REPORT TO THE COMMISSION ON RISK ASSESSMENT AND RISK MANAGEMENT: Health risk assessments prepared *per* the risk assessment reforms under consideration in the U.S. Congress

Prepared by

Edmund A.C. Crouch, Ph.D.,
Richard R. Lester,
Timothy L. Lash, M.P.H.,
Sarah R. Armstrong, M.S., M.S., and
Laura C. Green, Ph.D., D.A.B.T.

October 23, 1995

Acknowledgments

This work has been supported by the Commission on Risk Assessment and Risk Management. The opinions expressed herein are those of the authors, who accept responsibility for any errors or omissions.

Contents

List of Tables

List of Figures

1 Summary

The Commission on Risk Assessment and Risk Management retained Cambridge Environmental Inc. to conduct case studies of health risk assessment that conform with proposed regulatory reform legislation[1] and to comment, as risk assessors, on the required methods. The principal relevant mandate in these legislative proposals is that the conservative point estimates of risk currently generated and relied upon be augmented with estimates that are in some sense "best" — that are central tendency estimates, generated by taking better account of the uncertainties and variabilities in the underlying data and assumptions.

To illustrate the techniques required to satisfy such a mandate, we studied four cases. The objective of the first case study was to estimate incremental lifetime risk of cancer to an individual in a population whose water supply had been contaminated with part-per-billion levels of 1,1-dichloroethylene (1,1-DCE). The second case study differed from the first only in that 1,1-DCE was allowed, consistent with its dose-response data, to have either an anticarcinogenic or a carcinogenic potency, rather than being constrained to have only a carcinogenic potency, as is the current regulatory norm. The third case study differed from the first only in that it considered exposure to similar levels of vinyl chloride, a potent and known human carcinogen, rather than exposure to the equivocally carcinogenic 1,1-DCE. The fourth case study estimated incremental lifetime risk of cancer associated with occupational exposures, rather than low-level environmental exposures, to 1,1-DCE.

For each case study, we first estimated the incremental lifetime risk of cancer to a "reasonably maximally exposed individual" using the methods currently recommended by U.S. EPA. We then prepared a distribution of risk estimates by choosing parameter values for each variable from the distribution defined for that variable and combining these choices in the risk equation. These latter tasks required (1) significant research in the scientific literature, and (2) not a small amount of statistical and computational expertise. Using computer software we created, we repeated the risk calculation about 20,000 times, gathering up each estimate of incremental lifetime risk of cancer to define its distribution. From the distribution, we could estimate the mean, median, and 95th percentile (and other statistics) of the distribution for the incremental lifetime risk of cancer. Each of these might be considered a "best" estimate of risk.

The results of the four case studies are summarized in the following table.

[1] In particular, bills S 343 and HR 1022.

Table 1 Statistics of the distributions of risk estimates from the case studies

Case	Median (50^{th} percentile)	Mean	95^{th} percentile	Current EPA-style Point-estimate (reasonably maximum exposure)
1,1-DCE, standard	1.2×10^{-9}	1.6×10^{-6}	1.7×10^{-6}	1.3×10^{-4}
1,1-DCE, non-standard	-2.0×10^{-9}	-9.5×10^{-6}	1.7×10^{-6}	—
Vinyl chloride (standard)	1.4×10^{-6}	8.8×10^{-5}	2.0×10^{-4}	4.1×10^{-4}
1,1-DCE workers	1.4×10^{-8}	3.6×10^{-3}	8.4×10^{-3}	2.7×10^{-2}

Several comparisons are noteworthy. In the first case study, U.S. EPA methods (specifically, those used for risk assessment of Superfund sites) yielded a point-estimate of risk of 1.3×10^{-4}. Although such an upper-bound point estimate is typically assumed by many to be at about the 95^{th} percentile of the risk estimate distribution, it corresponded here to the 99.8th percentile of such a distribution. The probabilistic method employed here found that the 95^{th} percentile of the distribution was about 80-fold lower — 1.7×10^{-6}. These two different estimates — both upper-bound — would likely indicate dramatically different intervention strategies. Risks as high as the former often require extensive remediation, whereas risks as low as the latter usually do not.

The second case study, in which exposures to 1,1-DCE were allowed to confer either beneficial or detrimental effects on cancer risk, yielded two central tendency estimates of risk that were negative — so suggested that low levels of 1,1-DCE might confer no excess risk of cancer, and might even confer a small benefit. Nonetheless, the 95^{th} percentile of the distribution of risk estimates in the second case study was identical to that estimated in the first case study (1.7×10^{-6}). Thus, allowing the relevant portions of the bioassay data themselves to define the slope and bounds of the dose-response curve — as opposed to imposing standard, regulatory restrictions on that curve — yielded both dramatically different central tendency estimates and identical upper-bound estimates.

The third case study, in which exposures to vinyl chloride were substituted for dose-equivalent exposures to 1,1-dichloroethylene, yielded a point estimate of risk (4.1×10^{-4}) that was only three times larger than the point estimate generated in the first study for 1,1-DCE. Such a minor difference belies the substantial differences in the quality and quantity of data surrounding the carcinogenicity of these two chemicals. In contrast, the probabilistic methods yield a 95^{th} percentile estimate for the risks from vinyl chloride that is some 120-times larger than the estimate from 1,1-DCE.

Finally, the fourth case study suggested that (1) occupational exposures to 1,1-DCE were, as expected, substantially riskier than low-level environmental exposures, and (2) that the point-estimate of risk is only some three-fold larger than the 95th percentile estimate. Under certain circumstances, such as relatively high exposures, the deterministic and probabilistic methods may thus yield reasonably similar upper-bound estimates of risk.

Working through these case studies, we have reached certain conclusions about the proposed risk assessment reforms. Among these opinions are:

- **Performing risk assessment holistically and probabilistically is not easy.** Considerable research must be made into the ranges of plausible estimates for a vast number of inputs. Considerable quantitative expertise, including computer-programming skills, are required to design and implement the method. The risk assessor must genuinely understand — as opposed to merely use — many sorts of models — and perhaps be able to create some anew. He or she must combine distributions in valid manners.

- **Current point-estimates of risk may obscure underlying scientific complexities and other important information.** Public health policy demands upper-bound estimates of risk; but if these are calculated too crudely, they prevent efficient, health-protective decision-making.

- **Under various circumstances, probabilistic risk assessment may indeed be informative and worthwhile.** Techniques used to generate risk estimates should scale with the situation to be assessed. Some situations can be shown to be harmless under almost any method of risk analysis; running full Monte Carlo analyses on these would be inefficient. Other situations are much harder to call, have high stakes, or otherwise demand more sophisticated analysis. For such situations, probabilistic methods, carefully and honestly implemented, may offer the best current hope.

- **Health risk assessment is typically dominated by uncertainty, rather than by variability.** Distributions of estimates of health risk are remarkably broad; and most of that breadth is due to our fundamental uncertainty about the health effects of low-level exposures to environmental chemicals, not to variations in people's exposures. The high ends of a risk distribution are driven primarily by "pessimistic" interpretations of, but consistent with, the dose-response data. These data typically derive from over-exposed rodents whose responses may or may not predict human responses in the situation under analysis.

- **Central tendency, mean or median estimates of risk are unlikely to provide a full, useful basis for public health decision-making. One really needs the full distribution. However, a properly derived 95th percentile estimate of risk, supplemented with mean and median estimates, may provide a set of three bottom lines that can indeed be a basis for sound public policy.** There is no single estimator of risk appropriate to all situations, and the *definition* of the estimator

matters greatly. Further, no matter what estimator of risk might be chosen, the estimate must be compared with some standard for decision-making, and that choice of standard is also crucial.

- **An entirely scientific risk assessment is a mirage. There is no single right way to do it.** Sound policy should indeed rest on sound science. But risk assessment is not and cannot be a wholly scientific undertaking. Risk assessment often turns upon details that are inherently unknowable. In general, probabilistic and holistic risk assessments could lead to improved decision-making. Whether such assessments prove to be more defensible than the *status quo* is harder to say.

2 Introduction

Current standard methods for risk assessments — typified by approaches taken by the U.S. EPA — are designed to overestimate risks and are reported in such a way that the extent of overestimation in the risk estimates is impossible to discern. The former property — that of producing an overestimate — is often desirable and/or mandatory, depending upon the context. The latter property — obscuring the extent of overestimation — is neither desirable nor mandatory.

In standard risk assessments, the overestimates arise from a number of causes, principally:

- the selection of models that are inherently conservative,
- the selection for those models of parameter estimates that are conservative, and
- the concatenation in the risk assessment of these conservative parameters without taking account of the probability for all such parameters to be simultaneously at the high end of their range.

Without a mechanism to account for the implausibility of the concatenation of extreme values — or to moderate the conservatism of each parameter during such a concatenation — one cannot evaluate the extent of the overestimate.

The case studies presented below illustrate risk assessment techniques that can take account of the probabilities for extreme values in parameter estimates, and so can meet some of the requirements of legislation under consideration by Congress. The relevant requirements are to provide (1) more complete uncertainty analyses; and (2) more complete descriptions of the ranges of risk estimates, including identification of "best estimates." In particular, the case studies presented below incorporate quantitatively the uncertainty or variability in each parameter used in estimating risks to human health, by using probabilistic techniques. The uncertain or variable parameters include those used to describe toxicity, model fate and transport, and estimate doses to exposed humans. The resulting estimates of risk include "best estimates," together with the surrounding uncertainty distributions. The report describes the hypothetical exposure circumstances, outlines the methods used to estimate risks, and explains the concomitant estimates of risks to human health.

3 Summary of methods

The main features of the case studies may be summarized by a set of selections that we made:

- We initially selected a hazardous waste site for which we had previously prepared a health risk assessment. This Superfund site will remain unidentified, since it is used solely to provide a realistic set of concentration measurements from a situation resembling that adopted for the case study. The site has been thoroughly studied, and provides ample measurements of concentrations of hazardous contaminants in well water.

- We selected a single environmental medium for study — groundwater providing a supply of drinking water. We selected groundwater because it is frequently of interest for health risk assessments and because there are important pathways from this medium leading to the three major routes of exposure. Two of the pathways are direct: contaminated groundwater that supplies domestic water will entail exposure to contaminants through (1) ingestion of drinking water and cooking water, and (2) dermal absorption of contaminants during showering and bathing. The third pathway, *via* air, leads to inhalation of vapors released from the groundwater, with the largest exposures occurring during showering.

- We initially selected a single chemical of concern for study — 1,1-dichloroethene (1,1-DCE, also known as 1,1-dichloroethylene or vinylidene chloride). We selected 1,1-DCE for two reasons: (1) the groundwater at the selected hazardous waste site had been well-characterized for concentrations of 1,1-DCE at fifteen sample sites, some of which presented interesting problems of interpretation, and (2) the ambiguous carcinogenicity of 1,1-DCE presents challenges to the risk assessor. The generic issues involved in estimating additional risks to health associated with exposure to 1,1-DCE at the representative site are thus germane to many other chemicals and scenarios.

- We selected a single health endpoint for risk characterization — cancer. We focus on lifetime incremental risk of cancer: (1) because it is usually the focus of environmental health risk assessments; (2) because of the challenges to risk assessors in selecting and applying appropriate models of cancer causation; and (3) because of the particular challenges of assessing the carcinogenic potency of 1,1-DCE in humans. 1,1-DCE has shown some carcinogenic potency in a small minority of bioassays performed in laboratory rodents.

- We evaluated two additional exposure scenarios to illustrate further the character of the distributions of risk estimates that may be produced. The first additional case study postulated exposures to vinyl chloride at waterborne concentrations equal to those for 1,1-DCE to illustrate the effect of a very different set of animal bioassay results. The second additional case study postulated much higher exposures to 1,1-

DCE — exposures that traditional point estimates of risk would unequivocally label as unacceptable.

- The results of the case studies presented here are all conditional on the correctness of the models used in them. Some of these, like the exposure models, are likely to be good representations of reality, and could in principle be checked against measurements (although the parameter values may have been mis-estimated). Other models are, notoriously, much more uncertain. Principal among these are the assumed dose-response model(s) and the interspecies extrapolation model, the latter of which cannot be checked against measurements.[2] To illustrate the potential effects of choice of a different model, we also re-evaluated the initial case study for 1,1-DCE with a dose-response model assumption that allowed for the possibility of beneficial effects — that is, that allowed for the possibility, consistent with the data, that low levels of 1,1-DCE might protect against cancer.

In these case studies it is infeasible to illustrate all the interesting and possibly controversial aspects of risk assessments incorporating uncertainty. For example, some of the aspects not examined here are:

- What measures of risk should be evaluated? We selected the lifetime incremental cancer risk to a randomly chosen individual. We made this choice because: (1) individual risk estimates are typically the basis of regulatory action; and (2) we could then treat variability among the population as an uncertainty for a random individual. Other risk measures are just as interesting and/or relevant, but may require very careful definition to ensure that "variability" and "uncertainty" can be adequately defined.

- Of the many bioassays testing the carcinogenicity of 1,1-DCE in laboratory animals, very few have produced any positive results. We have here used only those positive bioassays (one for ingestion exposures, one for inhalation exposures) used for constructing point estimates to evaluate some of the uncertainties in carcinogenic potency estimates; although it is feasible to incorporate all the bioassay results in a consistent manner. The latter approach is likely what is envisioned in the legislation: but it requires considerable extra effort and represents a substantial departure from the usual restriction to the most sensitive species, strain, sex, and organ/endpoint.

- In some wells, there are clear variations of concentration of 1,1-DCE with time. We have here assumed that the variation in measured concentrations in each well are due to random fluctuations, with the distribution of such fluctuations unchanging in time. One could obviously make other assumptions, such as obtaining estimates of concentrations from a groundwater model.

---

[2] One may have to worry about two dose-response models (one for rodents and one for humans), together with an interspecies extrapolation model — see Appendix B. All of these can be checked only at high doses.

- Each of the case studies examined is for a single chemical. The evaluation of multiple simultaneous chemical exposures is not in principal very different (providing the effects of the chemicals do not interact), but the differences from the point estimates can become more exaggerated (see Crouch *et al.*, 1983). The point estimates of risk for multiple chemicals correspond to the addition of several extreme percentiles of the distributions for individual chemicals, while the probabilistic approach correctly takes account of the likelihoods for the various values.

3.1 The risk measure selected

Many measures of risk could be computed for the situation described. A probabilistic risk assessment incorporates uncertainty in the parameters of the exposure models, uncertainty in the potency of each contaminant, and variability among the individuals for whom risks are calculated. The treatment of these uncertainties and variabilities in the risk assessment depends on the risk measure(s) that are being assessed. To simplify our task, we have chosen to evaluate estimates of risk for a randomly chosen individual from the exposed population. The effect of this choice is that, by definition, all the variability distributions for characteristics of individuals within a population become uncertainty distributions when contributing to this measure of risk. *For this measure of risk* there can be no variability — only uncertainty.

It may be argued that this measure of risk is not the most useful one, or the one that is generally called for, or the one that is "obvious" for this situation, and myriad other objections to it may be raised. We agree. In general, many distinct measures of risk are, or should be, required to fully characterize that risk; and different measures may be important to different people, at different times for the same person, or for different policy or information purposes. It is essential that the correct measure(s) of risk be evaluated for any particular purpose, and a specification of the measure(s) generally required should perhaps be a principal feature of any new legislation.

For the purposes of this document, however, the measure chosen is adequate as an illustration. The conclusions we draw are strictly valid for this measure of risk, for the situation under analysis. It is likely that some or all of these conclusions may be extrapolated to other measures of risk and/or for other situations, since the conclusions are based on generic properties of the results, but there are plausible situations for which some of the conclusions may be invalid for some measures of risk.

In this risk assessment of 1,1-DCE, therefore, the result of the probabilistic analysis is a distribution of incremental carcinogenic risk estimates for an individual:

- randomly selected from the U.S. population (because of our assumption that the exposed population has characteristics identical to the U.S. population);
- exposed to a concentration of 1,1-DCE randomly selected from the 1,1-DCE concentrations in the fifteen wells at the site (because of our assumption that each well serves an identical population); and
- under conditions corresponding to our hypothetical exposure situation.

We reiterate that for an individual randomly selected from the U.S. population, the uncertainty in personal characteristics such as average shower duration, water ingestion rate, and weight is the same as the variability of these characteristics in the population as a whole. Because the risk assessment is calculating risks associated with individuals, all distributions used in the risk assessment represent uncertainty distributions for an individual, and a distinction need not be made between distributions resulting from uncertainty in parameters and those resulting from variability among individuals.

3.2 Approach

The approach taken here is to adhere to the usual structure of risk assessments: hazard identification, exposure assessment, dose-response assessment, and risk characterization. Most of the descriptions are abbreviated from what would be expected in a complete risk assessment, since the situation has been artificially constructed and such details are not the major focus of the effort. In each section, however, we detail the typical approach that would be used (the "point estimate"), and then explain the modifications required for a probabilistic approach. Several of the more technical details have been relegated to the appendices. Appendix A.1 gives details of the practical implementation of simple random number distributions, while the rest of Appendix A details the more complex distributions, together with descriptions of the data whence they are derived. Appendix B is devoted to the evaluation of carcinogenic potencies.

The result we seek — a risk estimate — involves the combination of many parameters, each of which has an associated probability distribution. While the models involved can usually be reduced to fairly simple mathematical forms, it is generally intractable to obtain analytically the probability distribution for the final result. The general problem is that of performing a multi-dimensional integral over the support of all the distributions involved. The method we have used for this multi-dimensional integration is the Monte Carlo method, with the precise steps involved detailed in the risk characterization (Section 7.2.1).

4 Hazard Identification

4.1 General description

The process of hazard identification typically: (1) begins with a description of the site and its surroundings, (2) goes on to identify the chemicals that may pose a hazard, (3) describes the configurations of the sources containing those chemicals, and (4) is completed by a summary of the concentrations in the sources of the chemicals to be included in the exposure assessment. Because of the nature of this risk assessment example, however, only a limited amount of detail need be provided.

The site contained a manufacturing facility that made use of various chlorinated organic solvents, and some of these solvents were discarded in a way that resulted in contamination of an aquifer that supplied domestic water to nearby residents.

Concentrations of chlorinated organics in groundwater at fifteen wells were measured over periods of two to six years, depending on the well. Concentrations of 1,1-DCE showed moderate variability across wells and, in some wells, through time. Concentrations at some times in some wells were below detection limits. It is possible to write plausible statistical models to account for the observations — for example, one might consider the statistical model:

$$C(t) = \mu + \beta_{time} t + \epsilon \qquad (1)$$

where

C	is the concentration measured at time t,
μ, β_{time}	are parameters of the statistical model, and
ϵ	is a random error term.

The error term ϵ is defined by the requirement that at each time t, $C(t)$ is a random variate with a distribution that is a truncated normal, the normal having mean $\mu + \beta_{time} t$ and standard deviation σ, with the truncation being at zero concentration. Table 2 gives maximum likelihood estimates of the μ, β_{time}, σ, and a test of the hypothesis that β_{time} is non-zero.[3] A better approach would be to use groundwater models to indicate the time variation of concentration to be expected in well waters; but for the current analysis, it was assumed that the long-term average concentration in each well was independent of time, the observed variations being fluctuations about this long-term average.

[3] We originally intended to perform an assessment where concentrations varied with time according to some simple model (although probably not this one). Other analyses were subsequently deemed more important and enlightening.

In a situation such as that postulated for exposure to the water from these wells,[4] current EPA guidance calls for estimating a (lifetime average) concentration estimate for a reasonably maximally exposed (RME) individual.[5] That concentration estimate is generally taken as an upper 95th percentile estimate for the concentration to which any individual may be exposed. For the hypothetical exposure situation at this site, 1/15 of the exposed population is exposed to each well, so that one should perhaps choose the highest mean concentration as the point estimate, or perhaps an upper confidence limit on that highest mean estimate. What would probably be done in a standard style assessment is indicated in Table 3, which lists the mean concentrations for each well, together with the grand mean of the means (average) and the standard deviation of the means (ignoring any uncertainty in each mean). In addition there is the typical calculation for "upper confidence limit." There are several policy decisions involved in such calculations:

- In calculating the mean concentration in each well, each measurement reported as non-detect has been treated as though the true concentration were 1/2 of the minimum detection limit. This policy usually has a small effect on estimates of average concentration, although in Well 15 (where all measurements were at or below the detection limit) it may have introduced a substantial bias.

- In evaluating the distribution of mean concentrations across wells, the uncertainties in each well's mean have been ignored, although such uncertainties could be taken into account with a more complex calculation. In this case, the effect is probably small.

- The upper confidence limit (UCL) concentration has been estimated as the grand mean (of means of well concentrations) plus the appropriate *t*–statistic (for a 95th percentile) multiplied by the estimated standard deviation of well means. For the UCL so calculated to be an accurate estimate of 95th percentile, there is an implicit assumption that the distribution of well means is normal, although it clearly is not normal in this case: a probability plot of the well means indicates that the distribution is closer to lognormal.

- Applying the "mean plus t-statistic times standard deviation" approach to this exposure situation is incorrect, since individuals are exposed (while in a single residence) to just one of the 15 wells.

---

[4] At the actual site, there was a substantial question as to whether anybody would in the future be exposed to the groundwater from the contaminated wells, since water from them had been replaced by water from other sources because of the contamination. We gloss over such essential details, as do most "baseline" risk assessments, by presuming exposure.

[5] The approach used by the EPA at the site whence came the measurements does not correspond to current guidance, so a direct comparison with that approach is not relevant here.

The UCL estimate in Table 3 is used as the "point estimate" for comparison with the results of the probabilistic assessment, since that is the value we think most likely to be used in an EPA risk assessment, despite its faults. Using the highest mean estimate of concentration in the wells would increase this to 19.74 ppb, and taking an upper confidence limit on this estimate of highest mean would increase the point estimate further still. Treating the distribution of mean values as lognormal (and assuming that somehow individuals could be exposed to concentrations higher than the mean in the worst case well) would give an estimate of approximately 27 ppb.

Table 2 Parameters of a time dependent model for concentrations of 1,1-DCE in wells

Well Number	Estimates for the statistical model of equation (1)			
	μ (ppb)	β_{time} (ppb/s)	σ (ppb)	p[a]
1	2.1	9.3e-04	1.2	0.836
2	1.1	4.2e-03	2.5	0.254
3	7.3	-5.3e-03	2.3	0.087
4	10.2	-1.5e-02	2.0	0.001
5	9.4	-7.9e-03	6.7	0.503
6	1.2	5.0e-03	2.7	0.254
7	26.0	-9.2e-03	7.9	0.017
8	0.0	1.1e-03	1.1	0.015
9	0.8	2.2e-05	0.3	0.857
10	0.0	8.2e-03	4.2	2.4e-07
11	17.3	-7.3e-03	14.7	0.056
12	19.0	-6.5e-03	7.5	0.009
13	0.9	-3.2e-04	0.3	0.182
14	0.7	-1.6e-04	0.3	0.505
15	0.2	4.3e-05	0.0	1.000

[a] probability of observing data under the null hypothesis that $\beta_{time} = 0$.

