimates for the default drinking water intake rates, and as upper percentiles as tap water intake rates. For distributions, most authors cite a USEPA *Exposure Factors Handbook* or the U.S. Department of Agriculture's (USDA's) 1977–1978 *Nationwide Food Consumption Survey* (NFCS)[32] as the source, and because of this, there is some consistency in the use of tap water intake.

The consistency of the data on fluid intake can be assessed only within reported strata. In these groups, which have only one to three estimates, the data are fairly consistent. For example, the point estimates cited from Pennington[64] are consistent with the medians cited from Ershow and Cantor. The estimates from Stern,[40] and from Cantor *et al.*[47] are all in the mid-1,000-ml/day range. As can be seen by the Ershow and Cantor data, however, data vary markedly by age and use of age-specific information is likely to improve the accuracy of Monte Carlo simulations.

Data cited from Cantor *et al.*[47] are based on more than 5,000 controls (75% of which were male) from a bladder cancer study. The controls were randomly selected from the population to match the cancer cases regarding age, gender, and geographic location. While this group is large enough to provide precise estimates, the accuracy of the estimates needs to be considered in light of the fact that this matched control population may not be representative of the U.S. population. Ershow and Cantor[44] estimated water intake rates based on data collected by the NFCS,[32] which included over 26,000 individuals surveyed over a 3-day period. Short-term-recall surveys often introduce a degree of variation into the accuracy of estimates. The USDA NFCS, however, is a large geographically and seasonally balanced survey of a representative sample of the U.S. population, and the estimates obtained by Ershow and Cantor are very precise. Since, however, the survey is 20 years

Table VIII. Summary

Parameter	Stratification[a]	Consistent	Accurate	Precise[b]	Specific	Distributions used[c]
Body weight	Adult	Yes	Yes	Yes	No	E, LN, P
	Child	No	Yes	Yes	No	LN, N, T, U
Food consumption	Meats	No	No	Yes	Adults only	E, N
	Eggs	No	No	Yes	Adults only	E
Breathing rates		Yes[d]	Mixed	No	Adults only	LN, P, T
Fluid intake	Tap	Yes[e]	Yes	Yes	Age	E, LN, P, T
	Total	Yes[e]	Yes	Yes	Age	E, LN, P, T
Soil ingestion		No	No	No	Child only	E, LN, N, P, T

[a] Stratification refers to large subdivisions of data for which several estimates were found in the literature.
[b] For precision: Yes if the standard deviation is consistently less than one half the mean, or there is a small standard error of the mean, or there is a large sample size; No if the standard deviation is generally as large as or larger than the mean, or lack of specific information.
[c] Distributions: E = empirical distribution; L = lognormal distribution; N = normal distribution; P = point estimates only provided; T = triangular distribution; U = uniform distribution.
[d] One study is an outlier.
[e] Only two studies per age stratum, but intake increases with age across studies.

old and there has been a large rise in the use of bottled water, it is unclear how accurate the data are.

Most of the estimates are based on empirical data. Some investigators fit distributions to these publicly available data. For example, Roseberry and Burmaster[46] fit the Ershow and Cantor[44] data to a lognormal distribution. Note that Ershow and Cantor data are "binned," and the Roseberry and Burmaster data are fitted to the "binned" data, and not the NFCS[32] raw data. No justification or measure of the goodness of fit of the distribution is provided. It was found that the fitted lognormal distributions for tap water consistently overestimate the lower tail and underestimate the median for all age groups, but fit the upper tail fairly well. Since it is often this upper tail that is of concern in Monte Carlo distributions, then these distributions are likely sufficient for a Monte Carlo simulation. Stern[40] proposes a triangle distribution because he did not believe that adequate data were available to accurately estimate the upper and lower percentage values of the distribution.

4.5. Soil Ingestion

The ingestion of soil is a potential source of exposure, especially for children. Several studies have attempted to measure the amount of soil ingestion by children. A few studies have looked at adults. Soil ingestion is a prime example of how the same exposure can be measured, interpreted, and extrapolated in different ways by different researchers. Given that different methodologies provide markedly different results, it is not possible to assess accuracy in this review. For example, data collection efforts vary in terms of length of observation (e.g., Binder *et al.*[49] observed for 3 days, while Calabrese *et al.*[53,56] observed for 2 weeks) and season of observation (e.g., Clausing *et al.*[55] sampled subjects at the beginning and end of the summer).