Table 3 Average concentrations of 1,1-DCE in wells

Well Number	Average Concentration (ppb)	Well Number	Average Concentration (ppb)
1	2.59	10	5.92
2	3.46	11	16.00
3	5.20	12	15.68
4	4.56	13	0.78
5	7.64	14	0.65
6	3.86	15	0.50
7	19.74	Average	5.90
8	1.13	Standard Dev.	6.25
9	0.85	UCL	16.91

For the probabilistic assessment, the distribution of average concentrations to which a random individual may be exposed was taken to be the distribution of average concentrations for the fifteen wells (*i.e.* each average well concentration was selected with equal probability). The averages used were those given in Table 3, so that once again the policy decision has been made to treat non-detects as 1/2 the detection limit. In addition, the uncertainty for each mean has been ignored (since it is small compared with the variability, and the variability of concentrations is treated as an uncertainty in this assessment).

4.2 Summary of source concentrations

The source concentrations used for the point estimate, and for the probabilistic assessment, are shown in Table 4.

Table 4 Summary of source concentrations

Source: Groundwater		
Chemical	Point estimate of concentration	Probabilistic estimate of concentration
1,1-DCE	16.91 ppb	15 well mean concentrations from Table 3, with equal probability

5 Exposure assessment

Exposure assessment ordinarily begins with fate and transport analyses. In this case, we assume that concentrations in well water are equivalent to concentrations in tap water, so that there is no other fate and transport modeling required. In general, modelling fate and transport within a water distribution system can be an extraordinarily complicated and uncertain undertaking. At the site supplying the concentration measurements, the 15 wells using the contaminated aquifer fed at various points into a distribution system that was also supplied from an uncontaminated aquifer at different points. Tap concentrations (some of which were also measured) thus did not necessarily correspond to well concentrations. Provided, however, that fate and transport can be modeled (as is required for point estimates), the probabilistic techniques used here could be applied to fate and transport analyses also.

For each exposure route included here, we describe below the mathematical representation of the physical models used to estimate doses and the distributions of parameter values that will be substituted in the equations to accomplish the probabilistic assessment. The point estimates for the parameters of these models are those recommended by U.S. EPA in its current guidance under Superfund (*e.g.*, EPA, 1989a).

5.1 Exposure by direct ingestion

The average lifetime exposure to 1,1-DCE through drinking contaminated tapwater (e_w, mg/kg-day) is estimated by:

$$e_w = \int_{t_1}^{t_2} \frac{C_w \cdot \gamma_w \cdot f_w}{BT}\, dt \qquad (2)$$

where the terms are:

e_w	lifetime average dose received through direct ingestion of water (mg/kg-day),
C_w	concentration of 1,1-DCE in drinking water (mg/l),
γ_w	consumption rate of tap water (l/day) at age t (yrs),
f_w	fraction of tap water consumed at home,
B	body weight (kg) at age t (yrs),
T	lifetime (yrs), and
t_1, t_2	the age range (yrs) of exposure.

The physical model is here the trivial one of expressing tap water intake consumed in the home as a fraction of body weight at each age during which exposure occurs, averaging over a lifetime by integrating with respect to age and dividing by total lifetime, and assuming that the concentration of 1,1-DCE in consumed tap water is identical to that at the tap. This last assumption is almost certainly incorrect — 1,1-DCE is sufficiently volatile that a considerable fraction will evaporate from tap water used for beverages and food — so that this model is itself inherently conservative.

5.1.1 Point estimate for ingestion

Tapwater consumption was taken to be the EPA-recommended value of 2 liters/day, independent of age (by policy, this is a standard value). The fraction f_w was set to unity (conservative upper bound estimate). Body weight was assumed to be 70 kg, and lifetime 70 years (both policy standards). The age range of exposure was defined simply as being any 30 year period ($t_2 - t_1 = 30$ years), since all terms in the integral in equation (2) are independent of age using the standard point estimate default values. The 30-year exposure period is a policy standard.

5.1.2 Probabilistic estimate for ingestion

Contaminated tap water intake (as a function of age and sex) was approximated by using total tap water intake (Appendix A.6), including direct and indirect (*e.g.* in beverages and cooking) uses of tapwater, although the values given are probably now overestimates (Appendix A.6.3). The fraction of tap water consumed at home was arbitrarily assumed to be a uniform (0.5,1) distribution, to take some account of the consumption of large fractions of their liquids at other locations by large fractions of the population. The distribution of body weight at each age was taken to be lognormal, with parameters given in Appendix Section A.5. The selection of the age range t_1, t_2 and the lifetime T for each sample of the Monte Carlo assessment are described in Section 7.2.1, since they are common to all the models. The duration of exposure ($t_2 - t_1$) was always arranged to be the residence period in a household, corresponding to the assumption that anybody moving would move far enough away to be no longer supplied by the same water system. The integration in equation (2) was approximated by summation at 1 year intervals.

The integral in equation (2) has to be applied to individuals, so that the variations with age of its various terms are required for individuals, not for populations (as are given in Appendix A). These variations were approximated by assuming that individuals remain at the same percentiles on the relevant distribution at all ages. Thus, for example, at an initial age selected as described in Section 7.2.1, a weight for an individual was obtained from the relevant age-sex-specific weight distribution, and the percentile on this distribution was noted. This individual's weight at other ages (as required in the integral in equation (2)) was then estimated by choosing the same percentile from the weight distribution at those other ages. The same approach applies for tap water consumption, and for fraction of tap water consumed at home.

5.2 Exposure by inhalation while showering

Various modeling efforts and measurements indicate that the majority of air exposures to volatile organic compounds (VOCs) present in tap water occurs during showers. When water containing VOCs such as 1,1-DCE is used for showering, the VOCs may volatilize from the water into surrounding air (Andelman, 1985). The amount of such transfer depends upon factors such as: duration of showering, volume of the shower stall and bathroom, ventilation rates and patterns of the shower and bathroom, water flow rate, shower head style, and water droplet size. Airborne concentrations depend also on chemical-specific

factors including: solubility in water, vapor pressure, Henry's law constant, and diffusion constants.

Simple physical models of the evaporation of VOCs at fairly high concentrations from shower water involve the factors just mentioned, and generally suggest a linear relation between air concentration and water concentration. More complex models should involve more complex phenomena, such as adsorption of the VOCs to walls and plastic components of showers, that become more important at low concentrations, and that may lead to more complex relationships between air concentration and water concentration (in particular, the slope of the air concentration versus water concentration relationship may be lower at low concentrations). In either case, attempting to predict air concentration from water concentration would require many parameters (such as ventilation rates, shower room sizes, and compositions of materials in shower stalls) that are specific to specific exposure circumstances, and that cannot be predicted for populations.

To estimate the average airborne concentrations of VOCs present during showering, we rely on experimental measurements by Jo *et al.* (1990). Jo *et al.* (1990) measured the concentration of chloroform (a chlorinated VOC similar to 1,1-DCE) in bathroom air after volunteers showered with water containing chloroform as a byproduct of chlorination. For each test, 10-minute air samples were collected from the breathing zone in the shower stall. There was limited ventilation between the shower stall and the rest of the bathroom — a shower curtain covered most of this space.

We adopt a simple approach in which we assume that the measurements of Jo *et al.* (1990) are representative of all individual exposures to VOCs, and assess the relationship between VOC concentration in air and VOC concentration in shower water (assumed to be identical to VOC concentration in tap water) by applying a statistical model to their measurements. The statistical model adopted is the simplest available for practical purposes (a quadratic relationship with zero constant term) that agrees with the theoretical ideas mentioned:

$$C_a = \beta C_w + \gamma C_w^2 \qquad (3)$$

where the terms are:

- C_a the shower air concentration (µg/m^3),
- β a constant coefficient, estimated as 2.07 ± 1.01 ℓ/m^3,
- γ a constant coefficient, estimated as 0.164 ± 0.035 ℓ^2/µg-m^3, and
- C_w the tap water concentration (µg/ℓ).

Our decision to select a quadratic correlation with zero constant term is a policy decision, based on theoretical ideas.

Application of the correlation derived from Jo *et al.* (1990) to the risk assessment requires the following assumptions:

- Differences in the shower size, ventilation characteristics, water temperature, water flow rate, and water droplet size between the Jo *et al.* experiment and other bathrooms will not significantly alter the correlation; and

- The volatilization of 1,1-DCE during shower use is similar to that of chloroform.

Lifetime average exposure to 1,1-DCE while showering was calculated from the 1,1-DCE concentration in shower air by averaging such exposure over a lifetime:

$$c_x = \int_{t_1}^{t_2} \frac{C_a f_d N}{T} dt \qquad (4)$$

where the terms are:

c_i	average 1,1-DCE concentration in air averaged over a lifetime (µg/m^3),
C_a	concentration of 1,1-DCE in shower air (µg/m^3),
f_d	fraction of the day spent in the shower,
N	number of showers per day,
t	age (yrs),
T	lifetime (yrs), and
t_1, t_2	the age range (yrs) of exposure.

5.2.1 Point estimate for inhalation

The best fit estimates for β and γ in equation (4) were used. Shower duration was taken to be 6.8 minutes (the average of the distribution of Appendix A.3), with 1 shower per day (a guess). Lifetime was fixed at the standard 70 years, and the age range was again taken to be any 30 year period — no terms in equation (4) depend on age, so that the integration is trivial.

5.2.2 Probabilistic estimate for inhalation

Normal uncertainty distributions were assumed for the parameters β and γ in equation (3), corresponding to the statistical model used to estimate them from the data, with parameters given below equation (3), except that these distributions were truncated at zero to ensure that positive concentrations of 1,1-DCE in water could never give negative concentrations in air. The distribution for shower duration (and hence the fraction the day spent in the shower) is given in Appendix A.3. The number of showers per day was assumed to have a triangular distribution with range 0.25 to 2 per day, and mode 1 per day. Treatment of lifetime T and age range t_1 to t_2 is described in Section 7.2.1.

As for the ingestion estimates, the variation with age for individuals was treated by assuming that all individuals stayed at the same percentile of each age-sex-specific distributions; although none of the distributions was assumed to vary with age for the terms in equation (4), so that the integral is trivial.

5.3 Exposure by dermal absorption while showering

The lifetime average absorbed dose of contaminants due to dermal absorption while showering can be evaluated using (EPA, 1989a):

$$e_d = \int_{t_1}^{t_2} \frac{C_w A \xi (24 f_d) N}{BT} dt \qquad (5)$$

where the terms are:

e_d	lifetime average dose received through dermal absorption while showering (mg/kg-day),
C_w	concentration of 1,1-DCE in shower water (µg/ℓ),
A	surface area (m^2),
ξ	skin permeability for 1,1-DCE (m/hr),
f_d	fraction of the day spent in the shower,
N	number of showers per day,
t	duration of exposure (yrs),
B	body weight (kg),
t_1, t_2	age range (yrs) of exposure, and
T	lifetime (yrs).

This equation may be derived from a physical model in which 1,1-DCE is assumed to permeate through the skin according to Fick's law, with the shower water washing over the whole of the body surface and containing a constant concentration of 1,1-DCE, while the concentration of 1,1-DCE in the interior of the body remains very much smaller than the concentration in the shower water.

Rough calculations show that the doses estimated using this model are small fractions of those occurring by other routes, so little time was spent on it. The model is, however, probably substantially in error, since measurements attempting to isolate the inhalation and dermal absorption components of doses during showering indicate that dermal absorption may contribute as much as 1/2 the total dose. More recent models of dermal absorption also predict much higher doses, but they appear to over-estimate.

5.3.1 Point estimate of dermal exposure

Human surface area was taken to be 1.94 m^2 (a measured average value for adults). Skin permeability for 1,1-DCE was taken to be 8×10^{-6} m/hr, the same as the permeability as water (EPA, 1988). Other parameters were as for inhalation, Section 5.2.1.

5.3.2 Probabilistic estimate of dermal exposure

Human surface area was taken to be age-dependent, with values given in Appendix A.8. Skin permeability for 1,1-DCE was taken to be 8×10^{-6} m/hr, the same as the permeability as water (EPA, 1988). Other parameters, and the approach, were as for inhalation (Section 5.2.2). The approach taken was substantially simpler here (surface area distributions as a function of weight and sex could be evaluated, for example), but since the calculated contribution of this route was small, less effort was directed to it.

5.4 Summary of parameter estimates for the exposure assessment

5.4.1 Point estimates

Table 5 summarizes the point estimates, and the sources of those estimates, that would most likely be used in a standard risk assessment and that are used for comparison here. The entries labelled with "Policy standard" are generally upper bound estimates that may have been derived from upper percentile estimates of some population parameters, but the distributions may not be relevant in the circumstances in which these "standards" are used.

Table 5 Parameters used to calculate risks under current U.S. EPA guidance

Symbol	Description	Value	Source
γ_w	Tap water consumption rate	2 ℓ/day	Policy standard
f_w	Fraction of water consumed at home	1	Upper bound
B	Body weight	70 kg	Policy standard
$t_2 - t_1$	Exposure duration (period of residence)	30 yrs	Policy standard
T	Lifetime	70 yrs	Policy standard
β	Coefficient in shower air concentration equation	2.07 ℓ/m^3	Fit to data
γ	Coefficient in shower air concentration equation	0.164 ℓ^2/μg-m^3	Fit to data
N	Number of showers taken per day	1	Guess
A	Body surface area	1.94 m^2	Measured average
ξ	Skin permeability for 1,1-DCE	8×10^{-6} m/hr	Based on water
d	Average shower duration	0.114 hr	Measured average

5.4.2 Probabilistic estimate

Table 6 summarizes the parameters used for the probabilistic estimate, and cross-references the details of the distributions used.

Table 6 Parameters used to calculate risks under current U.S. EPA guidance

Symbol	Description	Value	Source
γ_w	Tap water consumption rate	App. A.6	Measurements
f_w	Fraction of water consumed at home	App. A.6.4	Guess
B	Body weight	App. A.5	Measurements
$t_2 - t_1$	Exposure duration (period of residence)	See Sect. 7.2.1	
T	Lifetime		
β	Coefficient in shower air concentration equation	Section 5.2.2	Fit to data
γ	Coefficient in shower air concentration equation		Fit to data
N	Number of showers taken per day	App. A.3.3	Guess
A	Body surface area	App. A.8	50th percentile
ξ	Skin permeability for 1,1-DCE	8×10^{-6} m/hr	Based on water
d	Average shower duration	App. A.3	Measurements

6 Dose-response assessment

6.1 General approach

The following sections indicate the details of the evaluation of carcinogenic potencies and unit risks for 1,1-DCE and vinyl chloride. The methods are described in Appendix B for both point estimates of potency or unit risk, and probability distributions for these quantities. The point estimates correspond to those usually used by the EPA and published on-line in the Integrated Risk Information System (IRIS) or in the annual Health Effects Assessment Summary Tables (HEAST). The distributions are obtained using exactly the same models, and making minimal extensions to the methodology (even though a slightly more radical modification would be desirable — see Appendix B.3).

The computations for the standard analyses of both point estimates and distributions were made using MSTAGE 2.0 (Crouch, 1992). This program is designed to produce the required probability distributions in tabular form. In the Monte Carlo analysis, the distributions were handled by special-purpose software that read the tables and stored them as inverse cubic splines on an inverse lognormal scale.

For the non-standard analyses of 1,1-DCE described in Section 6.3, tables of the required distributions were obtained using the optimization routines available in Borland's Quattro Pro spreadsheet (similar routines are available in most competing products), and then again handled in the Monte-Carlo analysis by the special-purpose software. Other implementation issues are discussed in Section 11.3.

Throughout Section 6, various values are given to 4 significant figures solely to allow others to duplicate the analysis exactly. The true accuracy or reliability of most of the figures is approximately 10% (1 significant figure).

6.2 1,1-Dichloroethene (1,1-DCE), standard analyses

1,1-Dichloroethene has been assigned a carcinogen classification of C — possible human carcinogen — by EPA. Point estimates for the carcinogenic potency and unit risk of 1,1-DCE are given on IRIS (1995) as 0.6 kg-day/mg and 5.0×10^{-5} $m^3/\mu g$ respectively.

At the time of the EPA *Health Assessment Document for Vinylidene Chloride* (EPA, 1985a) there had been eighteen rodent bioassays reported, of which eleven were of inhalation exposure, five were of ingestion exposure, and two were of exposure topically or subcutaneously. Most of these studies were not carried out at maximum tolerated doses or exposures, and so had less than the possible maximum sensitivity. Just one of the inhalation studies yielded any positive results (according to the original authors), and the inhalation unit risk estimate is based on an analysis of this one study. None of the ingestion studies yielded any positive response (according to the original authors), with upper bounds on the possible potency estimate ranging from 0.16 kg-d/mg (when extrapolated to humans using the standard EPA surface area extrapolation) in a study using drinking water, to 0.6 kg-d/mg (extrapolated to humans) in a gavage study. The last — the highest upper bound — is the basis for the EPA ingestion potency estimate.

In what follows, we also use just the single studies selected by the EPA for its point estimates of carcinogenic potency or unit risk. Incorporation of the other studies in a probabilistic manner would have little effect on the *inhalation* unit risk estimates, since the other studies were generally of lower sensitivity than the study that produced a positive response. For the ingestion potency estimate, however, incorporation of the other studies might have a fairly large effect, since one drinking water study had a sensitivity four times higher than the study selected by the EPA as the basis for their potency estimate. Appendix B sketches how one may go about incorporating multiple bioassays to obtain a single distribution for carcinogenic potency or unit risk.

6.2.1 Carcinogenic potency of 1,1-DCE by ingestion

The data used, and EPA's analysis, are summarized in EPA (1985a) and IRIS (1995), and in Table 7. Analysis by the standard EPA multistage procedure (Appendix B, Section B.1) leads to an estimate of q_1^* in rats of 0.1072 kg-d/mg, and extrapolation to humans using the 1/3 power of body weight gives an estimate of human carcinogenic potency of Q_1^* of 0.5854 kg-d/mg. The value of Q_1^* published by EPA (EPA, 1985a; IRIS, 1995) is 0.6 kg-d/mg.

The dose-response curves assumed in the standard multistage procedure (maximum likelihood estimate, and that corresponding to q_1^*) are shown in Figure 1. The error bars around the experimental results are 95th percentiles on estimates of tumor probability, assuming binomial statistics. Application of the methodology described in Appendix B (Section B.2) leads to the uncertainty distribution (for experimental uncertainty alone) shown in Figure 2. The scale in Figure 2 incorporates the standard EPA extrapolation to humans, and so has a 95th percentile equal to the estimate of 0.5854 kg-d/mg. When incorporating the lognormal distribution for interspecies extrapolation described in Appendix B (Section B.8), the distribution for q_1 in rats is needed as the starting point, and this is equivalent to the distribution of Figure 2 with all potency values multiplied by the factor[6] $(0.43/70)^{1/3} = 0.1831$.

[6] This is the standard (cube root of relative body weight) factor.

Table 7 Experimental results for lifetime ingestion of 1,1-DCE by rats (NCI/NTP, 1982)

Experimental Dose mg/kg-d (by gavage, 5 d/wk, 104 wk)	Response Number of rats with tumor/ Number of rats examined
0	6/50
1	5/48
5	13/47

Period of Dosing: 104 wk
Length of animal lifetime: 111 wk
Tumor site and type: Adrenal pheochromocytomas
Animals: Male F344 rats, estimated average weight 0.43 kg

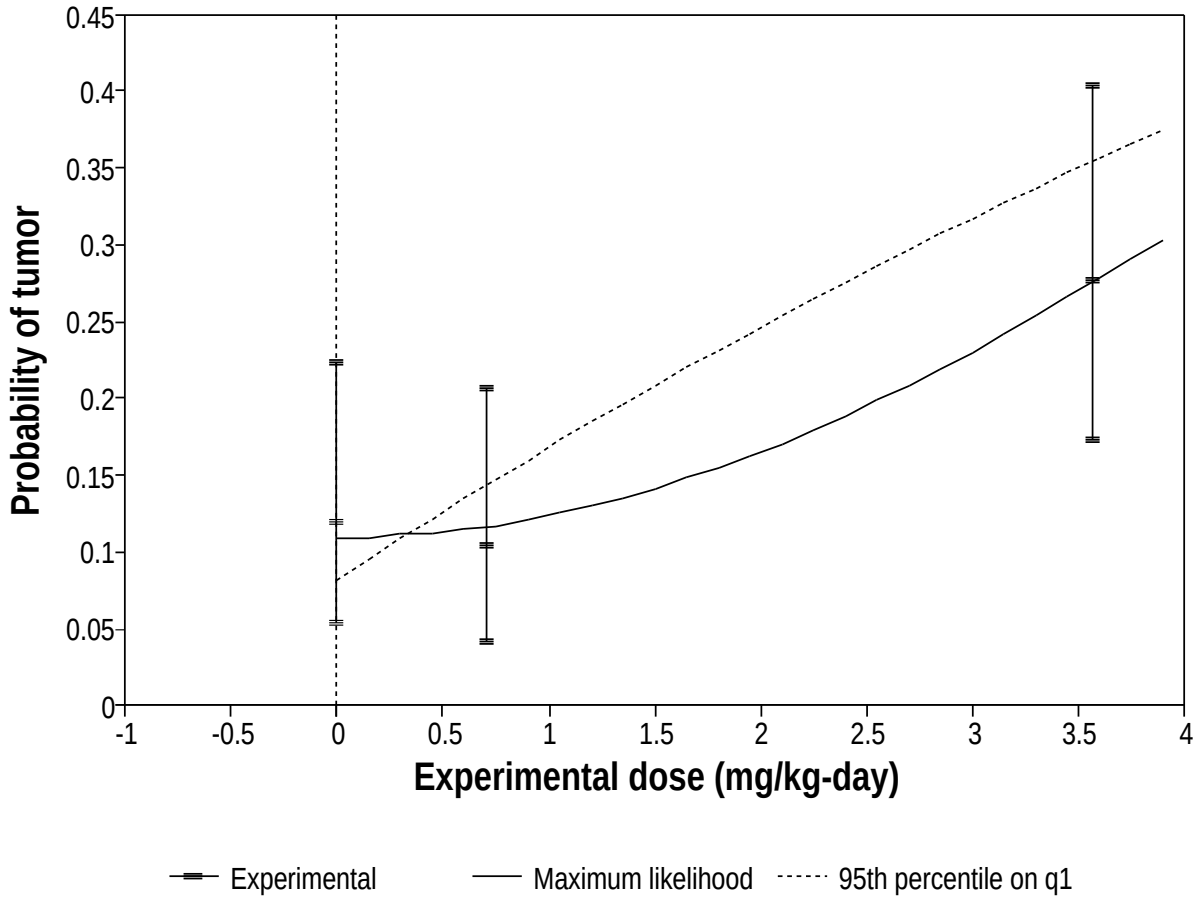

Figure 1 Dose-response data for 1,1-DCE ingestion (NCI/NTP, 1982: see Table 7).

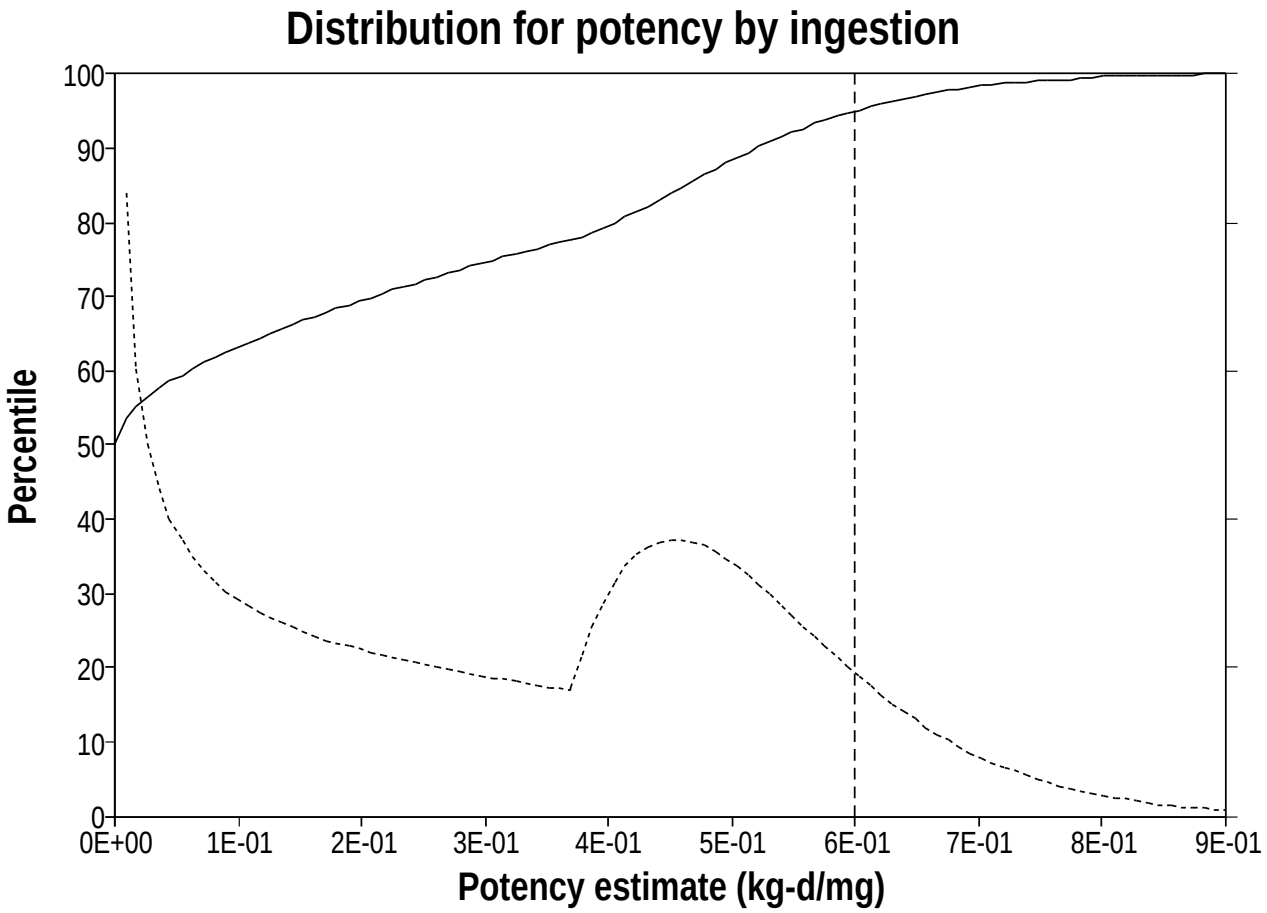

Figure 2 Cumulative (solid line) and differential (dashed line) distribution for the carcinogenic potency of 1,1-DCE by ingestion.

6.2.2 Inhalation unit risk for 1,1-DCE

The data used, and EPA's analysis, are summarized in EPA (1985a) and IRIS (1995), and in Table 8. Analysis by the standard EPA multistage procedure (Appendix B, Section B.1) leads to an estimate of unit risk in mice of 0.1959 ppm^{-1} or 4.846×10^{-5} $m^3/\mu g$. The EPA standard extrapolation to humans is unity on this dose measure, so that the EPA point estimate of carcinogenic potency in humans is also 4.846×10^{-5} $m^3/\mu g$. The value published by EPA (IRIS, 1995) is 5.0×10^{-5} $m^3/\mu g$.

The dose-response curves assumed in using the standard multistage procedure (maximum likelihood estimate, and upper 95th percentile on the linear term) are shown in Figure 3. Application of the methodology described in Appendix B (Section B.2) leads to the uncertainty distribution (for experimental uncertainty alone) shown in Figure 4. This figure incorporates the standard EPA extrapolation to humans, and so has a 95th percentile equal to 4.846×10^{-5} $m^3/\mu g$. When incorporating the lognormal distribution for interspecies extrapolation described in Appendix B (Section B.8), the distribution for unit risk based on equivalence of doses on a weight basis is needed as the starting point, and this is equivalent to the distribution of Figure 4 with all unit risk values multiplied by the factor[7] $(0.045/70)^{1/3} = 0.08631$.

Table 8 Experimental results for lifetime observation after 1 year exposure to 1,1-DCE by inhalation (Maltoni *et al.*, 1985)

Experimental exposure ppm 4 hr/day, 4 to 5 day/week, for 52 weeks	Response Number of mice with tumor/ Number of mice examined
0	0/126
10	0/25
25	28/119
Length of Experiment: 121 weeks Tumor site and type: Kidney adenocarcinomas Animals: Male Swiss mice, average weight 45 g (approx)	

[7] This is the standard (cube root of relative body weight) factor.

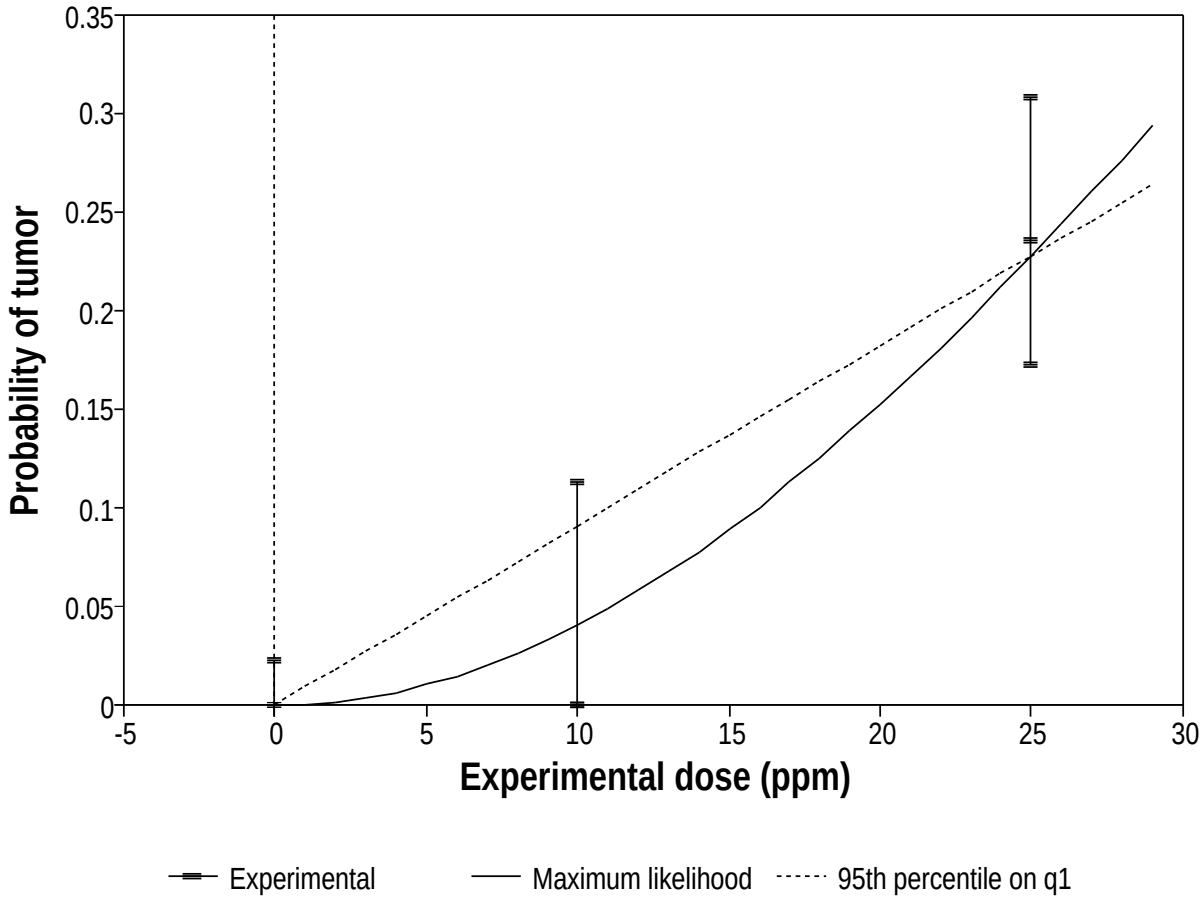

Figure 3 Dose-response curves for 1,1-DCE inhalation (Maltoni *et al.*, 1985: see Table 8).

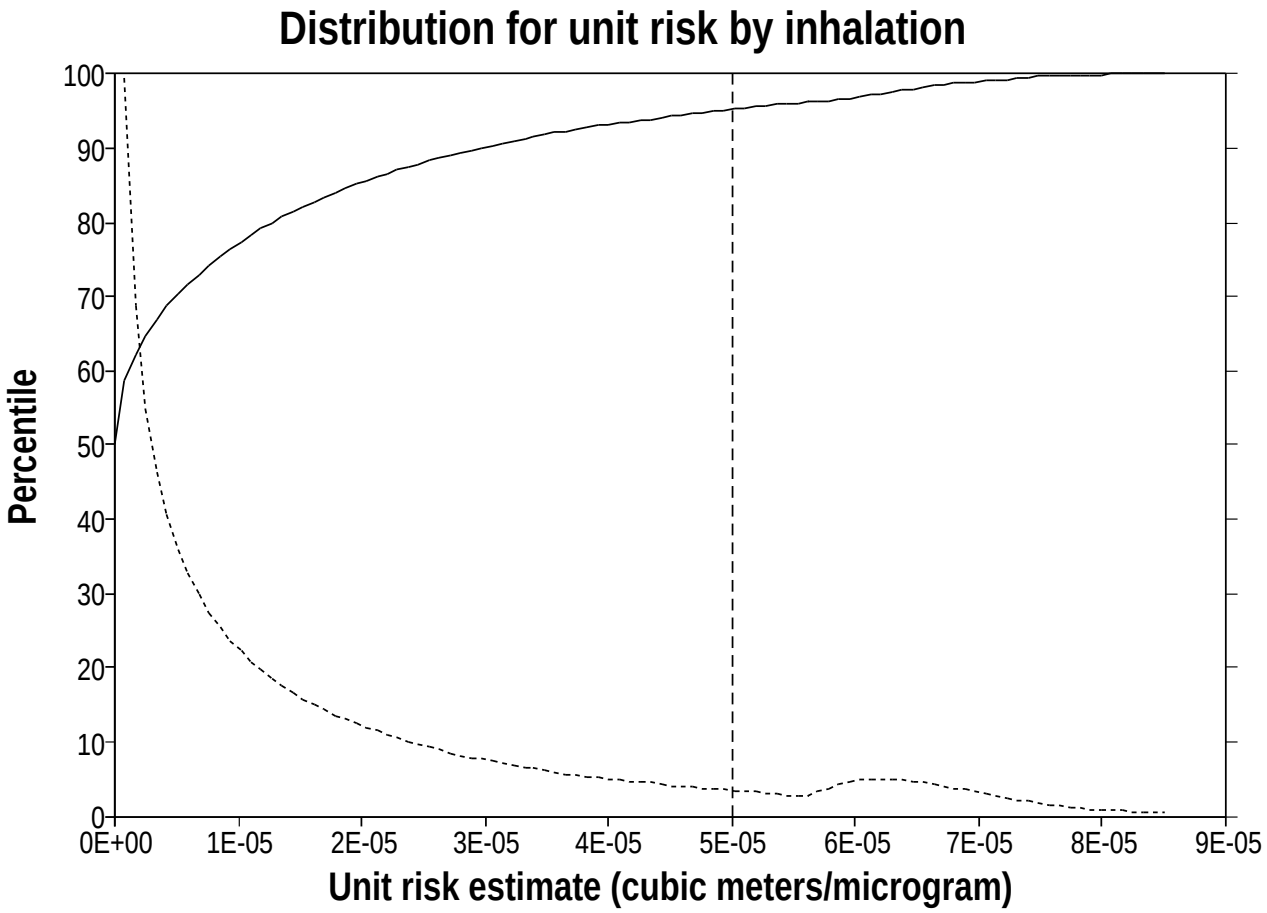

Figure 4 Cumulative (solid line) and differential (dashed line) distribution for inhalation unit risk for 1,1-DCE.

6.3 1,1-Dichloroethene, non-standard analyses

Section 6.2 has given details of the analysis of the bioassays of 1,1-DCE based on the standard dose-response (or exposure-response) curve currently in use. That dose-response model is of the form:

$$p = 1 - \exp(-q_\eta - q_\Theta d - q_\theta d^\theta \ldots - q_\gamma d^\gamma) \qquad (6)$$

where:

- p is the lifetime probability of tumor,
- d is the lifetime average dose rate of the material under test, and
- q_i $i = 0,1,\ldots,m$, are the $m + 1$ parameters of the model, which are subject to the conditions $q_i \geq 0$.

The effect of the positivity conditions on the parameters in this dose-response equation is to force the assumed probability of tumor to be non-decreasing as a function of dose. Indeed, if the probability of tumor increases at any dose, including the very high doses of animal bioassays, then it is necessarily assumed to be an increasing function of dose at all doses, including the very low doses of interest for human risk assessment. The effect of this assumption is to make a biased assessment of what might be happening at doses well below the range where experimentally observable responses occur — the possibility of a beneficial effect on response, for example, is completely denied.

In this section, we relax the constraint just discussed. Suppose that, at low enough doses, no assumption is made about the plausibility of increasing or decreasing response rates as a function of dose, so that a beneficial effect of low doses is considered as plausible as a deleterious effect. As in the standard approach, the simplest method of analysis is to assume some dose-response relationship, but one that does not contain all the positivity constraints generally used with equation (6). All that is logically required for equation (6) to be a dose-response relationship is that the polynomial term in the exponent of this equation remain non-negative, not that every term within the polynomial be non-negative. Therefore, we adopt equation (6) as the dose-response relationship, but with these relaxed constraints.

As a further simplification, we examine only the lowest dose tested (together with the control group) in order to remove to the extent possible the effect of what might happen at high doses, and restrict the polynomial in the exponent of equation (6) to two terms, so that we are effectively just fitting straight lines between two points (one of which is the zero-dose point) on the dose-response curve. Thus we examine the effect of assuming a dose-response curve of the form

$$p = 1 - \exp(-a_\eta - a_\Theta d)$$

where no restriction is placed on the sign of a_1, but we require that $a_0 + a_1 d$ be non-negative for all doses within the range examined (since $d = 0$ is always included, this requires that $a_0 \geq 0$).

With this dose-response curve, the approach described in Appendix B may also be used to obtain an uncertainty distribution (due to experimental uncertainties only) for the parameter a_1. Since a_1 is equivalent to q_1 in the standard analysis (except that it may be negative), it may also be called the potency (or unit risk, if the analysis is for inhalation exposure), and interpreted and used in the same way as the potency to evaluate the effect of exposures. In particular, we here assume that it may be extrapolated to humans in the same way as the standard potency estimate, and that the product of the extrapolated a_1 and human dose (or exposure) gives the increment (or decrement, if a_1 is negative) in risk of cancer for humans. That is, if a_1 is positive, the material tested is a carcinogen; if a_1 is negative, it is an anti-carcinogen.

Figure 5 and Figure 6 show the uncertainty distributions for the ingestion potency and unit risk respectively for 1,1-DCE, using the two lowest doses or exposures (including zero dose) from Table 7 and Table 8. These have been extrapolated to humans using the standard EPA approach, and so are directly comparable with Figure 2 and Figure 4.

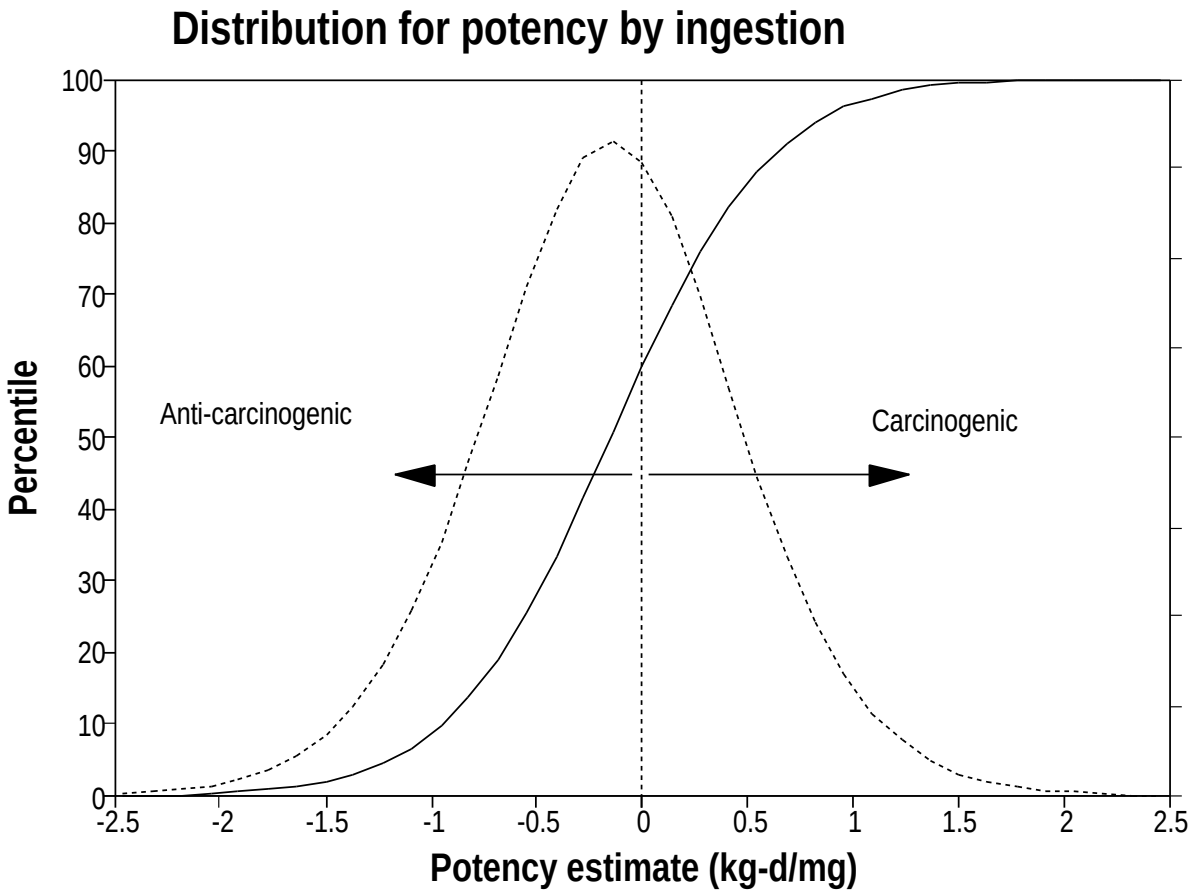

Figure 5 Cumulative (solid line) and differential (dashed line) distribution for the ingestion potency of 1,1-DCE using the non-standard approach, extrapolated to humans using the standard EPA method.

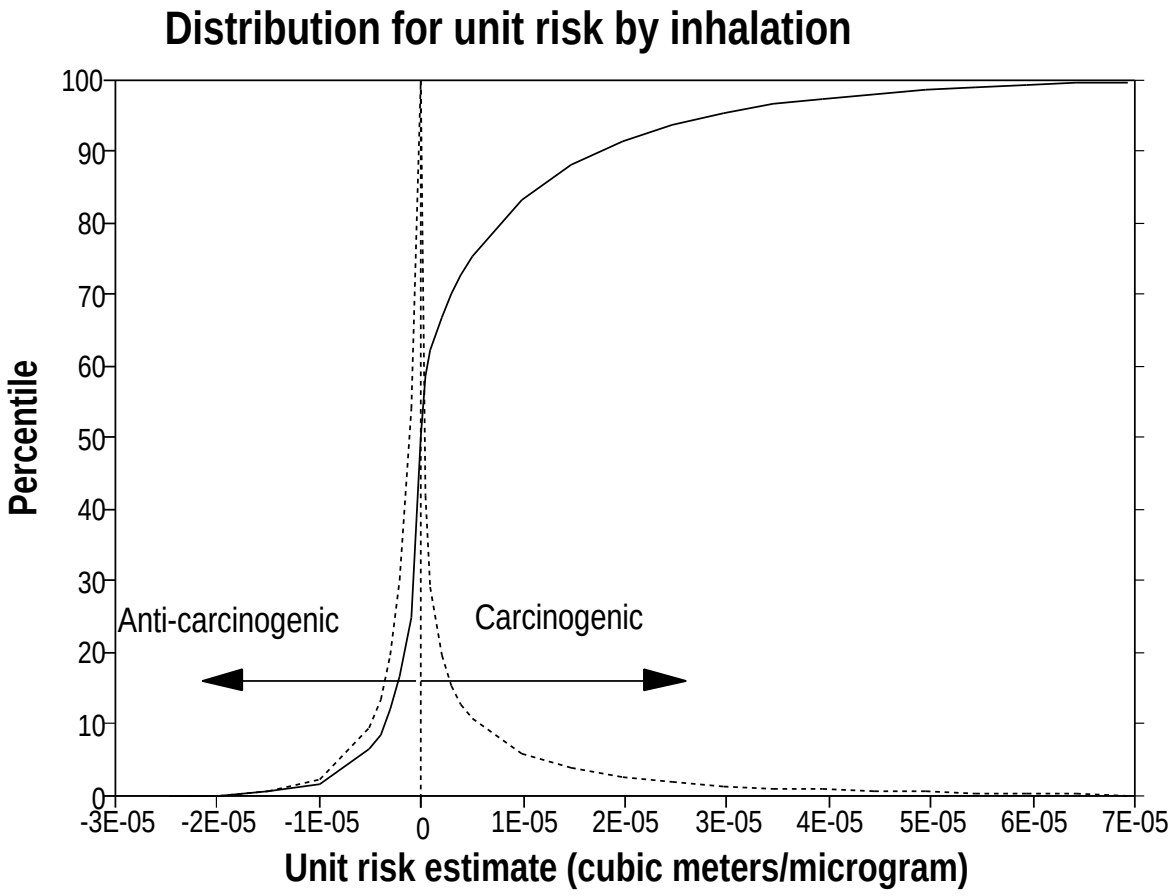

Figure 6 Cumulative (solid line) and differential (dashed line) distribution for the inhalation unit risk for 1,1-DCE using the non-standard approach, extrapolated to humans using the standard EPA method.