The data also are not consistent. This is primarily due to investigators using different methodological approaches to estimate soil ingestion, even using the same raw data. For example, Thompson and Burmaster[58] reevaluated the Binder *et al.*[49] data using actual fecal weights, instead of Binder *et al.*'s assumed fecal weight of 15 g/day. These estimates differ from those of Calabrese *et al.*,[53,56] especially at the upper tail.

Surprisingly, in spite of the accuracy limitations, the estimates are reasonably precise. Clausing *et al.*[55] sampled approximately 300 children, half of which were also sampled again a few months later. Calabrese *et al.*[53,56] examined eight tracer elements for reliability, showing that aluminum and silicon were among the most reliable based on precise estimates and percentage of recovery. These tracers were then used on 64 children, aged 1 to 4. Binder *et al.*[49] used 59 children, aged 1 to 3.

Site and population specificity are an important issue for soil ingestion, as children are a highly exposed group compared to adults. Childhood development phases (mouth exploration, crawling) change fairly quickly, and as they do, so does the opportunity for exposure. LaGoy[18] clarifies that point estimates of ingestion differ by a factor of ten depending on the site and source of the exposure. Further, some estimates are based on soil alone, while some also attempt to include dust.

For distributions, some investigators used normal distributions, some truncated distributions, and others lognormal distributions. Thompson and Burmaster[58] used a lognormal distribution. There is a lack of fit in the tails of the distribution resulting in differences from Calabrese *et al.*[53,56] estimates. Clausing *et al.*[55] uses a normal distribution although the data appear likely to better fit a truncated normal distribution. Finley *et al.*[57] fit a distribution to Calabrese *et al.*'s data, which, to prevent negative values, results in a "truncated" cumulative distribution.

One way of investigating the need for population-specific estimates is to compare estimates across populations. Using data reported by Lagoy,[18] soil ingestion estimates versus age were plotted (Fig. 1). When assessed with a linear regression, the data do not show a trend with age, even when the first observation is omitted. As one would expect mouthing activity to decrease with age, this calls into question the validity of the estimates. Further research is needed.

5. DISCUSSION

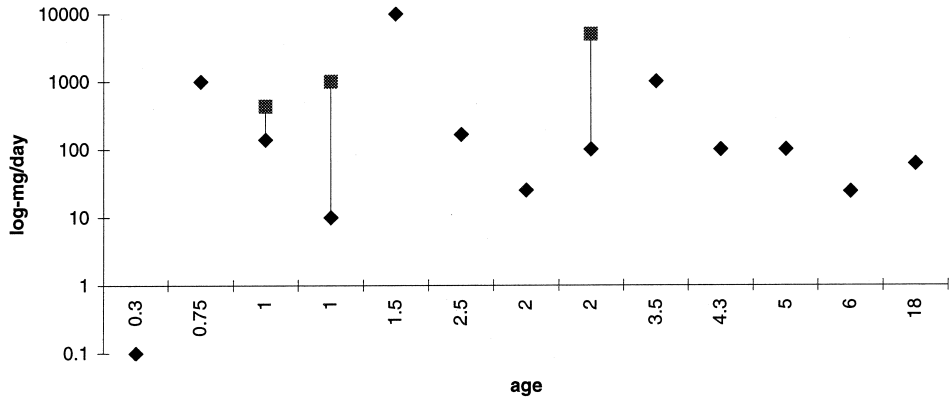
The goal of this article is to critically review and assess the use of input data to Monte Carlo simulations in QRA. It is recognized that specific situations or contexts may require specific choices of distributions. In most situations reviewed, however, the case was not made for tailoring distributional choices to the specific conditions under study. Thus, while such situations can be imagined, they do not seem relevant to this review.

The question is whether different choices made by different investigators can have a substantial impact on the resultant risk estimate. Specifically, for similar applications, do investigators use similar sources of data, treat them in similar manners, and derive similar results for use in this process? If results are similar, then results of the simulations can be com-

pared in terms of other issues, assumptions, and parameters. If not, then all evaluations and comparisons of results must address not only the traditional concerns with risk assessments,[2] but also the data and assumptions underlying the Monte Carlo simulations. This would greatly complicate, and likely obfuscate, such comparisons and make the critical evaluation of results all but impossible for policy analysts.