6.4 Vinyl chloride (standard analyses)

Vinyl chloride has been assigned a carcinogen classification of A, known human carcinogen. Point estimates for the carcinogenic potency and unit risk of vinyl chloride are provided in HEAST (1994). The values given are 1.9 kg-d/mg and 8.4×10^{-5} $m^3/\mu g$ respectively. These values are qualified as being under review, with the additional general comment:

> The most recently reviewed quantitative toxicity values listed here appear in EPA documents published in 1984 and 1985. The agency is aware that these values do not incorporate considerable information that is now available. The Office of Health and Environmental Assessment's position is that these toxicity values do not reflect state-of-the-art science for vinyl chloride. EPA now has individual animal data, not available when the oral unit risk was calculated, that may influence this value. Additional information that may be factored into a revised quantitative toxicity value includes data on increased sensitivity observed in young animals and data on metabolism/pharmacokinetics. A unit risk for air that considers information on young age exposure increases the risk (*i.e.* lowers the risk specific dose) by at least 3-fold. The consideration of metabolism pharmacokinetics will further increase the risk. One unpublished physiologically-based pharmacokinetic model prediction results in a 100-fold increased risk.

We here obtain uncertainty distributions for carcinogenic potency and unit risk, based on the same data as used by EPA to obtain the point estimates above. This approach specifically ignores any new data contributing to the general comment just cited.

6.4.1 Carcinogenic potency of vinyl chloride by ingestion

The data used, and EPA's analysis, are summarized in EPA (1985b). Feron *et al.* (1981) report a lifetime oral dosing study of rats, with the results summarized in Table 9. Analysis by the standard EPA multistage procedure (Appendix B, Section B.1) leads to an estimate of q_1^* in rats of 0.2942 kg-d/mg, and extrapolation to humans using the 1/3 power of body weight gives an estimate of human carcinogenic potency Q_1^* of 1.925 kg-d/mg. The value published by EPA (EPA, 1985b; HEAST, 1994) is 1.9 kg-d/mg.

The dose-response curves assumed in using the standard multistage procedure (maximum likelihood estimate, and to obtain q_1^*) are shown in Figure 7. Application of the methodology described in Appendix B (Section B.2) leads to the uncertainty distribution (for experimental uncertainty alone) shown in Figure 8. This figure incorporates the standard EPA extrapolation to humans, and so has a 95^{th} percentile equal to the estimate of 1.925 kg-d/mg. When incorporating the lognormal distribution for interspecies extrapolation described in Appendix B (Section B.8), the distribution for q_1 in rats is needed as the starting point, and this is equivalent to the distribution of Figure 8 with all potency values multiplied by the factor[8] $(0.25/70)^{1/3} = 0.1529$.

[8] This is the standard (cube root of relative body weight) factor.

Table 9 Experimental results for lifetime ingestion of vinyl chloride by rats (Feron *et al.*, 1981)

Experimental Dose mg/kg-d (daily, for life)	Response Number of rats with tumor/ Number of rats examined
0	2/57
1.8	26/58
5.6	42/59
17	56/57

Length of Experiment: 1001 days
Tumor site and type: Liver neoplastic nodules, hepatocellular carcinoma and angiosarcoma, or lung angiosarcoma
Animals: Female Wistar rats, assumed body weight 0.25 kg

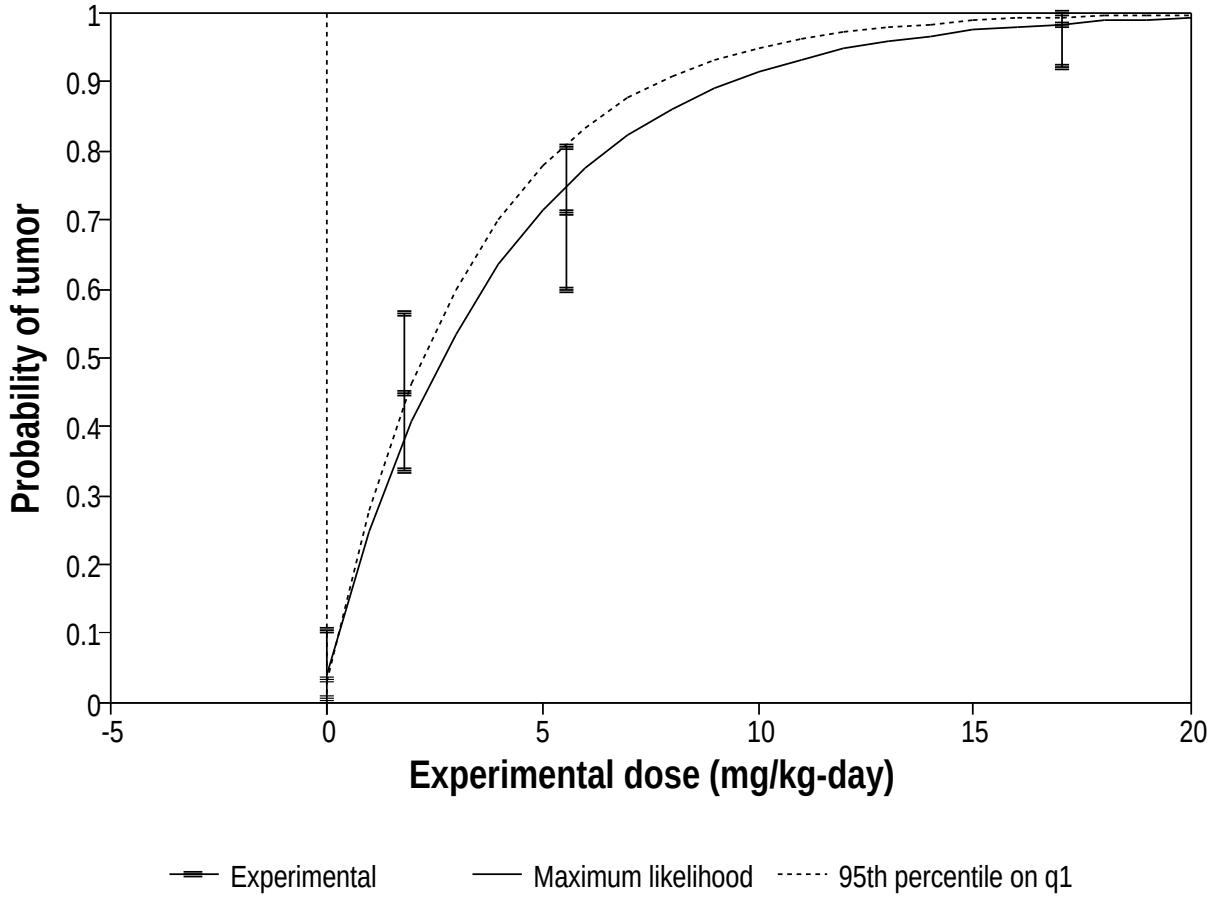

Figure 7 Dose-response data for vinyl chloride ingestion (Feron *et al.*, 1981: see Table 9).

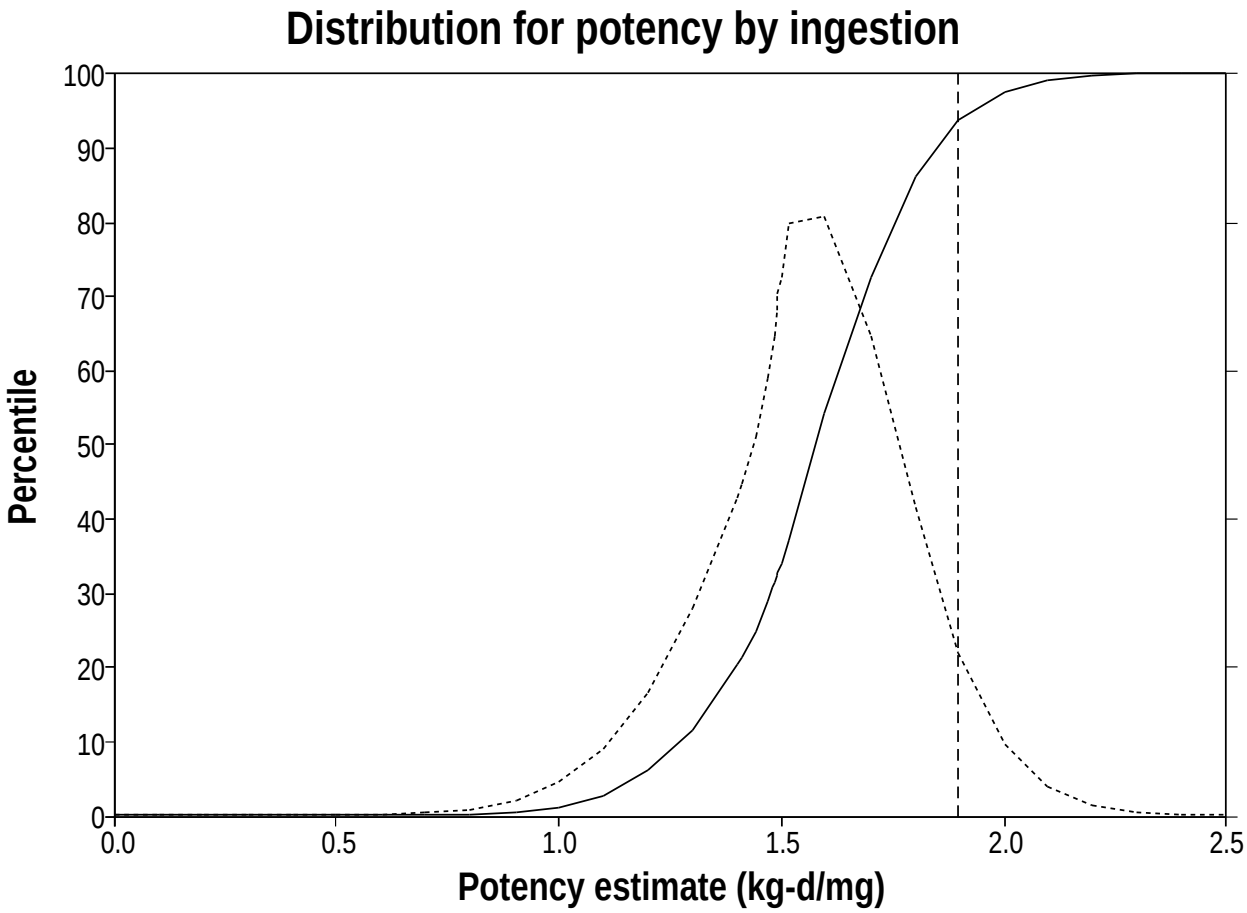

Figure 8 Cumulative (solid line) and differential (dashed line) distribution for the ingestion potency of vinyl chloride.

6.4.2 Unit risk of vinyl chloride by inhalation

The original data used by EPA to obtain a unit risk by inhalation are summarized in EPA (1985b). The methods used by EPA are summarized in EPA (1985b) and HEAST (1994). Unfortunately, the combination of two approaches has led to use of methods at variance with those that would now be used by EPA. The data used, shown in Table 10 and Figure 9, are a subset of those reported by Maltoni *et al.* (1980), who also gave results for 1 and 5 ppm exposures, and for several exposures above 50 ppm. Incorporating the experimental results from the lower exposures has little effect on the calculations that follow, and the higher doses are excluded by the methodology itself.

The approach taken by EPA was convoluted. An equivalent lifetime exposure concentration C_e was computed from the experimentally applied concentration C_a to take account of the intermittent and less-than-lifetime exposures:

$$C_e = f_1 C_a$$

where:

f_1 = fraction of lifetime exposed
= ((4 hours/day)/(24 hours/day)
× (5 days/week)/(7 days/week)
× (365 days exposure)/(1029 days lifetime)
= 0.04223.

Table 10 Experimental results for lifetime observation after 1 year exposure to vinyl chloride by inhalation (Maltoni *et al.*, 1980)

Experimental exposure ppm 4 hr/day, 5 day/week, for 1 year	Response Number of rats with tumor/ Number of rats examined
0	0/363
10	1/119
25	5/120
50	1/60
Length of Experiment: 1029 days Tumor site and type: Liver angiosarcoma Animals: Male and female Sprague-Dawley rats, assumed body weight 0.35 kg	

This equivalent lifetime exposure was then converted to an equivalent lifetime average absorbed dose *d* in mg/kg-d by:

$$d = \frac{\eta C_{a,m} V_r}{M_r}$$

where:

- η = 0.5 is an estimate of the fraction of inhaled vinyl chloride that is absorbed,
- V_r = 0.223 m^3/day is the assumed breathing rate of the rats, and
- M_r = 0.35 kg is the assumed body weight of the rats.

With such doses obtained from the exposures shown in Table 10, and the results shown there, a "potency" q_1^* in rats was obtained using the standard EPA procedure (see Appendix B, Section B.1). This potency may be represented as:

$$q_1^* = \frac{q^*}{d_m}$$

where:

- q^* is the result obtained from the standard EPA procedure applied to the results of Table 10, but with doses (exposures) normalized to a maximum value of unity, and
- d_m is the dose corresponding to the maximum exposure[9] ($C_{a,m}$ = 50 ppm = 1.278×10^5 µg/m^3).

This potency estimate in rats was then extrapolated to humans using the standard method:

$$Q_1^* = \left(\frac{M_h}{M_r} \right)^{1/3} q_1^*$$

where:

- Q_1^* is the estimate of potency in humans, and
- M_h = 70 kg is the standard human body weight.

Finally, to obtain the value in HEAST (1994), EPA applied a standard conversion, assuming a 70 kg person breathing 20 m^3/day, to obtain the estimated unit risk for humans *U* as

[9] Any exposure could be used here, so long as all exposures are normalized to it and the same exposure is used later in the analysis wherever $C_{a,m}$ is used.

$$U = \frac{V_h}{M_h} Q_0^*$$

where:

V_h = 20 m^3/day is the estimate of human breathing rate.

The net effect of this sequence of manipulations was to obtain the estimated unit risk U in humans as:

$$U = \frac{V_h}{V_r}\left(\frac{M_r}{M_h}\right)^{0.75} \frac{1}{\eta} \frac{q^*}{f_0 C_{risk}}$$

This expression (and hence the reported EPA unit risk of 8.4×10^{-5} m^3/µg, HEAST 1994) shows two differences from the current standard EPA approach to estimating a unit risk from inhalation data. First, it incorporates the term $1/\eta = 2$, accounting for absorption, which is strictly incorrect for a standard "unit risk" that relates external exposure to risk. Second, the term:

$$\frac{V_h}{V_r}\left(\frac{M_r}{M_h}\right)^{0.75}$$

which has the value 2.622 with the assumed values used for vinyl chloride, is generally taken to be unity in performing such extrapolations, since breathing rates are expected to extrapolate between species approximately as the 2/3 power of body weight. Thus the current standard approach for estimating unit risk would give:

$$U = \frac{q^*}{f_0 C_{risk}}$$

To maintain consistency with the published EPA point estimate, we have used that estimate as the point estimate, and normalized the uncertainty distribution for potency (experimental uncertainty) shown in Figure 10 so that its 95th percentile is equal to the EPA point estimate. To incorporate the interspecies extrapolation uncertainty (see Appendix B (Section B.8)), the distribution in Figure 10 is multiplied by a lognormal uncertainty distribution with median 1/30.67 and geometric standard deviation exp(2.2). The factor 30.67 comes from the extraneous factor 2 × 2.622 introduced by the EPA's manipulations, together with a factor of 5.848 (= $(M_h/M_r)^{1/3}$) to reduce the results to a body weight basis, as required by the approach of Appendix B.8.

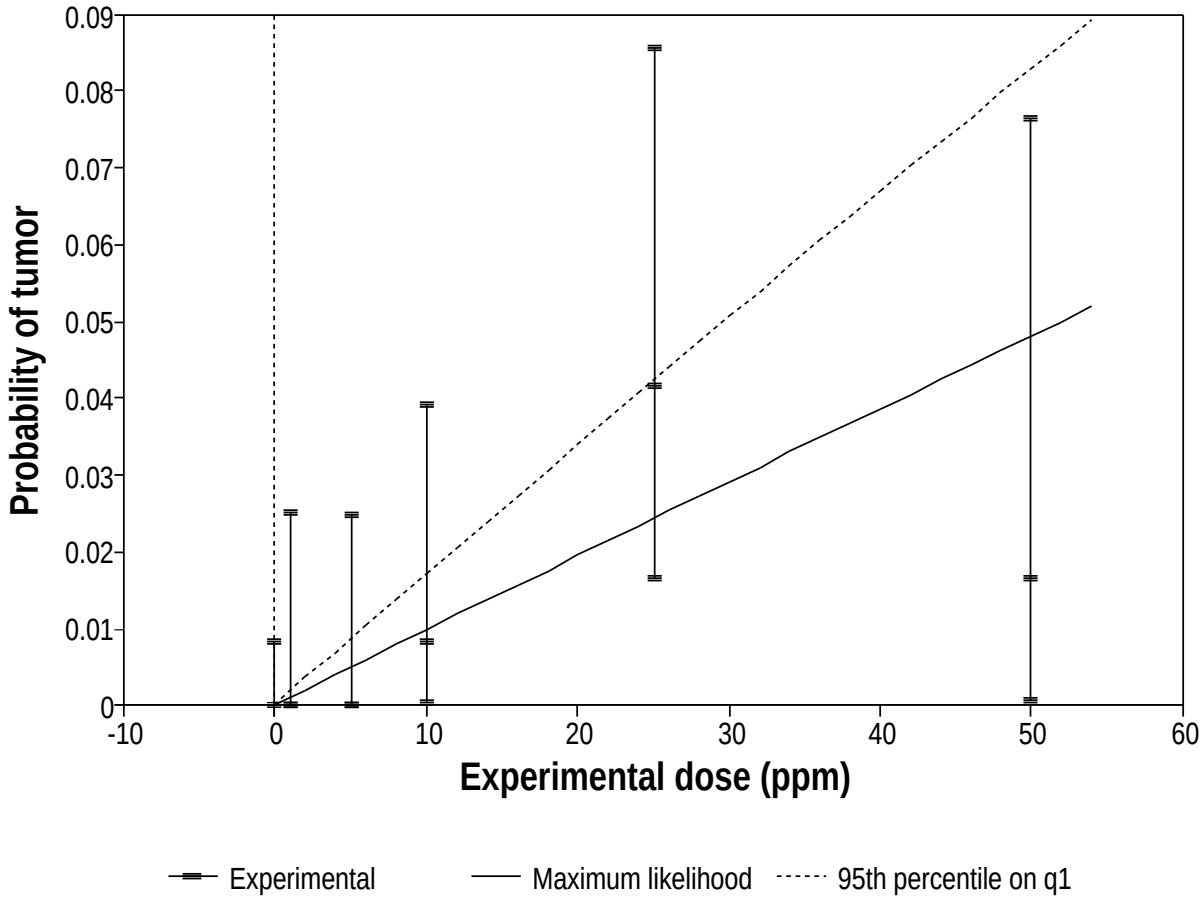

Figure 9 Dose-response data for vinyl chloride inhalation (Maltoni *et al.*, 1980: see Table 10).

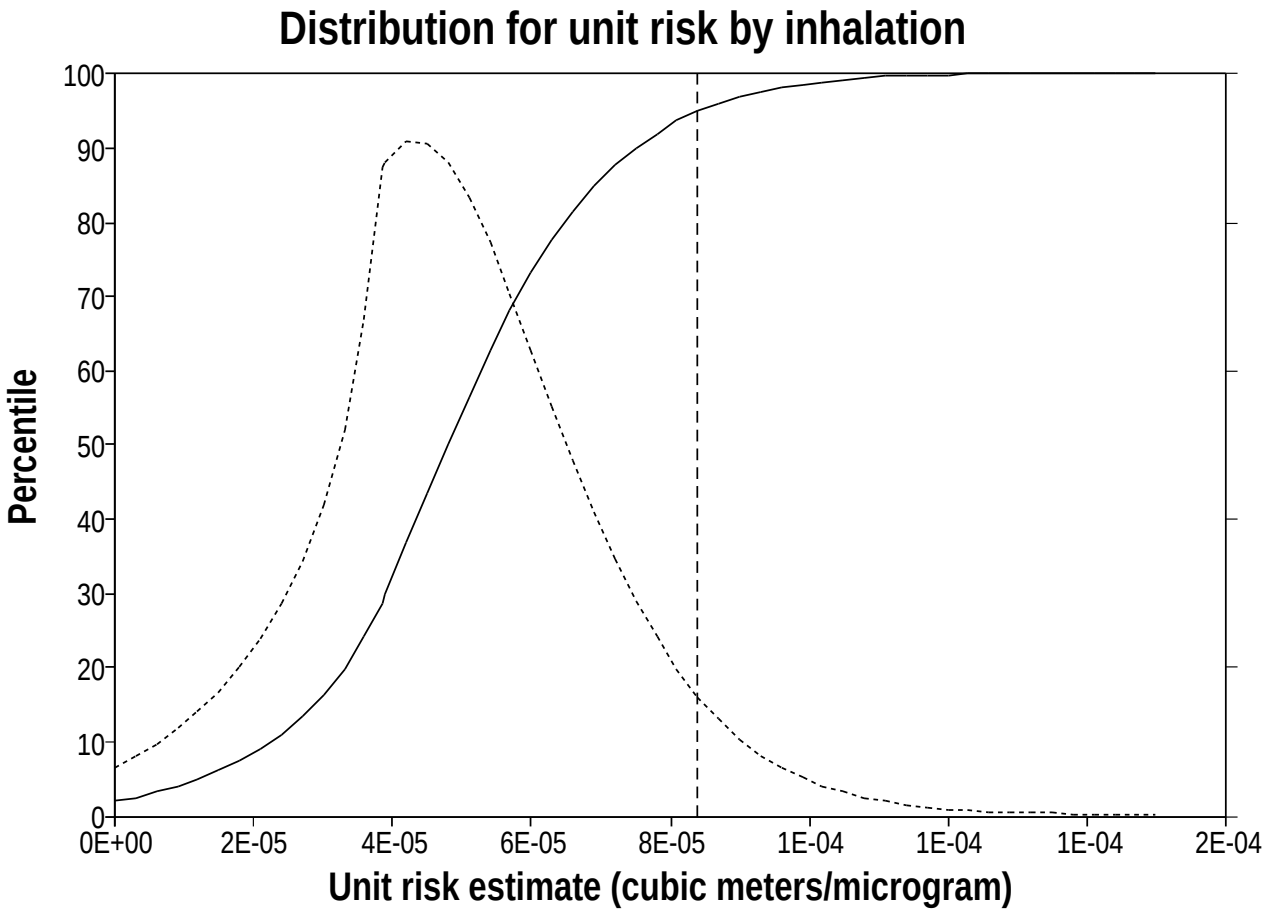

Figure 10 Cumulative (solid line) and differential (dashed line) distribution for the inhalation unit risk of vinyl chloride.