For this evaluation, the published literature was reviewed, the most frequently estimated parameters were identified, and the distributions and parameter estimates used were summarized. Not surprisingly, the results were mixed. Where large, carefully collected data sets exist (e.g., body weight, food, and fluid intake), investigators usually use these as their starting point. Sometimes approaches of different investigators are consistent, sometimes not. What is somewhat troubling is that even where extensive, well-collected data exist, investigators used a variety of different distributional shapes to approximate the same data. Various investigators have reported that alternative choices can make a substantial difference. Where large, well-collected data sets do not exist, investigators developed their own data. This, too, led to substantial disparity in parameter estimates.

The implication of even this limited survey is that the choices investigators have made for input distributions may substantially affect the output of their analyses. QRAs using different input assumptions may result in substantially different risk estimates. Yet, the cause of these differences, because they are buried in assumptions of the inputs to the simulations, may not be apparent to the readers or users of the risk assessments. The failure of risk assessors to be forthright about the sensitivity of their results to these choices, and the validity of alternative distributions, may impact on the critical communication and trust between risk assessors and policy makers. Particularly in simulations that focus attention on the tails of the distributions, differences may be observed that reflect only this assumption of distributional shape, which, likely is unknown to the policy makers using these results as a basis for reaching decisions about risk. Thus, rather than clarifying the results of risk assessments without adequate disclosure, Monte Carlo simulations may further obscure details of the



0-9 months	0	based on estimates of mouthing behavior and amount of soil on hands
9-18 months	1000	based on estimates of mouthing behavior and amount of soil on hands
age 1 to 3 years	140-430	based on estimate of paint consumption by children with pica
age 1 to 3 years	10-1000	based on soil ingestion on candy
age 1.5-3.5 years	10000	based on estimates of mouthing behavior and amount of soil on hands
2.5 years	165	based on estimated ingestion of both soil and dust
age 2 to 6 years	25	based on mouthing behavior and amount of soil on hands
age 2 to 6 years	100-5000	based on normal children and children with habitual pica
3.5-5 years	1000	based on estimates of mouthing behavior and amount of soil on hands
Mean age 4.3 yrs	100	based on observation of mouthing behavior and amount of soil on hands
>5 years	100	based on estimates of mouthing behavior and amount of soil on hands
6 years	24	based on estimated ingestion of both soil and dust
adults	61	based on estimated ingestion of both soil and dust

Fig. 1. Soil ingestion estimates (mg/day) versus age reported in literature, as listed in Lagoy.[18]

methodology used to derive the risk estimates of use to regulators and lead to substantive misinterpretations. Potentially compounding misinterpretations are the use of raw data or distribution parameters where a standard error is needed, and not a standard deviation. The lack of disclosure of sample sizes and standard errors can also be misleading.

The basic intent of Monte Carlo simulations is a good one. But, to make it a more effective tool, a clear set of guidelines must be developed, as well as practices and disclosures that allow readers and users to appreciate and evaluate the choices made and their implications on the final results. While there has been some activity in this direction, for example, Burmaster and Anderson,[19] it appears to be little used. There should be more emphasis placed on sensitivity analyses, quantifying the impact of assumptions, and a more thorough discussion (and tabulation) of sources of variation as part of the presentation of any risk assessment results.

In addition, more attention must be paid to the data underlying these estimates. Since the scant data provided here show that parameters can differ markedly by age, gender, and location, among other parameters, appropriate data sources must be developed from which to extract relevant data or parameters.

If such practices and disclosures are followed, the authors believe that Monte Carlo simulations can greatly enhance the accuracy and appropriateness of specific risk assessments. The price for this increase and improvement in information, however, is the necessity for greater detail and complexity in the presentation of results. Without such disclosures, it is believed that researchers will be increasing the size of the risk assessment "black box," a concern already raised by many critics of more traditional risk assessments.

ACKNOWLEDGMENTS

The authors would like to acknowledge the suggestions and encouragement of the area editor. This work is partially funded by the New Jersey Department of Environmental Protection and Consortium for Risk Evaluation with Stakeholder Participation.

REFERENCES

1. National Research Council (NRC). (1983). *Risk assessment in the federal government: Managing the process.* Washington, DC: National Academy Press.
2. Wartenberg, D., & Chess, C. (1992, March/April). Risky business: The inexact art of hazard assessment. *The Sciences,* 17–21.
3. Fiering, M. B., Wilson, R., Kleiman, E., & Zeise, L. (1984). Statistical distributions of health risks. *Civil Engineering Systems 1*, 129–138.
4. Eschenroeder, A. Q., & Faeder, E. J. (1988). A Monte Carlo analysis of health risks from PCB-contaminated mineral oil transformer fires. *Risk Analysis, 8*, 291–297.
5. Finkel, A. M. (1990). *Confronting uncertainty in risk management, a guide for decision makers.* Washington, DC: Center for Risk Management Resources for the Future.
6. Burmaster, D. E., & von Stackelberg, K. (1991). Using Monte Carlo simulations in public health risk assessments: Estimating and presenting full distributions of risk. *Journal of Exposure Analysis and Environmental Epidemiology, 1*(4), 491–512.
7. Burmaster, D. E., & Harris, R. (1993). The magnitude of compounding conservatisms in Superfund risk assessments. *Risk Analysis, 13*, 131–134.
8. McKone, T., & Bogen, K. (1991). Predicting the uncertainties in risk assessment. *Environmental Science and Technology, 25*(10), 1674–1681.
9. Cullen, A. C., & Frey H. C. (1999). *Probabilistic techniques in exposure assessment.* New York: Plenum Press.
10. American Industrial Health Council (AIHC). (1994). *Exposure factors sourcebook.* Washington, DC: Author.
11. U.S. Environmental Protection Agency (USEPA). (1995). *Exposure factors handbook.* Washington, DC: Office of Research and Development.
12. Morgan, M. G., & Henrion, M. (1992). *Uncertainty, a guide to dealing with uncertainty in quantitative risk and policy analysis* (1st paperback ed.). New York: Cambridge University Press.
13. Haas, C. N. (1997). Importance of distributional form in characterizing inputs to Monte Carlo risk assessments. *Risk Analysis, 17*, 107–113.
14. Hattis, D., White, P., Marmorstein, L., & Koch, P. (1990). Uncertainties in pharmacokinetic modeling for perchloroethylene. I. Comparison of model structure, parameters, and predictions for low-dose metabolism rates for models derived by different authors. *Risk Analysis, 10*, 449–455.
15. Seiler, F. A., & Alvarez, J. L. (1996). On the selection of distributions for stochastic variables. *Risk Analysis, 16*, 5–18.
16. Hamed, M. M., & Bedient, P. B. (1997). On the effect of probability distributions of input variables in public health risk assessment. *Risk Analysis, 17*, 97–105.
17. Bukowski, J., Korn, L., & Wartenberg, D. (1995). Correlated inputs in quantitative risk assessment: The effects of distributional shape. *Risk Analysis, 15*, 215–219.
18. LaGoy, P. K. (1987). Estimated soil ingestion rates for use in risk assessment. *Risk Analysis, 7*, 355–359.
19. Burmaster, D. E., & 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*, 477–482.
20. National Center for Health Statistics. (1987). *Anthropometric reference data and prevalence of overweight, United States, 1976–1980.* (DHHS Publication No. (PHS) 87-1688. Series 11, No. 238). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics.
21. Brainard, J., & Burmaster, D. E. (1992). Bivariate distributions for height and weight of men and women in the United States. *Risk Analysis, 12*, 267–275.
22. McKone, T. E., & Bogen, K. (1992). Uncertainties in health-risk assessment: An integrated case study based on tetrachloroethylene in California groundwater. *Regulatory Toxicology and Pharmacology, 15*, 86–103.
23. Najjar, M., & Roland, M. (1987). Anthropometric reference data and prevalence of overweight (Publication No. (PHS) 87-1688). Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics.
24. McKone, T. (1994). Uncertainty and variability in human exposures to soil contaminants through home-grown food: A Monte Carlo assessment. *Risk Analysis, 14*, 449–463.