7 Risk Characterization

Once exposures and doses are calculated for an individual, the incremental carcinogenic risk for that individual is calculated using the equation:

$$R_c = e_w Q + c_i U + e_d Q \qquad (16)$$

where the terms are:

- R_c incremental carcinogenic risk,
- e_w lifetime average dose received through direct ingestion of water (mg/kg-day),
- c_i lifetime average air concentration (µg/m^3),
- e_d lifetime average dose received through dermal absorption (mg/kg-day),
- Q carcinogenic oral slope factor (kg-day/mg), and
- U inhalation unit risk (m^3/µg).

Substituting the doses and exposure (e_w, e_d, and c_i) obtained from Section 5, together with potency and unit risk estimates from Section 6 gives the risk estimate.

7.1 Point estimates of risk

Using the point estimates of doses and exposures detailed in Section 5 calculated using the point estimate of water concentration given in Section 4, together with the point estimates for carcinogenic potency and unit risk given in Section 6, yields the point estimate of incremental carcinogenic risk of 1.3×10^{-4}. This is the usual value that would be calculated, and the correct statement of its use would be something like:

> The calculated upper bound incremental carcinogenic risk estimate is 1.3×10^{-4}. It is likely that the true risk lies somewhere below this value, and may be zero.

The majority of this risk estimate is due to the direct ingestion of tap water, accounting for 93.6% of the total risk.

7.2 Probabilistic estimate of risk

To obtain the probabilistic estimate of risk, the probabilistic estimates given in Sections 4, 5, and 6 must be correctly combined. The method of combination used was the Monte Carlo procedure to be described (Section 7.2.1), the resulting distribution being given in Section 7.2.2.

7.2.1 Monte Carlo procedure

The Monte Carlo algorithm can be described by the pseudo-code (each entry is further described below):

```
repeat
     Select an exposure point concentration
     Select an age for an individual
     Select the sex of the individual
     Define the individual's lifetime
     Select the exposure duration
     Insert the individual into the exposure duration
     Select an average shower duration
     Select the number of showers per day
     Select a weight percentile for this individual
     Select a water ingestion rate percentile for this individual
     Select the fraction of water consumed at home for this
          individual
     Define the individual's surface area
     Select model parameters
     Select a potency and unit risk
     Select an interspecies extrapolation factor
     Integrate exposures over exposure duration
     Calculate incremental carcinogenic risk for the individual
     Store the incremental carcinogenic risk in a list of values
until 20,000 values have been stored.
     Sort the list of incremental carcinogenic risks
     Output the percentiles of the distribution of risk estimates
```

The individual operations in this pseudo-code are:

- Select the exposure point concentration.
  Randomly select, with equal probability, an exposure point concentration from the 15 well mean concentrations of 1,1-DCE in groundwater (Section 4.2).

- Select an age for the individual.
  Randomly select an age for the individual from the age distribution for the population of the United States (Appendix A.4).

- Select the sex of the individual.
  Randomly select the sex for the individual from the age-dependent sex distribution for the population of the United States (Appendix A.4).

- Define the individual's lifetime.
  The lifetime of the individual was set equal to the selected age plus the average life expectancy for individuals in the United States at that age (Appendix A.7). This approximation was deemed adequate for the current analysis. A better approach would be to use the age-sex-specific distribution for life expectancy at every age; but we suspect the extra complexity would not be worth the effort.

- Select the exposure duration.
  The exposure duration was assumed to be equal to the length of stay in a residence, corresponding to the assumption that any move would be to another residence that is not supplied by the same water system. For a large enough water system, this assumption might not apply, and a more complex model of the exposure situation might then be in order. Exposure duration was thus randomly selected from the distribution of household residence times (all residences) in the United States (Appendix A.2).

- Insert the individual into the exposure duration.
  The individual was randomly placed (a uniform distribution) at some point during the exposure duration. For example, if the age of the individual is selected to be 35 years and the duration of exposure was 5 years, the individual was assumed to be exactly age 35 at some randomly selected point during the 5 year period of exposure. If the period of exposure then started before the birth of the individual, the exposure period was truncated to the date of birth. Likewise, if the exposure period lasted beyond the death of the individual, the exposure period was truncated at the age of death. This procedure was selected somewhat arbitrarily, but no data appear to be available to design a better procedure.

- Select an average shower duration.
  Randomly select the shower duration (to be used in the shower inhalation and dermal absorption models) from the distribution of shower lengths (Appendix A.3).

- Select the number of showers per day.
  Randomly select the number of showers per day from a triangular distribution with a minimum value of 0.25, a maximum value of 2, and a peak at 1 shower per day (Appendix A.3).

- Select a weight percentile.
  Randomly select (uniformly) a percentile between 0 and 100 to be used for the individual's weight. A percentile is selected for the weight of an individual, and then the individual is assumed to remain at the same percentile for his or her entire lifetime, although the absolute value of his or her weight will change as he or she ages. The weight is calculated from sex, age, and the weight percentile (Appendix A.5).

- Select a water ingestion rate percentile.
  Randomly select (uniformly) a percentile between 0 and 100 to be used for the individual's tap water ingestion rate. Again, the individual is assumed to remain at the same percentile for his or her entire lifetime, although the absolute value of his or her tap water ingestion rate will change as he or she ages. The water ingestion rate is calculated from sex, age, and the water ingestion rate percentile (Appendix A.6).

- Select the fraction of water consumed at home.
  Randomly select the fraction of water consumed at home from a uniform distribution ranging from 0.5 to 1 (Section 5.1.2).

- Define the individual's surface area.
  Calculate the individual's surface area from the age and sex (Appendix A.8). We used the 50^{th} percentile estimate of body surface area where a more complete analysis would use distributions of surface area by age, weight, and sex (or more probably, by age and sex only). However, since surface area is only used in a minor exposure route, the failure to use the full distributions will have negligible effect on the result.

- Select model parameters.
  Randomly select values for the shower inhalation model parameters β and γ in equation (3) of Section 5.2 from the distributions for those parameters (Section 5.2.2). In the dermal absorption model, the permeability of 1,1-DCE across the skin was fixed at the value for water, 8×10^{-6} m/hr. Again, because dermal absorption is a small contributor according to the models used here, this failure to use an uncertainty distribution has only a small effect.

- Select a potency and unit risk.
  Randomly select a potency and unit risk from the uncertainty distributions for these values based on the bioassay data (Section 6).

- Select an interspecies extrapolation factor.
  Randomly select an interspecies extrapolation factor from the uncertainty distribution for the interspecies extrapolation factor (Appendix B). The same extrapolation factor was used for both inhalation unit risk and carcinogenic potency, corresponding to the assumption that no component of the uncertainty is associated with the route of exposure. There is no logical reason for this assumption, and this problem has not been examined empirically. For 1,1-DCE it is plausible that the mechanisms of action are independent of exposure route. For vinyl chloride, there appears to be a

substantial amount of metabolism in the rat lung, so that the mechanisms of action could be different by the different routes.

- Integrate exposures over exposure duration.
  After all parameters had been selected from their respective distributions, the exposure models were applied. Because certain characteristics of an individual change as the individual's age changes, total exposures from each exposure pathway were integrated over the periods of exposure (approximating the integrals by sums over annual intervals). The characteristics for individuals were changed with age by maintaining them at the same percentiles of their respective distributions.

- Calculate incremental carcinogenic risk for the individual.
  Equation (16) was used to calculate the incremental carcinogenic risk for the individual.

- Store the incremental carcinogenic risk in a list of values.
  Incremental carcinogenic risks for each individual were stored in a large array of values.

- Sort the list of incremental carcinogenic risks.
  The array of risk values was sorted from lowest to highest to obtain percentile estimates for the incremental carcinogenic risk.

- Output the percentiles of the distribution of risk estimates.
  The percentile values were printed at every 0.2 percentile (every 40 entries), and plotted graphically.

7.2.2 Probabilistic risk estimates, standard dose-response model

The procedure defined in Section 7.2.1 was performed with several variations to evaluate the effects of several of the uncertainty distributions on the resultant distribution for risk.

- All distributions were used exactly as specified in Section 7.2.1. Figure 11 shows the resulting distribution for risks to a random individual exposed to the contamination. This graph (and all subsequent graphs of distributions of risk estimates) also shows the expected value of the distribution (mean estimate) and the point estimate calculated as described in Section 7.1. The median estimate of risk is 1.2×10^{-9}, the mean estimate is 1.6×10^{-6}, the 95^{th} percentile estimate is at 1.7×10^{-6}, and the point estimate of 1.3×10^{-4} is above the 99.8^{th} percentile.

- Subsequent cases allow some of the parameters to be selected from their uncertainty distributions while holding others constant at their point estimate values. This allows an evaluation of the extent to which each distribution affects the final distribution of risk values.

Figure 12 shows the distribution of risks resulting from sampling the distribution for carcinogenic potency and unit risk for humans, while holding all other parameters constant at the value used in calculating the point estimate. The similarity with the overall distribution (Figure 11) shows that the uncertainty of carcinogenic potency and unit risk values is a major contributor to the overall uncertainty in estimated carcinogenic risks. The human potency/unit risk estimates are obtained by evaluating potency/unit risk estimates in laboratory animals (rats or mice), and multiplying by an interspecies conversion factor. Figure 13 shows the effect of sampling only from the uncertainty in animal potency/unit risk, while setting all other parameters to their point estimates (*i.e.* incorporating only the uncertainty due to the small numbers of animals tested). While this uncertainty is relatively large, it is clear that the majority arises from the interspecies factor.

To emphasize that most of the uncertainty is in the estimate of toxicity, Figure 14 through Figure 17 illustrate the uncertainty distributions due to other major parameters in the assessment (all other parameters were held at their point estimates to make these graphs). Figure 14 is the uncertainty due to the water concentration distribution (the fifteen steps in the distribution correspond to the fifteen separate well concentrations from which the exposure point concentration was selected), while Figure 15 shows the effect of all personal characteristics like age, sex, weight, water ingestion rate, and average shower duration. Finally, Figure 16 and Figure 17 (the same graph on different scales) show the minor uncertainty in this simple exposure situation of the small amount of fate and transport modeling — parameters such as β and γ in Equation (3).

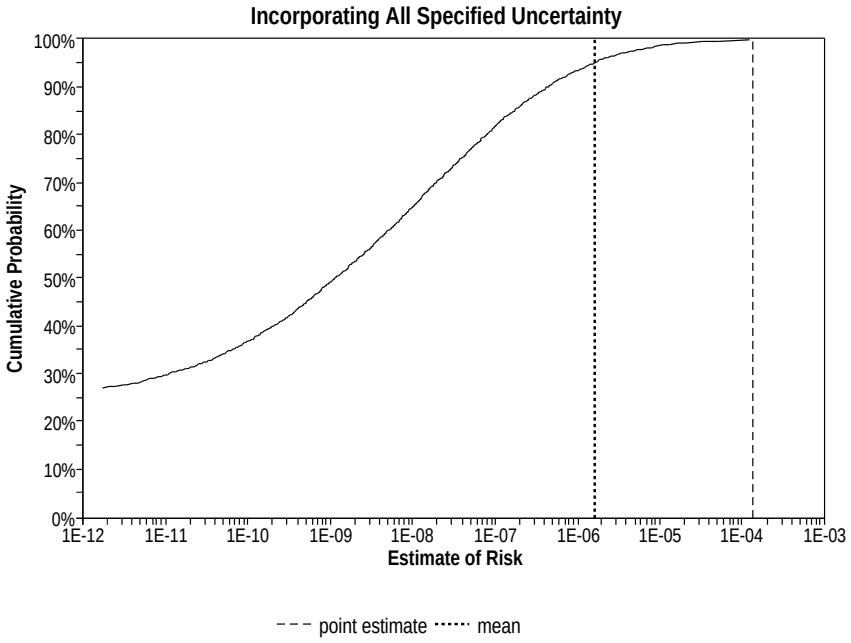

Figure 11 1,1-DCE: distribution of incremental carcinogenic risk: all uncertainties included.

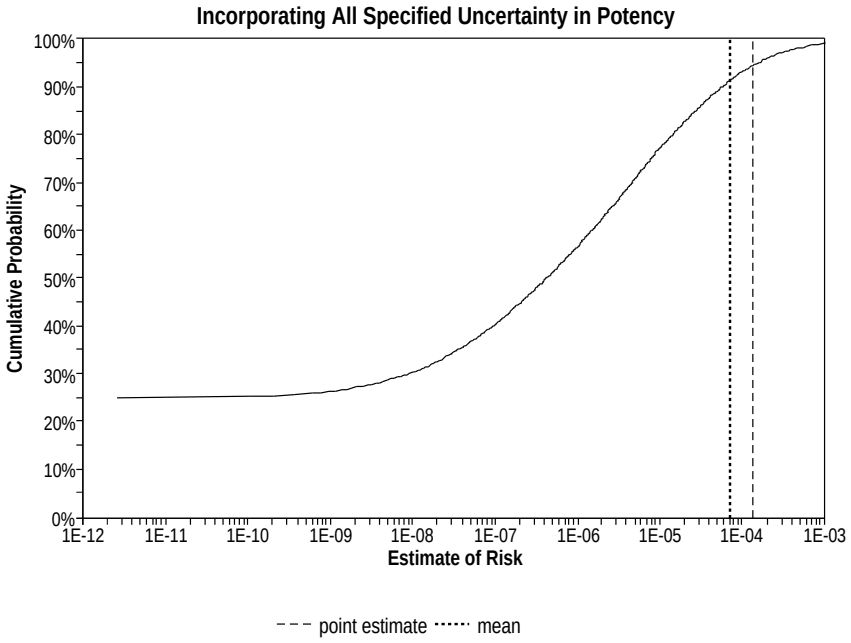

Figure 12 1,1-DCE: distribution of carcinogenic risk: uncertainty in carcinogenic potency values only.

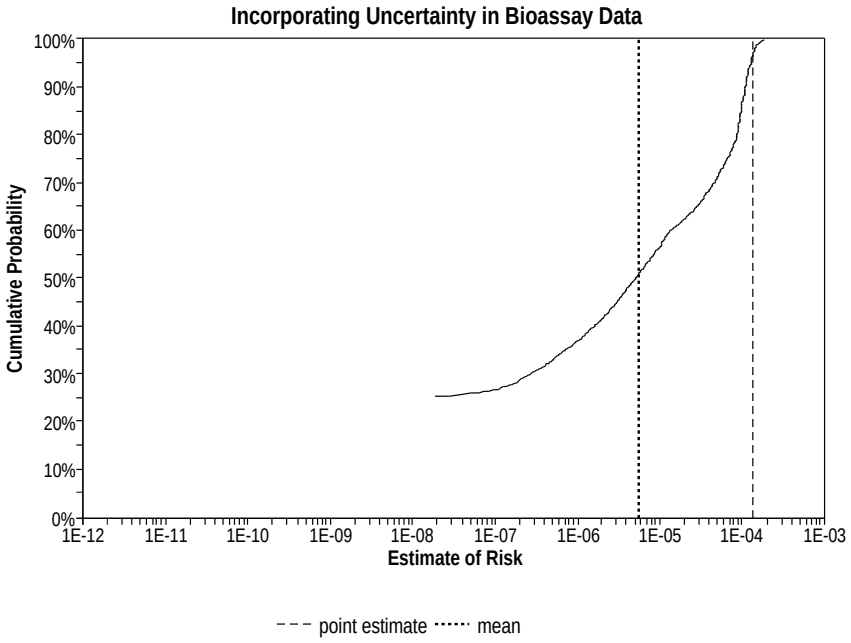

Figure 13 1,1-DCE: distribution of risk: uncertainty in bioassay analysis only.

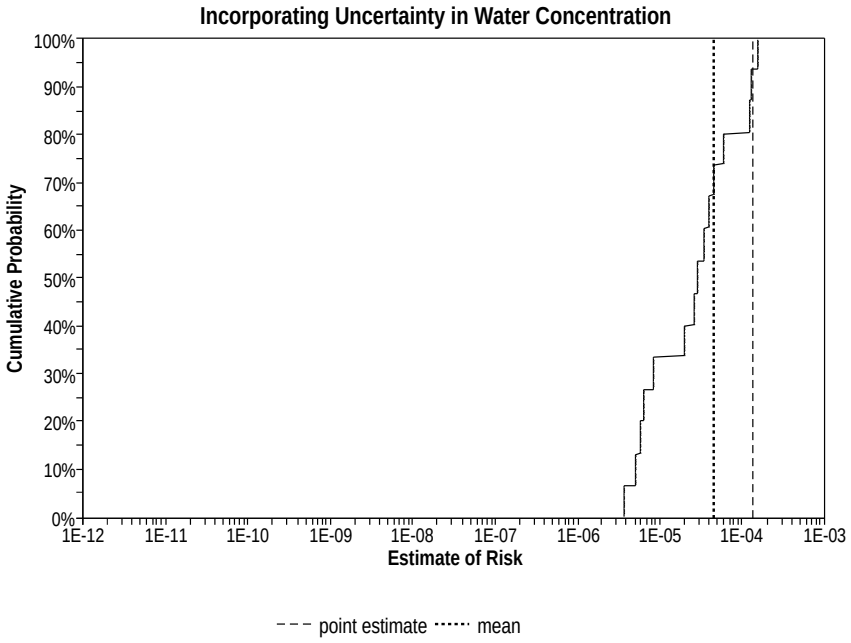

Figure 14 1,1-DCE: distribution of risk: uncertainty in water concentrations only.

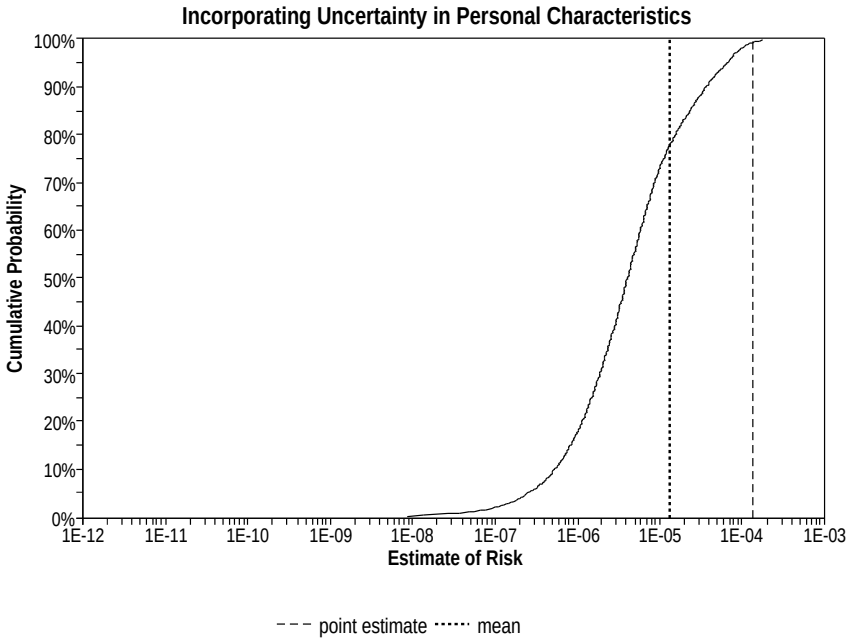

Figure 15 1,1-DCE: distribution of risk: uncertainty in individual characteristics only.

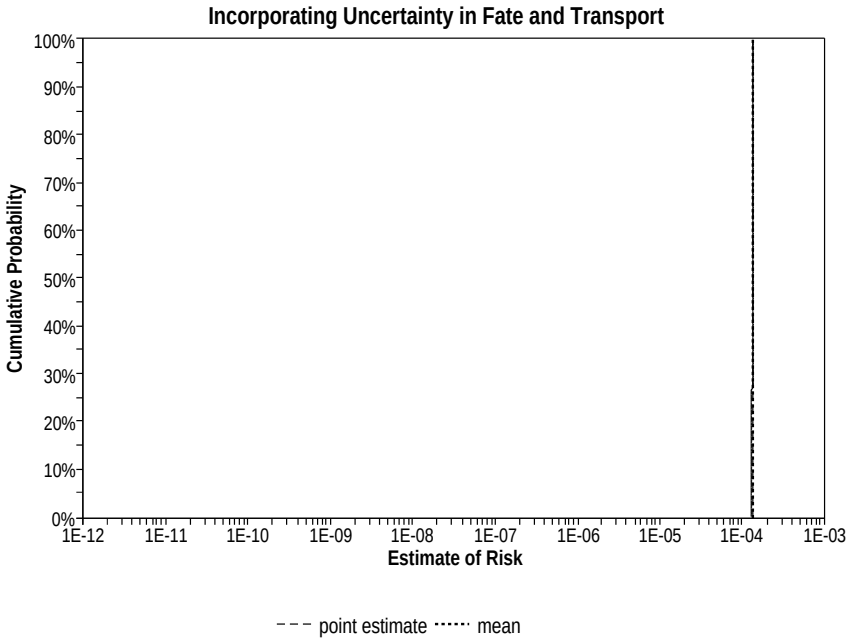

Figure 16 1,1-DCE: distribution of risk: uncertainty in fate, transport, and exposure model parameters only.