25. Stephens, T., & Craig, C. L. (1990). The well-being of Canadians: Highlights of the 1988 Campbell's survey. Ottawa, Canada: Canadian Fitness and Lifestyle Research Institute.
26. Richardson, G. M., & Allen, M. (1996). A Monte Carlo assessment of mercury exposure and risks from dental amalgam. *Human and Ecological Risk Assessment, 2*(4), 709–761.
27. Kerr, G. D. (1979). Organ dose estimates for the Japanese atomic-bomb survivors. *Health Physics, 37*, 487–508.
28. Bogen, K., & McKone, T. (1988). Linking indoor air and pharmacokinetic models to assess tetrachloroethylene risk. *Risk Analysis, 8*, 509–520.
29. GCA Corp. (1985). *Development of statistical distributions or ranges of standard factors used in exposure assessments.* Washington, DC: U.S. Environmental Protection Agency, Office of Health and Environmental Assessment, Office of Research and Development.
30. Finley, B., & Paustenbach, D. (1994). The benefits of probabilistic exposure assessment: Three case studies involving contaminated air, water, and soil. *Risk Analysis, 14*, 53–73.
31. Oregon Department of Environmental Quality. (1997). *DRAFT guidance for use of probabilistic analysis in human health exposure assessments.* Portland, OR: Waste Management & Cleanup Division.
32. U.S. Department of Agriculture (USDA). (1978). *Food and nutrient intakes of individuals in one day in the United States, spring 1977. Nationwide Food Consumption Survey 1977–78* (Preliminary Report No. 2). Hyattsville, MD: Author.
33. Pao, E. M., Fleming, K. H., Guenther, P. M., & Mickle, S. J. (1982). *Foods commonly eaten by individuals: Amount per day and per eating occasion* (Home Economics Report No. 44). Hyattsville, MD: U.S. Department of Agriculture.
34. U.S. Department of Agriculture (USDA). *1987–88 dataset. Nationwide food consumption survey 1987/88 household food use* (1987/88 NFCS database). Hyattsville, MD: Author.
35. Finley, B., & Paustenbach, D. (1992, February 10–11). *Opportunities for improving exposure assessments using population distribution estimates.* Paper presented for the Committee on Risk Assessment Methodology, Washington, DC.
36. Thompson, K. M., & Evans, J. S. (1993). Workers breath as a source of perchloroethylene (PERC) in the home. *Journal of Exposure Analysis and Environmental Epidemiology, 3*(4), 417–430.
37. International Commission on Radiological Protection (ICRP). (1975). *ICRP publication 23: Report of the task group on reference man.* Oxford, U.K.: Pergamon Press.
38. Checkoway, H., Pearce, N., & Crawford-Brown, D. J. (1989). *Research methods in occupational epidemiology.* New York: Oxford University Press.
39. U.S. Environmental Protection Agency (USEPA). (1989). *Exposure factors handbook.* Washington, DC: Office of Health and Environmental Assessment.
40. Stern, A. H. (1993). Monte Carlo analysis of the U.S. USEPA model of human exposure to cadmium in sewage sludge through consumption of garden crops. *Journal of Exposure Analysis and Environmental Epidemiology, 3*(4), 449–469.
41. Layton, D. W. (1993). Metabolically consistent breathing rates for use in dose assessments. *Health Physics, 64*(1), 23–36.
42. Smith, R. L. (1994). Use of Monte Carlo simulation for human exposure assessment at a Superfund site. *Risk Analysis, 14*, 433–439.
43. Stahl, W. R. (1967). Scaling of respiratory variables in animals. *Journal of Applied Physiology, 22*(3), 453–460.
44. Ershow, A. G., & Cantor, K. P. (1989). *Total water and tapwater intake in the United States: Population based estimates of quantities and sources.* Washington DC: Life Sciences Research Office, Federation of American Societies for Experimental Biology.
45. Goodrum, P. E., Diamond, G. L., Hassett, J. M., & Johnson, D. L. (1996). Monte Carlo modeling of childhood lead exposure: Development of a probabilistic methodology for use with the USEPA IEUBK model for lead in children. *Human and Ecological Risk Assessment, 2*(4), 681–708.