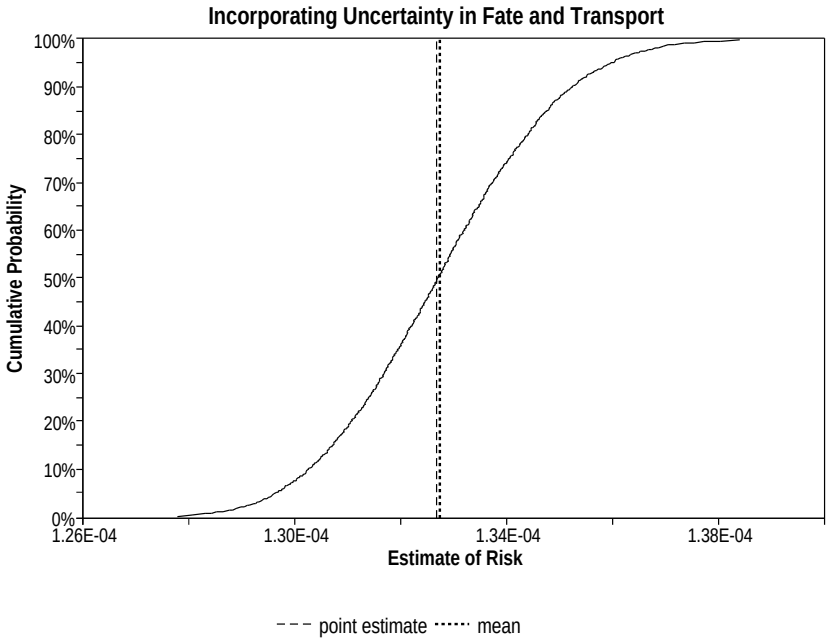

Figure 17 1,1-DCE: distribution of risk: uncertainty in fate, transport, and exposure model parameters only.

8 Probabilistic analysis with a non-standard dose-response model

There have been suggestions that an alternative approach to the evaluation of carcinogens may be to examine only what could happen in the low-dose region, and to drop the assumption that the dose-response curve for carcinogenesis is non-decreasing — so that beneficial effects are possible at low doses. To illustrate what might happen with such an alternative approach, we implemented one example of this idea and applied it to our case study on 1,1-DCE. The only change in the case study was the substitution of the non-standard dose-response models (for carcinogenic potency and unit risk) described in Section 6.3. There are myriad assumptions involved in this simple substitution that we do not list or evaluate (such as whether the interspecies extrapolation, which was empirically evaluated using the standard dose-response model, still applies with this alternative dose-response model).

The resultant distributions for risk are vastly different in some respects from those obtained in Section 7, but very similar in others. They are illustrated in Figure 17 and Figure 18, the latter being a blown-up version of the former, showing the central region more clearly. The difference in shape between Figure 17 and Figure 11 is obvious, Figure 17 showing that with the non-standard dose-response curve there is almost equal probability for carcinogenic and anti-carcinogenic effects, but with high probability the overall effect (positive or negative) is very small. The median of the distribution is very slightly negative, -2×10^{-9}. The mean is also slightly negative, about -9.5×10^{-7}, but the distribution has very long positive and negative tails, so the mean estimate will be of minor interest if there is any asymmetry in the values assigned to positive and negative effects.[10] The 95th percentile is at 1.7×10^{-6}, and the standard point estimate of 1.3×10^{-4} is at the 99.8th percentile.

[10] The long positive and negative tails on the distribution also make the mean extremely difficult to estimate accurately. The standard error in estimating its value is still about $\pm 4 \times 10^{-7}$, or 40%, after 100,000 iterations of the Monte Carlo analysis. Obtaining it to within 10% would take another 1,500,000 iterations!

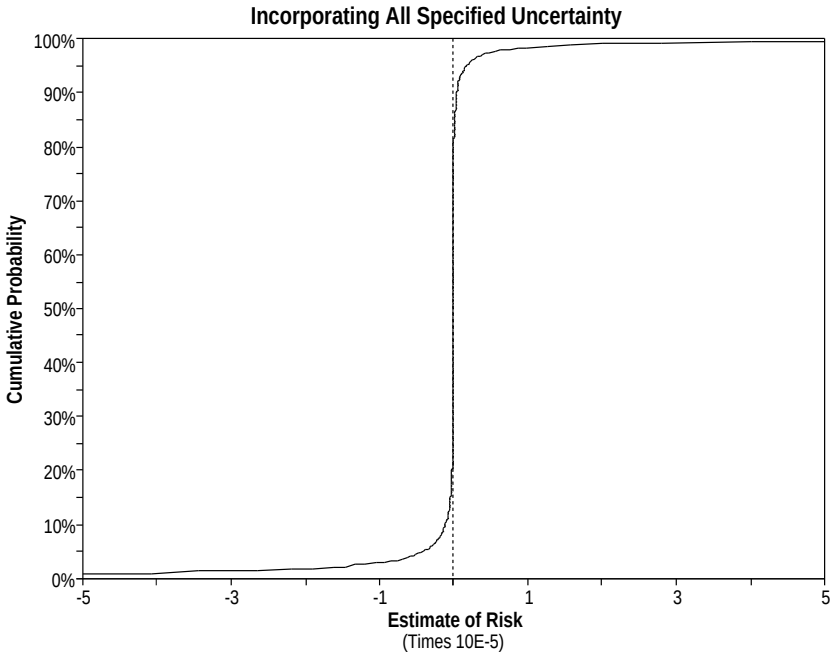

Figure 18 Distribution of risks for 1,1-DCE allowing negative potency values.

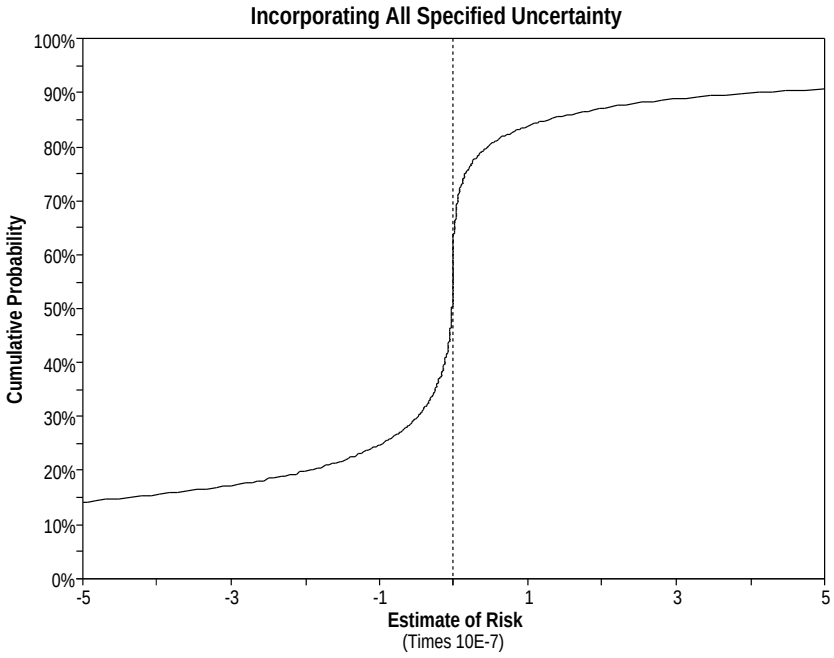

Figure 19 An expanded view of Figure 18.

9 Risk assessment for vinyl chloride

The Monte Carlo analysis of incremental carcinogenic risk associated with 1,1-dichloroethylene was repeated with vinyl chloride replacing 1,1-DCE as the chemical of concern — *i.e.* all the concentration values were assumed to be concentrations of vinyl chloride, not 1,1-DCE. For simplicity, all of the models in the risk assessment were assumed to apply to vinyl chloride exactly as they apply to 1,1-DCE. The only difference between the two risk assessments occurs in the distributions of carcinogenic potency values for the two compounds. Vinyl chloride is a much more potent carcinogen than 1,1-DCE and there is less uncertainty in the results of bioassays on vinyl chloride than there is in bioassays on 1,1-DCE.

Figure 20 through Figure 26 below correspond to the similar figures in Section 7 for 1,1-DCE, showing the effect of the various uncertainties taken singly. A comparison of Figure 22 and Figure 13 reveals the most significant difference between the data for the two compounds. The contribution of uncertainty in the bioassay analysis for vinyl chloride to the uncertainty in the final distribution of risks for a random individual is small relative to other uncertainties such as those in personal characteristics and or in the extrapolation from animals to humans. For 1,1-DCE, uncertainty in the bioassay analysis is a larger contributor to the total uncertainty.

The risk estimates obtained are: median 1.4×10^{-6}, mean 8.8×10^{-5}, 95th percentile 2.0×10^{-4}, and a point estimate 4.1×10^{-4}, which is at the 97.0th percentile of the distribution.

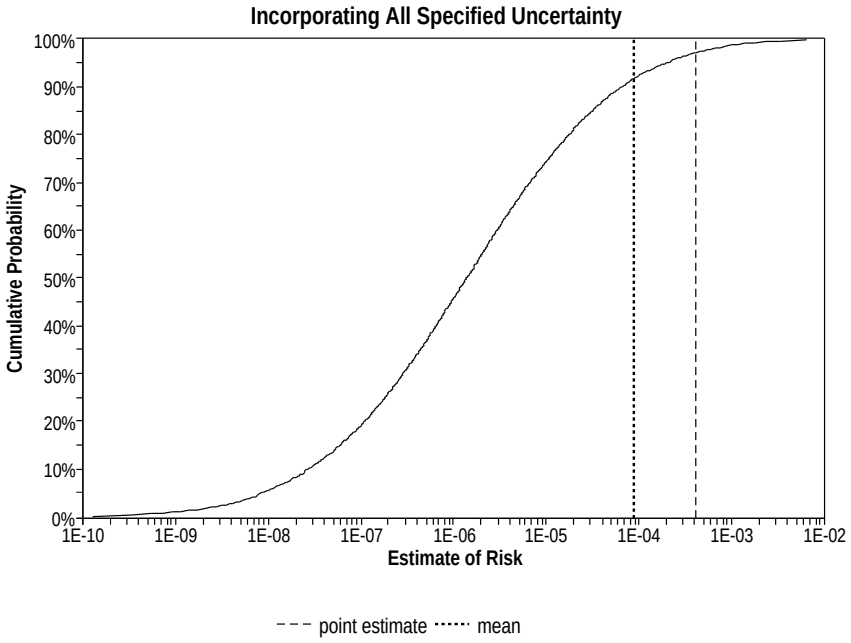

Figure 20 Vinyl chloride: distribution of incremental carcinogenic risk: incorporating all uncertainties.

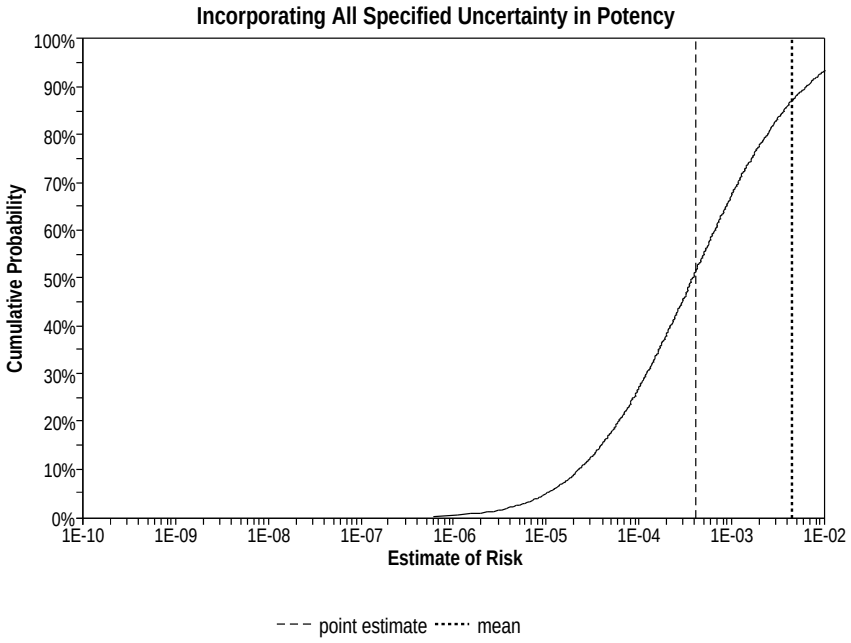

Figure 21 Vinyl chloride: distribution of risk: uncertainty in carcinogenic potency values only.

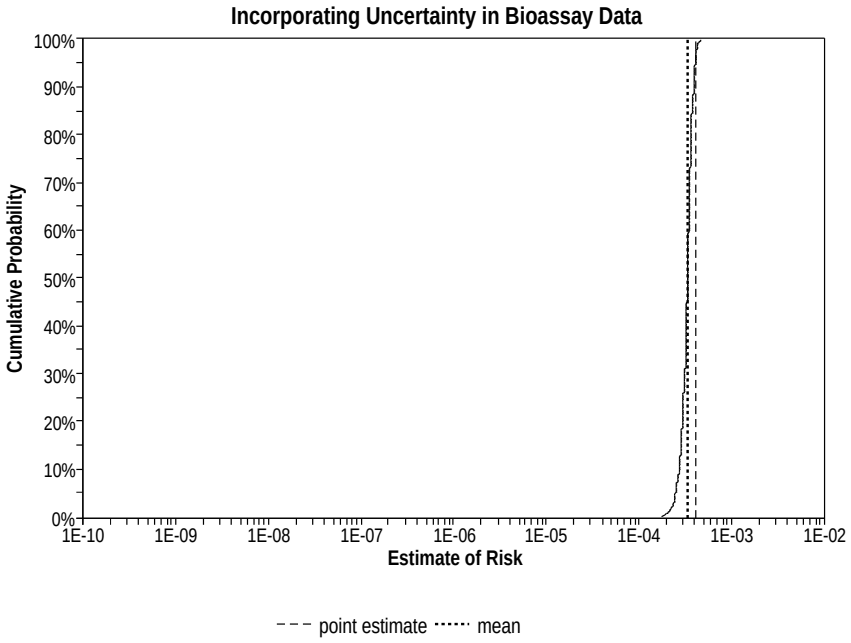

Figure 22 Vinyl chloride: distribution of risk: uncertainty in bioassay analysis only.

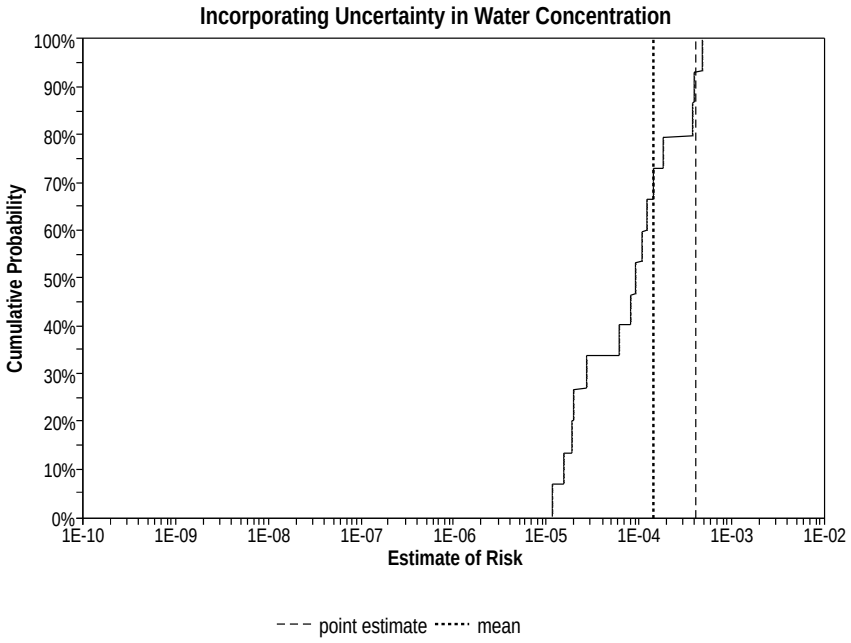

Figure 23 Vinyl chloride: distribution of risk: uncertainty in water concentrations only.

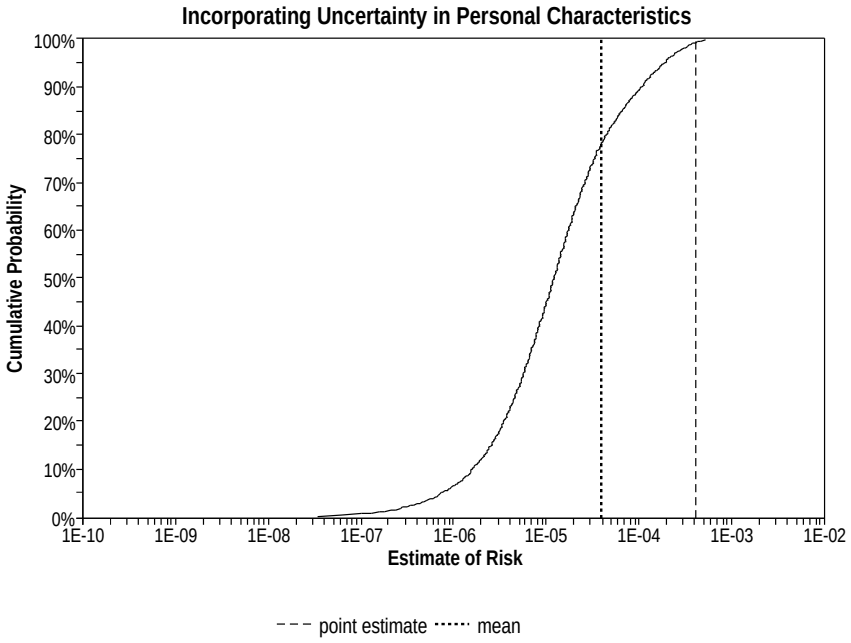

Figure 24 Vinyl chloride: distribution of risk: uncertainty in individual characteristics only.

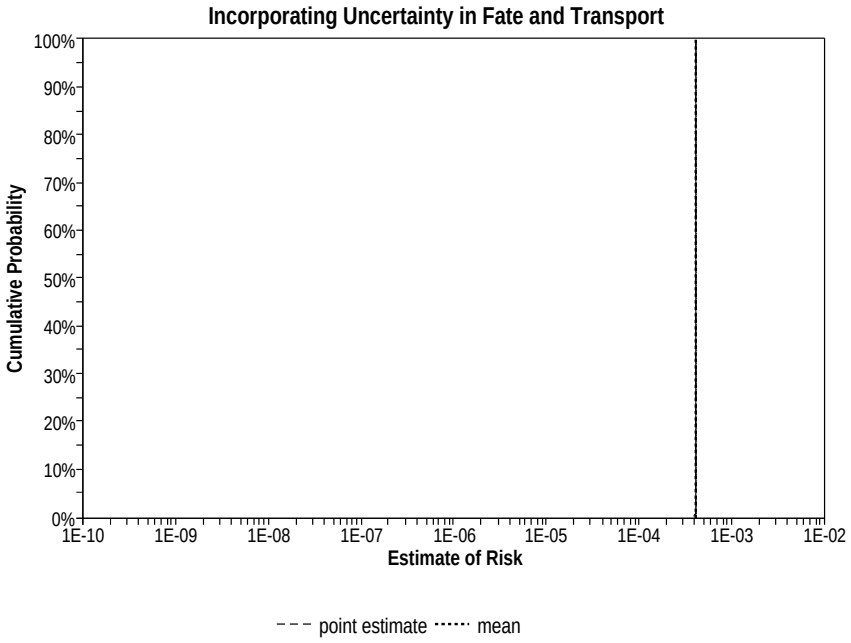

Figure 25 Vinyl chloride: distribution of risk: uncertainty in fate, transport, and exposure model parameters only.

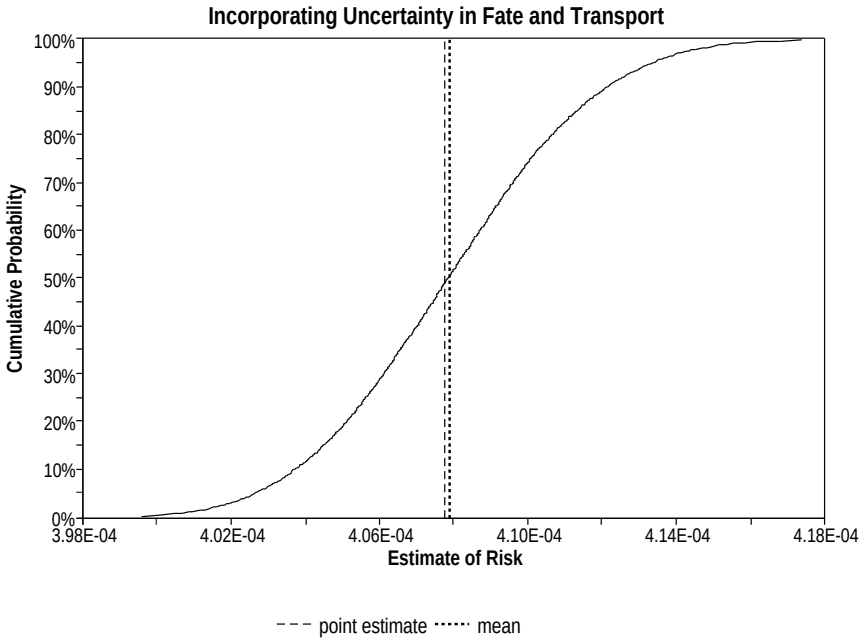

Figure 26 Vinyl chloride: distribution of risk: uncertainty in fate, transport, and exposure model parameters only.

10 Inhalation risks to 1,1-DCE workers

The methodology described in Section 7 was repeated to evaluate a very simple case study of risks to workers exposed to 1,1-DCE through the inhalation pathway, to illustrate what might happen at higher exposures. This case study was designed solely to show the type of distribution obtained at higher exposures, so it has been highly simplified.

Since only the inhalation pathway was evaluated, only three distributions were needed for the Monte Carlo risk assessment — exposure, the inhalation unit risk of 1,1-DCE, and the length of exposure. The inhalation unit risk distribution was obtained previously (Section 6.2.2), and the duration of exposure was fixed at 40 years.

Risks associated with the inhalation of 1,1-DCE by workers were estimated using (the model is the same as that used for the general population):

$$R_{carc} = C_{air} f_d f_w f_l U \qquad (17)$$

where the terms are:

R_{carc}	incremental carcinogenic risk,
C_{air}	average workplace concentration of 1,1-DCE in air (µg/m^3),
f_d	fraction of the day spent in the workplace (1/3),
f_w	fraction of the week spent in the workplace (5/7),
f_l	fraction of lifetime spent working with 1,1-DCE (4/7), and
U	unit risk for 1,1-DCE (m^3/µg).

The fractions that appear in Equation (17) correspond to an 8 hour workday with 5 days/week spent at the workplace (corresponding to approximately 2087 hours per year — vacations have been omitted, but so has overtime). It was assumed that a worker will work with 1,1-DCE for 40 years out of a 70 year lifetime, a conservative estimate of the length of time that someone might work at one location. The average worker would probably spend considerably less than 40 years working in a situation involving exposure to 1,1-DCE.

10.1 Concentrations of 1,1-DCE in workplace air

The distribution of average 1,1-DCE concentrations in workplace air was assumed to be lognormal, with median 0.74 ppm, mean 0.89 ppm (corresponding to a geometric standard deviation of exp(0.6)), so that the majority of workers were assumed to be exposed to concentrations less than OSHA's limit of 1 ppm for an 8 hour time weighted average. For the point estimate an air concentration of 1 ppm is assumed, corresponding to the OSHA limit for workplace concentrations. The distribution of 1,1-DCE concentrations used in the case study is illustrated in Figure 27.

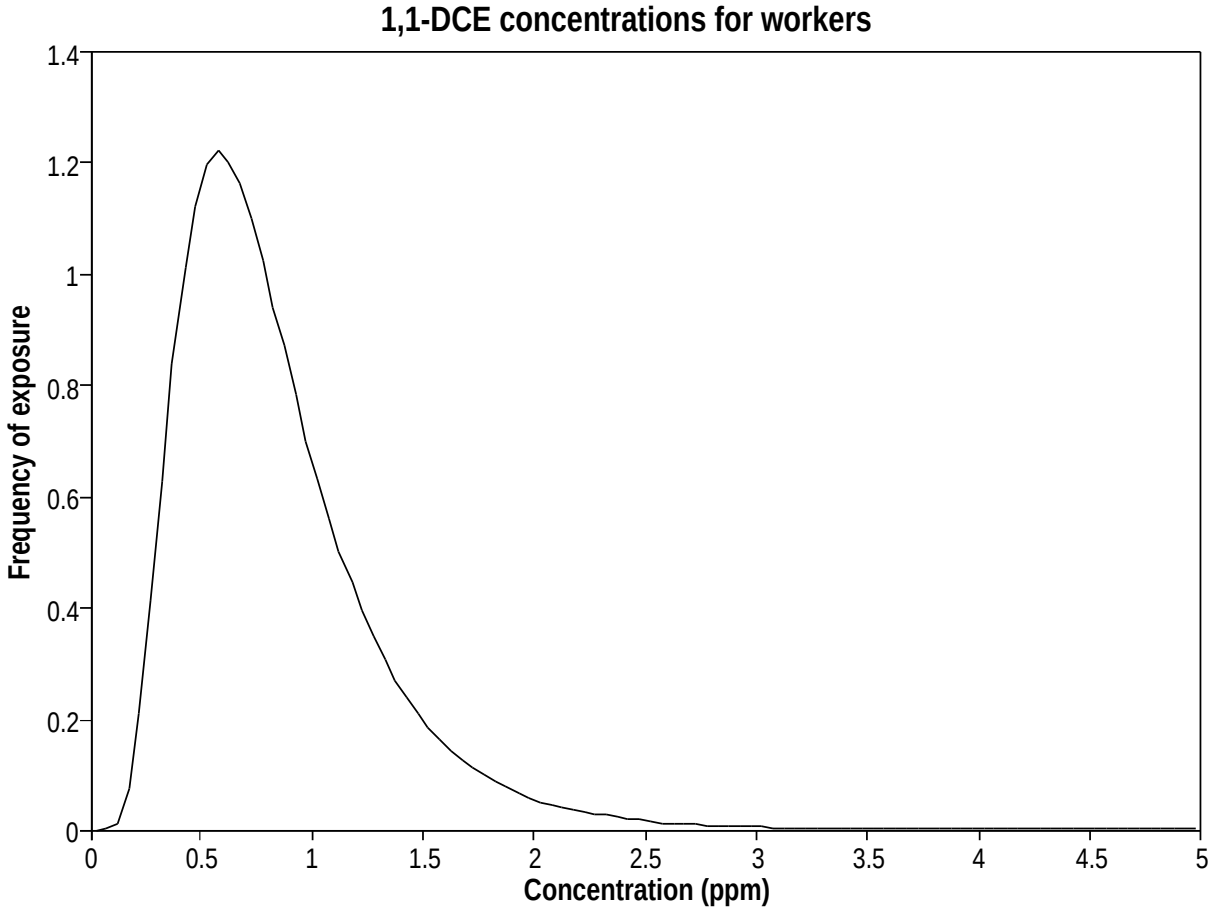

Figure 27 Lognormal distribution of 1,1-DCE concentrations in the workplace

10.2 Incremental carcinogenic risks

Figure 28 is a graph showing the distribution of incremental carcinogenic risks obtained in this case study. The median estimate of risk is 1.4×10^{-8}, the mean is 3.6×10^{-3}, and the 95th percentile of the distribution is 8.4×10^{-3}. Assuming a workplace air concentration of 1 ppm, a duration of exposure of 40 years, and a unit risk corresponding to the EPA value (5.0×10^{-5} m^3/µg), the point estimate of incremental carcinogenic risk for these workers is 2.7×10^{-2}, at the 97.7th percentile.

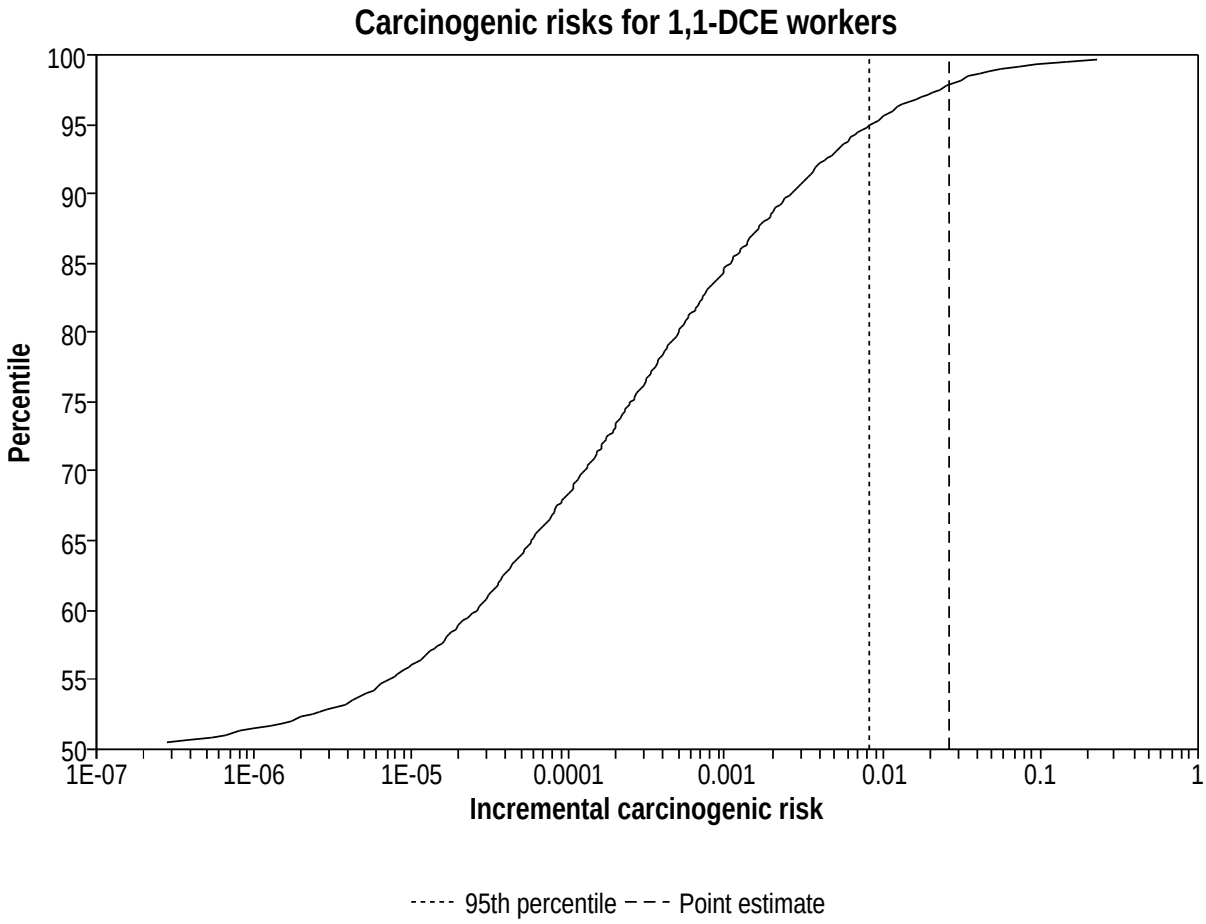

Figure 28 Distribution of incremental carcinogenic risks for workers

11 Comparisons, implications, and comments

11.1 Results and their implications

Some comparisons from the case studies may be illuminating. In Table 11, we summarize the median, mean, 95th percentile, and point estimates of individual carcinogenic risk from 1,1-DCE and vinyl chloride for the four case studies examined.

Table 11 Statistics of the distributions of risk estimates from the case studies

Case	Median (50th percentile)	Mean	95th percentile	Current EPA-style Point-estimate (reasonably maximum exposure)
1,1-DCE, standard	1.2×10^{-9}	1.6×10^{-6}	1.7×10^{-6}	1.3×10^{-4}
1,1-DCE, non-standard	-2.0×10^{-9}	-9.5×10^{-6}	1.7×10^{-6}	—
Vinyl chloride (standard)	1.4×10^{-6}	8.8×10^{-5}	2.0×10^{-4}	4.1×10^{-4}
1,1-DCE workers	1.4×10^{-8}	3.6×10^{-3}	8.4×10^{-3}	2.7×10^{-2}

For the primary case study — that involving a probabilistic assessment of risk from 1,1-DCE in contaminated water at a Superfund site — the difference between the 95th percentile estimate of risk and the point estimate may make all the difference in the world. The point estimate of risk — 1.3×10^{-4} — would be taken (and indeed was taken, in the actual case upon which our study was based) to be a clear, risk-based signal for groundwater remediation; and that remediation might well cost on the order of $10,000,000 – $20,000,000. It would not matter that this point estimate is at the 99.8th percentile of the distribution of risk estimates — indeed, one would not even know its percentile rank using only point-estimate techniques. On the other hand, the 95th percentile estimate of risk — shown here to be only slightly more than one-in-one million — might be taken as risk-based grounds for *not* undertaking remediation; or at least for choosing a "remedy" that would be far less expensive and/or extensive.

This point requires some elaboration. Risk management decisions, as they are called, change with time, across regions, and among decision-makers. At some times and in some regions, individual, upper-bound risk estimates of greater than 10^{-6} suffice for demanding remediation at Superfund sites; but for other situations, remediation is only required if risk estimates exceed 10^{-5}. Several states now use the 10^{-5} risk level as their "bright line" for decision-making; and some of those had preferred the 10^{-6} benchmark until it became clear how many sites/situations violated such a guideline, at which point their expectations relaxed. The 10^{-6} benchmark has not been shunned entirely, though; so that even doing risk assessment "right" — if that is what the sort of Monte Carlo approach we have taken here is

taken to be — those responsible may still have to spend the same sort of money on remediation and so on as at present.

The "standard" and "non-standard" analyses of 1,1-DCE (which differ only in the model used to interpolate the dose-response data from the rodent bioassay) show substantial — indeed, diametric — differences in the median and mean estimates. These two "best estimates" of risk in the "non-standard" analysis have a negative sign, indicating that exposure to 1,1-DCE at the low levels postulated might confer a net benefit! Clearly the definition of what statistic is the "best" estimator has to be carefully made to match the desired end point, since different choices of statistic may lead to diametrically opposed choices for policy.

In either the "standard" or "non-standard" analyses, the median estimates of risk (for carcinogenesis in the standard case, and for anti-carcinogenesis in the non-standard case) are extremely small — that is, quite close to zero. Moreover, the *upper end* of the uncertainty distribution is not substantially changed by the change in analysis method. The 95^{th} percentile of the distribution of risk estimates for both the standard and the non-standard approaches are identical for all practical purposes.

Such identical results for the high-end estimate might seem peculiar. After all, the standard approach assumes that all doses of 1,1-DCE may be carcinogenic, even though there is no evidence of carcinogenicity between the zero-dose and low-dose groups. In contrast, the non-standard approach makes no such assumption, and indeed allows the dose-response line in this region — the negative slope of which indicates anti-carcinogenesis — to be as negative as the underlying uncertainty suggests. These different approaches do indeed give vastly different sorts of risk estimate distributions at the low and mid-ranges (compare Figure 11 with Figure 17); but if one wishes to be 95% certain that a random individual's risk is a certain value or lower, then that value turns out to be the same for both approaches.

This point is lost on many critics of EPA-style risk assessment. Many such critics, believing that standard, regulatory risk assessments are heavily biased toward over-estimating the "real" risk, suppose that a full and honest incorporation of uncertainty and variability will vastly lower estimates of risk, even for those at the high end. Forgotten in this notion is that uncertainties and variabilities run in both directions. A compound may indeed be much less carcinogenic than the EPA-style point-estimate of potency suggests, for example; but it may also be more carcinogenic than EPA's point estimate allows. We return to this theme in the subsection following.

Comparison of the standard analyses of 1,1-DCE and vinyl chloride is also illuminating. The only differences between these two case studies are the carcinogenic potency estimates for the two substances — all other details are assumed to be identical. The standard EPA point estimate of carcinogenic potency for vinyl chloride is approximately a factor of three higher than that for 1,1-DCE, and the unit risk estimate is only 60% higher. These differences account for the relatively small factor of three difference in the point estimates of risk.

The proximity of the point estimates for risks from equal exposures to 1,1-DCE and vinyl chloride signals the erroneous nature of the point estimate approach. Vinyl chloride is a clear and strong carcinogen in both rodents and people; 1,1-DCE is a marginal rodent carcinogen at best. Surely the risks from the two compounds should differ by more than a factor of three. The distributions provided by the Monte Carlo approach do indeed display a much greater disparity between the standard 1,1-DCE and vinyl chloride cases.[11]

11.2 Implications of, and general comments on, the proposed risk assessment reforms

Working through these case studies, we have developed certain opinions about the proposed risk assessment reforms. As we have made clear, we wonder about the new emphasis on best estimate(s) of risk, as opposed to the old (and current) emphasis on an upper-bound estimate of risk. Best estimate, *per* the bills, means a "central estimate of risk," arrived at using the "most plausible and realistic assumptions . . . given current scientific information." We suppose that this emphasis, and the specific language, is an expression of the belief that current risk assessments inappropriately concatenate many upper-bound factors — so that the product is a highly implausible estimate of risk that applies to few or no real people exposed in the real world. We know of at least three groups who often share this view.

---

[11] Even the Monte Carlo distributions given here do not capture all of the differences between the carcinogenic potencies of vinyl chloride and 1,1-DCE. Vinyl chloride is classified as a known human carcinogen, whereas 1,1-DCE is classified as a possible human carcinogen. The qualitative distinction, while important, is not reflected in the width of the distribution of risk estimates because both potency estimates derive from rodent bioassay data so have the same uncertainty in interspecies extrapolation. This uncertainty is the same *despite* the fact that most analysts feel far more confident that vinyl chloride is a human carcinogen than that 1,1-DCE is a human carcinogen. One could remove the uncertainty in interspecies extrapolation for vinyl chloride by basing the potency estimate on epidemiologic data. One would then have to introduce, however, the uncertainty in the estimates of human exposure in the epidemiologic studies, which may be as large or larger than the uncertainty in interspecies extrapolation.

First are the regulated communities — whether private companies, municipalities, or others — who bear many of the costs of complying with certain regulations that may do little to reduce actual risk to the public health. The second is the research toxicology community, at least some members of which wonder why information they generate regarding mechanisms of carcinogenesis, for example, is largely ignored in favor of default assumptions such as that essentially all chemical carcinogens are nearly equally potent at all non-zero levels of exposure.

The third group, perhaps less well known, are the local regulatory authorities. Some of those responsible, for example, for writing or re-writing permits for facilities express frustration at U.S. EPA risk estimates for regulated chemicals. The Agency explains that the risk estimates are upper-bound values; and that the true value is likely to be lower, and may even be zero. But the local permit writer may not wish to know what the risk "might be," or "might be less than;" he or she wants to know what the risk *is*, at least given current scientific evidence and expert judgement.

Some people therefore believe that a focus on a best estimate of risk would help. Perhaps those paying for compliance with regulations would feel (if not be) less burdened; perhaps results from the laboratory would be translated into more chemical-specific inputs for risk assessment; and perhaps permit writers would be better able to match permit conditions with actual, or at least expected, risk reduction.

But there are two distinct problems that are not acknowledged by such people or in the proposed legislation. First, there is no single estimator of risk that should be used in all situations, and the definition of the estimator is what is most important. Second, no matter what estimator of risk might be chosen, the estimate has to be compared with some standard or another, and the choice of standard is also important.

It is plainly inadequate to perform a risk assessment such as the one we have performed (for the risk to a randomly chosen individual), and then use a central tendency (*e.g.* the median estimate) estimator, if the legislative requirement is, for example, to set a standard that protects everybody with an adequate margin of safety. What is required in such situation is to analyze the legislative requirement (protection of everybody with a margin of safety), and then to choose a suitable measure of risk — for example, an estimate of the risk to the most exposed member of the population might be a suitable choice. Given such a choice, one could demand that the estimator for this most exposed member of the population be the median for the uncertainty distribution of risk to that person (and then apply a safety factor of some sort — to afford the "adequate margin"). However, many of the parameters going into the evaluation of risk would certainly not be central tendency values, since the most exposed person will be extreme in some ways. Similar considerations apply in every other situation.

The second problem is the selection of the standard against which to compare. Why perform a risk assessment if not to compare the result to some standard? The current standard is an upper bound "risk estimate" of 10^{-4} to 10^{-7}, depending on the situation. If one simply demands that the current upper bound estimate be replaced with a central tendency value, what is to prevent the standard of comparison from changing to 10^{-8} or 10^{-11} or whatever? If

one is going to change the rules of evaluation of risk estimates, one also has to take care to accurately define the standards against which those estimates will be compared.

More useful might be new rules of thumb based on both central-tendency and 95th percentile estimates of risk, or on other properties of the distribution of risk. Imagine two scenarios; air pollutant α poses a central-tendency risk estimate of 1 in 100,000,000 and a 95th percentile risk estimate of 1 in 1,000; while air pollutant β poses a central-tendency risk of 1 in 10,000 and an upper-bound risk that is the same as for air pollutant α, 1 in 1,000. The current practice in risk management would be to regard both air pollutants as equally risky — equally in need of reduction *via* regulation. Perhaps the new approach would make air pollutant β a much stronger candidate for regulation — because the "real" risk (*per* risk assessment focussed on "best estimates") is 10,000 times greater for β than for α, even though, at the upper-bound, each could be equally risky.

More fundamentally, credible efforts at risk assessment reform should recognize that the notion of an entirely scientifically objective risk assessment is a mirage. To the extent that the authors of the reform legislation believe otherwise, we believe they have been misled.

Science is a method of observing the natural world and conducting tests of various hypotheses about how the world operates. Good science elicits a high degree of consensus among those who analyze it exactly because of its objectivity, reproducibility, and explanatory power.

Of course, science contributes heavily to risk assessment: it does so in such areas as identifying physical relationships that govern the transport of underground pollutants to the air, testing the ability of a chemical to cause cancer in laboratory animals, and measuring the volume of air inhaled in a day by persons of different ages. In each case, a scientific goal is clearly specified and can be reached by measurement of the quantities or processes of interest.

Unfortunately for those desiring certainty and consensus, risk assessment problems are frequently unresolvable by purely scientific methods. Only scientific questions can be answered by scientific methods, and risk assessment problems are often largely or entirely non-scientific. Thus, the solutions (or resolutions) to risk assessment questions vary greatly in the degrees to which they are rooted in scientific observations and deduction. The critical factors determining whether a risk assessment question can be satisfactorily answered using only scientific tools include: has careful science been done on this or similar problems in the past? If so, have scientists been able to observe this problem on a scale that is directly useful for public policy? Usually, the answers are "no."

A typical risk assessment problem is one like that examined in the case studies presented here: what are the chances that exposure to a certain amount of a chemical will cause cancer in people? Were a scientist assigned this problem, given no additional information about the particular concerns of the questioner, and forced to answer it as "scientifically" as possible, he would be limited in his work to one area of study: epidemiology. He would find people exposed to the chemical at the proper level, and comparable people with no exposure, monitor their health, and use statistics to evaluate the significance of any

differences in cancer rates between the two groups. In other words, he would look to actual instances of the problem of interest. Few inferences would be needed in applying the results of exposure to one group of people to another.

Alas, epidemiology is like a low-powered telescope — it does not allow one to see very far. The scientist's unarguable, objective conclusion might be that the exposure of interest may, at most, double one's rate of cancer, or may have no effect. This sound scientific answer[12] to the question posed may be useless for public policy, which might wish to discriminate very small increases in cancer risk.

In order to produce results that are useful for policy, scientists must turn away from the most relevant and objective information — that gained from deliberate or accidental human exposure — towards other branches of science such as rodent toxicology. The study of carcinogenesis in laboratory animals is a science in its own right — but the demand that it be applied to the human situation *ipso facto* introduces unscientific elements into the analysis. The less like the situation of interest are the phenomena that science has been able to observe and test, the more inferences and assumptions must be made. Many inferences are chains of reasoning that have not been, or cannot be, adequately tested by scientific method.

Risk assessment problems of concern to legislators and regulators are rife with such questions that science cannot answer. Whether a novel chemical that causes cancer in one strain of rat, but not in others or in mice, is able to cause cancer in humans is an issue science can not answer at the present. It is conceivable, however, that this question could be answered in the future, either because accidental human exposure supplies the response, or because the biochemistry and molecular biology of rodents and humans is learned in such detail that the inferential gaps narrow to insignificance.

Importantly, other elements of risk assessment problems will *never* be answerable by scientific methods. Such elements are future behaviors by persons, institutions, molecules, the weather, and cannot be predicted in detail. Past experience guides scientists in making long-term predictions about average or typical behaviors, but is of less use in predicting outcomes for particular individuals at particular points in time; indeed, there is no guarantee that even past trends will continue.

Limits to the knowable are inescapable. As Ralph Gomory (1995) has written:

> We are all taught what is known, but we rarely learn about what is not known, and we almost never learn about the unknowable. That bias can lead to misconceptions about the world around us. . . .

---

[12] Of course, we are simplifying even here. The soundness of the epidemiologic study itself is subject to many things: how well the exposures can or have been quantified; how well exposures to other potentially harmful or helpful substances have been accounted for; and so on.

> [People do not] try to predict the automobile accidents they will be involved in. To know that we will be struck by a car next year, we would have to know, with impossible accuracy, the particulars of the life of the driver, his habits, his timing, his way of pressing the accelerator and so forth. . . . It is clear that all these details are unknown, and we do not try very hard to learn about them because we instinctively realize they are also unknowable.
>
> In distinguishing the known or the unknown from the unknowable, the level of detail can be decisive. The level of detail is what separates the delusion of the gambler from the wealth of the casino owner. The gambler attempts to predict the individual and unpredictable spins of the roulette wheel; the owner concerns himself with the quite predictable average outcome.

Risk assessment almost always depends upon details that are inherently unknowable. Accordingly, how could anyone suppose there is one way to do it right?