46. Roseberry, A. M., & Burmaster, D. E. (1992). Lognormal distributions for water intake by children and adults. *Risk Analysis, 12*, 99–104.
47. Cantor, K. P., Hoover, R., Hartge, P., Mason, T., Silverman, D. T., *et al.* (1987). Bladder cancer, drinking water source, and tap water consumption: A case-control study. *Journal of the National Cancer Institute, 79*(6), 1269–1279.
48. Swan, S. H., Nuetra, R. R., Wrensch, M., Hertz-Picciotto, I., Windham, G. C., Fenster, L., Epstein, D. M., & Deane, M. (1992). Is drinking water related to spontaneous abortion? Reviewing the evidence from the California Department of Health studies. *Epidemiology, 3*, 83–93.
49. Binder, S., Sokal, D., & Maughan, D. (1986). Estimating soil ingestion: The use of tracer elements in estimating the amount of soil ingested by young children. *Archives of Environmental Health, 41*(6), 341–345.
50. Thompson, K. M., Burmaster, D. E., & Crouch, E. A. C. (1992). Monte Carlo techniques for quantitative uncertainty analysis in public health risk assessment. *Risk Analysis, 12*, 53–63.
51. Hawley, J. (1985). Assessment of health risk from exposure to contaminated soil. *Risk Analysis, 5*, 289–302.
52. Kimbrough, R. D., Falk, H., Stehr, P., & Fries, G. (1984). Health implications of 2,3,7,8-tetrachlorodibenzodioxin (TCDD) contamination of residential soil. *Journal of Toxicology and Environmental Health, 14*, 47–93.
53. Calabrese, E. J., Pastides, H., Barnes, R., Edwards, C., Kostecki, P., Stanek, E., *et al.* (1989). How much soil do young children ingest: An epidemiological study. *Journal of Regulatory Toxicology and Applied Pharmacology, 10*, 123–127.
54. Stern, A. H. (1996). Derivation of a target concentration of Pb in soil based on elevation of adult blood pressure. *Risk Analysis, 16*, 201–210.
55. Clausing, P., Brunekreef, B., & van Wijnen, J. H. (1987). A method for estimating soil ingestion by children. *International Archives of Occupational & Environmental Health, 59*(1), 73–82.
56. Calabrese, E. J., Pastides, H., Barnes, R., *et al.* (1991). How much soil do young children ingest: An epidemiologic study. In E. J. Calabrese & P. Kostecki (Eds.), *Petroleum contaminated soils* (p. 30). Chelsea, MI: Lewis.
57. Finley, B., Proctor, D., Scott, P., Harrington, N., Paustenbach, D., & Price, P. (1994). Recommended distributions for exposure factors frequently used in health risk assessment. *Risk Analysis, 14*, 533–553.
58. Thompson, K. M., & Burmaster, D. E. (1991). Parametric distributions for soil ingestion by children. *Risk Analysis, 11*, 339–342.
59. Lepow, M. L., Bruckman, L., Gillete, M., Markowitz, S., Rubino, R., & Kapish, J. (1975). Investigations into sources of lead in the environment of urban children. *Environmental Research, 10*, 415–426.
60. U.S. Environmental Protection Agency (USEPA). (1996). *Exposure factors handbook.* Washington, DC: Office of Research and Development.
61. Davis, S., Waller, P., Buschbom, R., Ballou, J., & White, P. (1990). Quantitative estimates of soil ingestion in normal children between the ages of 2 and 7 years: Population-based estimates using aluminum, silicon, and titanium as soil tracer elements. *Archives of Environmental Health, 45*(2), 112–122.
62. U.S. Environmental Protection Agency (USEPA). (1985). *Methods for assessing exposure to chemical substances: Vol. 2. Methods for assessing exposure to chemical substances in the ambient air* (prepared by Versar, Inc. under contract 68-01-6271). Washington DC: USEPA, Office of Toxic Substances.
63. Travis, C. C., & Land, M. (1990). Estimating the mean of data sets with nondetectable values. *Environmental Science and Technology*, 24, 961–962.
64. Pennington, J. A. T. (1983). Revision of the total diet study—food lists and diets. *Journal of the American Dietetic Association, 82*, 166–173.