11.3 Implementation issues

Another issue of interest to us is the substantially increased complexity of risk assessment performed *per* the proposed reforms. First, it is a fairly complex undertaking to derive the many distributions about model parameters that serve as inputs to these risk assessments. The task requires a great deal of research and some statistical expertise; although once defined and derived, these distribution parameters could be made available to the risk assessment community in ways similar to the publication of toxicity values in EPA's Integrated Risk Information System (IRIS) and in its various guidance documents.

Second, a more-than-typical breadth and depth of expertise is needed to perform probabilistic risk assessments. The risk assessor must genuinely understand (as opposed to merely use) the models (and perhaps be able to create some anew); and he or she must combine distributions in valid manners. In our experience, there are very few risk analysts who can do this well, whether inside of EPA or elsewhere.

Third, even simple probabilistic risk assessments require considerable computational complexity. Although software exists to perform probabilistic assessments as add-ons to spreadsheet programs that run on personal computers, our experience has been that these programs become prohibitively slow when applied to assessments of any appreciable size. Our main case study, for example, involves only one chemical of concern, only one contaminated medium, three exposure routes, and only one end-point — carcinogenic response. Existing software would require approximately one-half hour to calculate the 10,000 - 20,000 iterations necessary to adequately characterize the distribution of risks. Risk assessments of the usual size may require analysis of ten to fifty chemicals of concern, three to five contaminated media, ten to twenty exposure routes, and at least two dose-response distributions (carcinogenic and non-carcinogenic responses). These combine multiplicatively, so that commercially available software soon becomes prohibitively slow.

The implementation of some of the more complex distributions discussed in Appendix A is also not straightforward in commercially available software. Such distributions can be approximated by tabular data, but we worry that certain important features may be lost. As an example, EPA published (EPA, 1989b) a table corresponding to the shower data shown in Appendix A.3, but omitted the top 2% of households with average shower lengths of more than 20 minutes. If extreme values are of interest, then such extreme values must obviously be incorporated into the input data.

To prepare such assessments, then, one must write one's own software. We have done so for this case; and in this case, our software conducts the 20,000 iterations in three minutes or so (on a 486/66 PC compatible machine with all language debugging checks turned on).

12 References

Andelman, J.B. (1985). Inhalation exposure in the home to volatile organic contaminants of drinking water. *The Science of the Total Environment 47*:443-460.

Anderson, E., Browne, N., Duletsky, S., Ramig, J., and Warn, T. (1985). *Development of Statistical Distributions or Ranges of Standard Factors Used in Exposure Assessments*. Washington, DC: U.S. Environmental Protection Agency. EPA/600/8-85/010.

Crouch, E.A.C., R. Wilson, and L. Zeise, 1983. The risks of drinking water. Water Resources Research 19:1359–1375.

Crouch, E.A.C., 1992. MSTAGE — a program to fit end-of-life carcinogenesis bioassay data to the multistage formula, Version 2.01. Cambridge Environmental Inc. (Available from the author).

EPA, 1985a. Health Assessment Document for Vinylidene Chloride. Final report. OHEA, EPA, Washington DC 20460. August 1985. EPA/600/8083/031F.

EPA, 1985b. Health and environmental effects profile for chloroethene. ECAO, ORD, US EPA, Cincinnati, OH 45268. Sept. 1985. EPA/600/22.

EPA, 1988. Superfund Exposure Assessment Manual. Office of Remedial Response, Washington, DC 20460. EPA/560/1-88/001. OSWER Directive 9285.5-1.

EPA, 1989a. Risk Assessment Guidance for Superfund, Volume I, Human Health Evaluation Manual (Part A) EPA/540/1-89/002.

EPA, 1989b. Exposure Factors Handbook. Office of Health and Environmental Assessment, Washington, DC 20460. EPA 600-8-089-043.

Feron, V.J., C.F.M. Hendriksen, A.J. Speek, H.P. Til, and B.J. Sp't, 1981. Lifespan oral toxicity study of vinyl chloride in rats. Food Cosmet. Toxicol. 19(3):317–333.

Gomory, R.E. (1995) The Known, the Unknown and the Unknowable. *Scientific American 272*:120.

Jo, W.K., Weisel, C.P., and Lioy, P.J. (1990). Routes of chloroform exposure and body burden from showering with chlorinated tap water. *Risk Analysis 10*:575-580.

Maltoni, C., G. Lefemine, A. Ciliberti, G. Cotti, and D. Carretti, 1980. Vinyl chloride carcinogenicity bioassays (BT project) as an experimental model for risk identification and assessment in environmental and occupational carcinogenesis. Epidemiol. Anim. Epidemiol. Hum.: Cas Chlorure Vinyle Monomere, (Reun. Club Cancerog. Chim), 20th Meeting Data 1979, 11-112. Publ. Essent., Paris, France.

Maltoni, C., G. Lefemine, P. Chieco, G. Cotti, and V. Patella, 1985. Experimental research on vinylidene chloride carcinogenesis. In: C. Maltoni and M.A. Mehlman, eds. Archives of Research on Industrial Carcinogenesis, Vol. 3. Princeton Scientific Publishers.

NCI/NTP (National Cancer Institute/National Toxicology Program), 1982. NTP technical report on the carcinogenesis bioassay of vinylidene chloride in F344/N rats and B6C3F1 mice (Gavage study). Technical Report No. 228. NTP No. 80-82. NIH Pub. No. 82-1784. NTP, Research Triangle Park, NC. and Bethesda MD. U.S. Department of Health and Human Services, Public Health Service, National Institute of Health.

U.S. Bureau of the Census (1993). *Statistical Abstract of the United States 1993*. (113th edition). Washington, DC: U.S. Department of Commerce.

Yang, Y. and Nelson, C.B. (1985). An estimation of the daily average food intake by age and sex for use in assessing the radionuclide intake of individuals in the general population. EPA 520/1-84-021.

Appendix A: Random numbers and distributions

This appendix describes many of the distributions used in the case studies. Section A.1 gives details of some of the standard distributions, while subsequent sections describe specific distributions that build on the material of Section A.1.

A.1 Random number generation

The Monte Carlo technique strictly requires random numbers for its implementation. For our implementation, we used the standard computer technique of approximating this ideal with pseudo-random numbers. The basic generator is for the uniform distribution. All the other distributions are generated from the uniform distribution, so that given a uniform random number generator, the other distributions are generated exactly (Devroye, 1986).

A.1.1 Standard uniform pseudo-random variate generation

The generator used is a linear congruential generator of the form:

$$x_{\varphi+\Theta} = Ax_{\varphi} + 1 \mod B$$

where:

x_n is a sequence of 8 byte (64 bit) integers,
A = 6,364,136,223,846,793,005 (hexadecimal 5851F42D4C957F2D }, and
B = 2^{64}

and all integer arithmetic is performed exactly. The multiplier A is an "excellent" one that passes the spectral test (Knuth, 1981).

A sequence of real values in the range [0,1) is obtained by multiplying the 8 byte (64 bit) integer x_n by 2^{-64}, and these real values are returned with full 64 bit precision by using the 80 bit extended precision real format of the INTEL® iAPX® coprocessors.

The seeds used for each run may be chosen arbitrarily by hand or approximately randomly from the system clock. Four word (2 byte) values are obtained, and adjoined to form the 8-byte seed.

During all calculations, intermediate real values are held with 64 bit precision, again using the 80 bit extended precision real data format. Results are stored with the 52 bit precision of the IEEE standard double precision real number, all conversions from higher precision being performed by rounding.

A.1.2 Arbitrary uniform random variates

Given a standard uniform random variate *U*, a random variate uniform on the range [*a*,*b*) is obtained as:

$$X = a + (b - a)\,U$$

A.1.3 Triangular distribution random variates

The triangular distribution is defined by a range (*a*,*b*) and a point *p* where the distribution has its mode. Given a standard uniform random variate *U*, a random variate *T* from such a triangular distribution is obtained as:

$$\text{if } U < (p-a)/(b-a) \text{ then } T = a + \sqrt{(b-a)(p-a)\,U}$$
$$\text{else } T = b - \sqrt{(b-a)(b-p)(1-U)}$$

A.1.4 Exponential random variates

Given a standard uniform random variate *U*, an exponential random variate *E* with density function $e^{-\lambda x}/\lambda$ is obtained as:

$$E = -\frac{1}{\lambda}\ln U$$

A.1.5 Normal random variates

Given standard uniform variates U_1, U_2, two independent standard (mean zero, standard deviation 1) random normal variates are generated as:

$$N_\Theta = \sqrt{-2\ln U_\Theta}\sin 2\pi U_\theta$$
$$N_\theta = \sqrt{-2\ln U_\Theta}\cos 2\pi U_\theta$$

Given a standard normal random variate *N*, a variate *n* from a normal distribution with mean μ and standard deviation σ is generated as:

$$n = \mu + N\sigma$$

A.1.6 Lognormal random variates

Given a standard normal random variate *N*, a lognormal variate with median exp(μ) and geometric standard deviation σ is generated as:

$$L = e^{\mu + \phi\sigma}$$

A.1.7 Gamma random variates

A Γ(*a*,*b*) random variate has density function:

$$\frac{b^{-\gamma} x^{\gamma-\Theta} e^{-\eta\nu}}{\Gamma(a)}$$

In an obvious pseudo-code (where `random` returns a standard uniform variate), the algorithm for a gamma random variate is (Devroye, 1986):

```
c:=a-ln(4); λ:=Sqrt(2*a-1); d:=a+λ;
if (a<1) then begin  { Johnk's generator }
      repeat
            y:=random^(1/a); z:=random^(1/(1-a));
      until (y+z) ≤ 1;
      return -ln(random)*y/(y+z);
end else begin   { Cheng's rejection algorithm GB }
      repeat
            u:=random; v:=random;
            y:=ln(v/(1-v))/λ; x:=a*exp(y); z:=u*Sqr(v); r:=c+d*y-x;
      until r ≥ ln(z);
      return b*x;
end
```

A.1.8 References for Section A.1

Knuth, D.E., 1981. The art of computer programming, Vol. 2: Seminumerical Algorithms. Addison-Wesley Publishing Co., Reading, MA.

Devroye, L., 1986. Non-uniform random variate generation. Springer-Verlag, NY.

A.2 Expected time of residence in U.S. housing

A.2.1 Analysis (Israeli and Nelson 1992)

Israeli and Nelson (1992) obtained data from the American Housing Surveys of 1985 and 1987 (U.S. Dept. of Commerce, 1985,1987) on the numbers of people who remained in their houses after various periods (Table A.1 and Table A.2).

Let $S(t)$ be the fraction of households found in a survey to have moved into their current residences at least t years before the survey. Assume that S is ergodic, with $S(0) = 1$, and $S(\infty) = 0$, and let $P = 1 - S$ be the probability for a household to have lived in its current residence for less than t years. Then the density function $p(t)$ is given by:

$$p(t) = \frac{dP}{dt} = -\frac{dS}{dt} \tag{A.9}$$

and the average time T_{CR} for a household being in a current residence is:

$$T_{P\psi} = \int_{\eta}^{\infty} tp(t)\,dt = \int_{\eta}^{\infty} S(t)\,dt \tag{A.10}$$

provided $tS(t) \to 0$ as $t \to \infty$.

Now let $M(t)$ be the probability that the total residence time for a household is less than t years, and $m(t)$ its corresponding density function. Then the ergodic hypothesis implies that:

$$R(t) = 1 - M(t) = \frac{p(t)}{p(0)}$$

so that:

$$m(t) = \frac{dM}{dt} = -\frac{1}{p(0)}\frac{dp}{dt} \tag{A.12}$$

and the expected total residence time T is given by:

$$T = \int_{\eta}^{\infty} tm(t)\,dt = \frac{1}{p(0)}\int_{\eta}^{\infty} p(t)\,dt \tag{A.13}$$

provided $tR(t) \to 0$ as $t \to \infty$.

A.2.2 Numerical fitting to data

Israeli and Nelson (1992) used the data of Table A.1 and Table A.2 to fit the functional form:

$$S(t) = \exp(-a_0 b_0 (1 - e^{-b_1}) - a_1 t - a_2 b_2 (e^{-b_3} - 1)) \qquad \text{(A.14)}$$

with a_i, b_i non-negative. This functional form was chosen by examining the empirical values S_e for $S(t)$ obtained from the data, and noting that they resembled survival functions in mortality analysis. This functional form was fitted by minimization of $(S - S_e)^2/S$. No reason was given for choosing $1/S$ as the weight function — it was simply baldly stated that this was the weight function to be used. In addition, when certain of the parameters would have been selected to be zero by the minimization, they were constrained at positive value.

There are three problems with Israeli and Nelson's analysis. First, the functional form (A.14) is not very convenient for integration or generation of random variables. Second, the fitting procedure arbitrarily chose $1/S(t)$ as the weight function. Third, Israeli and Nelson apparently did not choose very stringent termination criteria for their optimization — the values they give are distinctly different from the true optima. There are very high correlations between the values of the parameters when fitting data with this parameterization, so that very different values of the parameters give almost identical results.

Choosing different weight functions for the optimization (*e.g.* $1/S^2$, or minimizing $(\ln(S/S_e))^2$)gives substantially different values for many of the parameters, so different in fact that the estimated mean time of residence T is changed more than the estimated standard deviation of this value. Thus the standard deviations given by Israeli and Nelson must be treated with the utmost caution.

There does not seem to be any theoretical reason for choosing any particular weight function. The data could be considered a multinomial distribution, but the error terms turn out to be much larger than would be expected. For this analysis, a somewhat different functional form was chosen for S:

$$S(t) = p_1 e^{-b_1 t} + p_2 e^{-b_2 t} + 2 p_3 (1 - \Phi(ct)) \quad \text{with} \quad p_1 + p_2 + p_3 = 1 \qquad \text{(A.15)}$$

where Φ is the cumulative standard normal function. The minimization was of the sum of squares of the logarithm of the ratio S/S_e, in order to minimize the relative deviations of S. The sum of squared relative errors obtained using this procedure are of (on average) slightly smaller than those obtained using Israeli and Nelson's function (A.14) — it varies from household type to household type. The functional form (A.15) has the advantage that it, and the required derivatives of it, are easy to generate using standard random number generators.

Graphs of $S(t)$ and relative residuals $(S - S_e)/S_e$ similar to those given by Israeli and Nelson (1992) are given in Figure A.1 and Figure A.2 respectively. The relative residuals are smaller on average than those obtained by Israeli and Nelson, principally because their total

sum of squares has been minimized for each fit (Israeli and Nelson's fitting procedure did not do this).

With the functional form of equation (A.15), the fitted parameters for *S(t)* are given in Table A.3. Differentiating, and using equation (A.11), leads to:

$$R(t) = 1 - M(t) = q_1 e^{-b_1 t} + q_2 e^{-b_2 t} + q_3 e^{-c^2 t^2/2} \text{ with } q_1 + q_2 + q_3 = 1 \qquad \text{(A.16)}$$

where obviously the q_1, q_2, and q_3 may be expressed in terms of p_1, p_2, p_3, b_1, b_2, and c. Table A.4 lists these new parameter combinations. Evidently, *M* is easy to generate, since, as indicated by equation (A.16), it is basically a linear combination of three exponentials.

Table A.1 Number of households (in thousands) in 1985 in residence for given periods

Years at current residence	0-1	1-6	6-11	11-16	16-26	26-36	36-46	46+
Allhses	13762	28539	15729	9221	10800	6257	2563	1554
renters	10181	14407	3960	1645	1292	463	184	147
owners	3581	14132	11769	7576	9508	5794	2379	1407
farms	82	370	364	245	267	224	169	133
urban	11213	21797	11339	6472	8155	4649	1691	915
rural	2549	6743	4390	2749	2646	1608	872	639
ne-rgn	1993	5679	3373	2058	2608	1734	779	506
mw-rgn	3196	6986	4017	2283	2859	1654	684	464
s-rgn	5216	9717	5147	3206	3568	1957	793	460
w-rgn	3357	6157	3193	1674	1766	912	307	125

Table A.2 Number of households (in thousands) in 1987 in residence for given periods

Years at current residence	0-3	3-8	8-13	13-18	18-28	28-38	38-48	48-0
Allhses	33920	17436	12613	7907	9711	5720	2292	1289
renters	20748	6462	2674	1212	997	369	162	101
owners	13172	10973	9939	6695	8714	5352	2130	1188
farms	254	263	273	230	245	228	132	115
urban	26925	12750	8920	5446	7223	4208	1511	763
rural	6995	4686	3693	2461	2488	1512	781	525
ne-rgn	5715	3795	2677	1766	2369	1522	679	429
mw-rgn	7717	4300	3226	1975	2510	1562	628	349
s-rgn	12216	5849	4191	2795	3307	1791	721	401
w-rgn	8273	3492	2519	1371	1525	845	263	110

Table A.3 Parameters values for *S(t)*

Parameter	p_1	p_2	p_3	b_1	b_2	c
Allhses	0.17644	0.37879	0.44477	0.867172	0.152536	0.044961
renters	0.721652	0.278348	0	0.414087	0.092416	0
owners	0.026999	0.319026	0.653975	0.837465	0.131792	0.045264
farms	0	0.174707	0.825293	0	0.198312	0.036499
urban	0.437141	0.002709	0.560151	0.387579	0	0.050599
rural	0.033778	0.521369	0.444853	1.749067	0.15848	0.040468
ne-rgn	0.057417	0.342395	0.600188	0.850485	0.217075	0.043474
mw-rgn	0.135745	0.371613	0.492642	0.925708	0.16766	0.044606
s-rgn	0.243396	0.440004	0.316601	0.784831	0.10414	0.045152
w-rgn	0.205905	0.414007	0.380088	0.993516	0.182416	0.050828

Table A.4 Parameters values for *R(t)*

Parameter	q_1	b_1	q_2	b_2	q_3	c
Allhses	0.674803	0.867172	0.254826	0.152536	0.07037	0.044961
renters	0.92074	0.414087	0.07926	0.092416	0	0
owners	0.256138	0.837465	0.476301	0.131792	0.26756	0.045264
farms	0	0	0.590423	0.198312	0.409577	0.036499
urban	0.882241	0.387579	0	0	0.117759	0.050599
rural	0.378544	1.749067	0.529421	0.15848	0.092035	0.040468
ne-rgn	0.339167	0.850485	0.516235	0.217075	0.144598	0.043474
mw-rgn	0.61149	0.925708	0.303189	0.16766	0.085321	0.044606
s-rgn	0.769477	0.784831	0.184578	0.10414	0.045945	0.045152
w-rgn	0.692271	0.993516	0.255567	0.182416	0.052163	0.050828

A.2.3 Important caveats

- The two housing surveys were in 1985 and 1987, so these results may be out of date.
- The analysis requires an ergodic assumption.
- There is no theoretical justification for the shape of the fitted curves.
- We do not know the correct error structure for fitting parametric curves, so there is some element of arbitrariness here. This is emphasized by the finding that different fitting methods give estimates of mean values that differ by more than the estimated standard deviation of the estimated mean values.

A.2.4 References for Section A.2

Israeli, M., and C.B. Nelson, 1992. Distribution and expected time of residence for U.S. households. Risk Analysis 12:65–72.

U.S. Department of Commerce, 1985. With the Bureau of the Census, and U.S. Department of Housing and Urban Development. American Housing Survey for the United States 1985, Current Housing Reports (H-150-85, 28–29, 1988).

U.S. Department of Commerce, 1987. With the Bureau of the Census, and U.S. Department of Housing and Urban Development. American Housing Survey for the United States 1987, Current Housing Reports (H-150-87, 52–53, 1989).

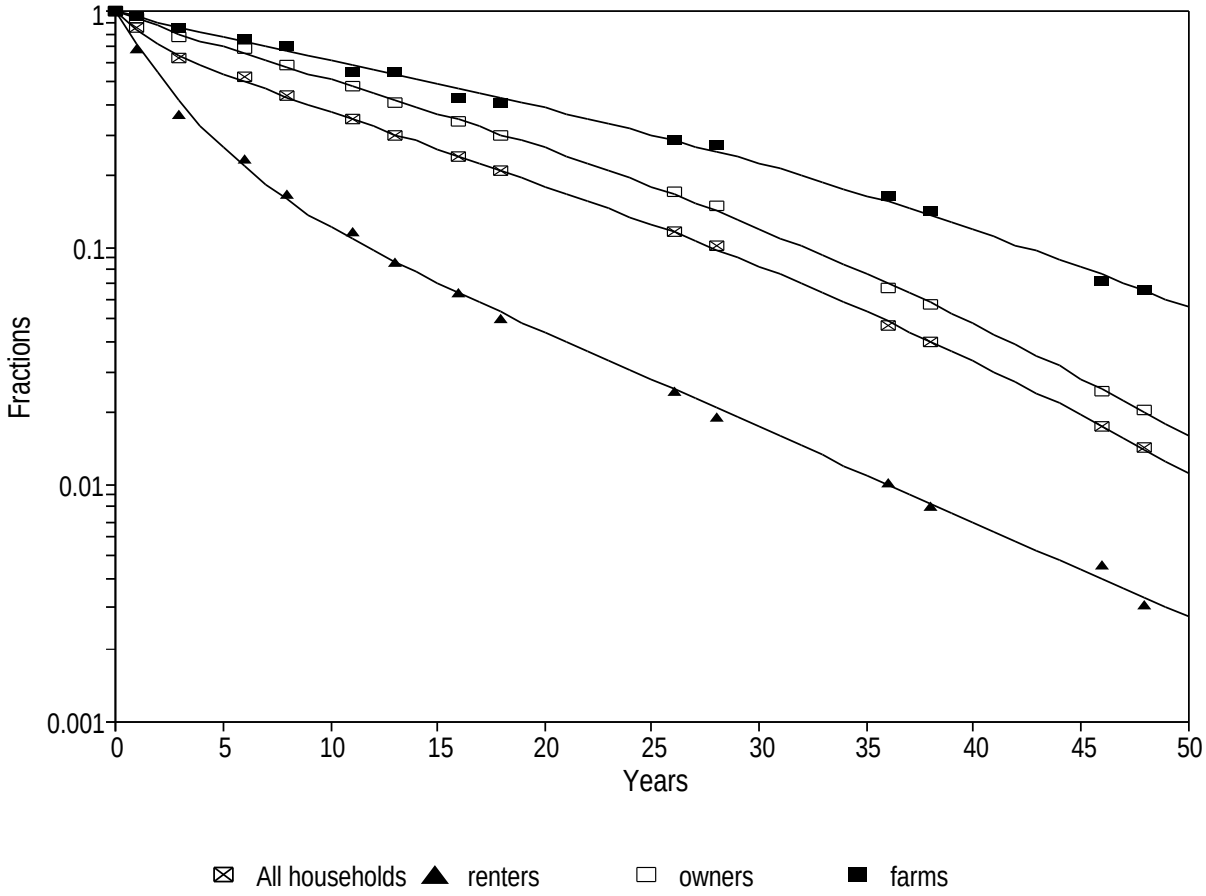

Figure A.1 Observed and fitted values for *S(t)*

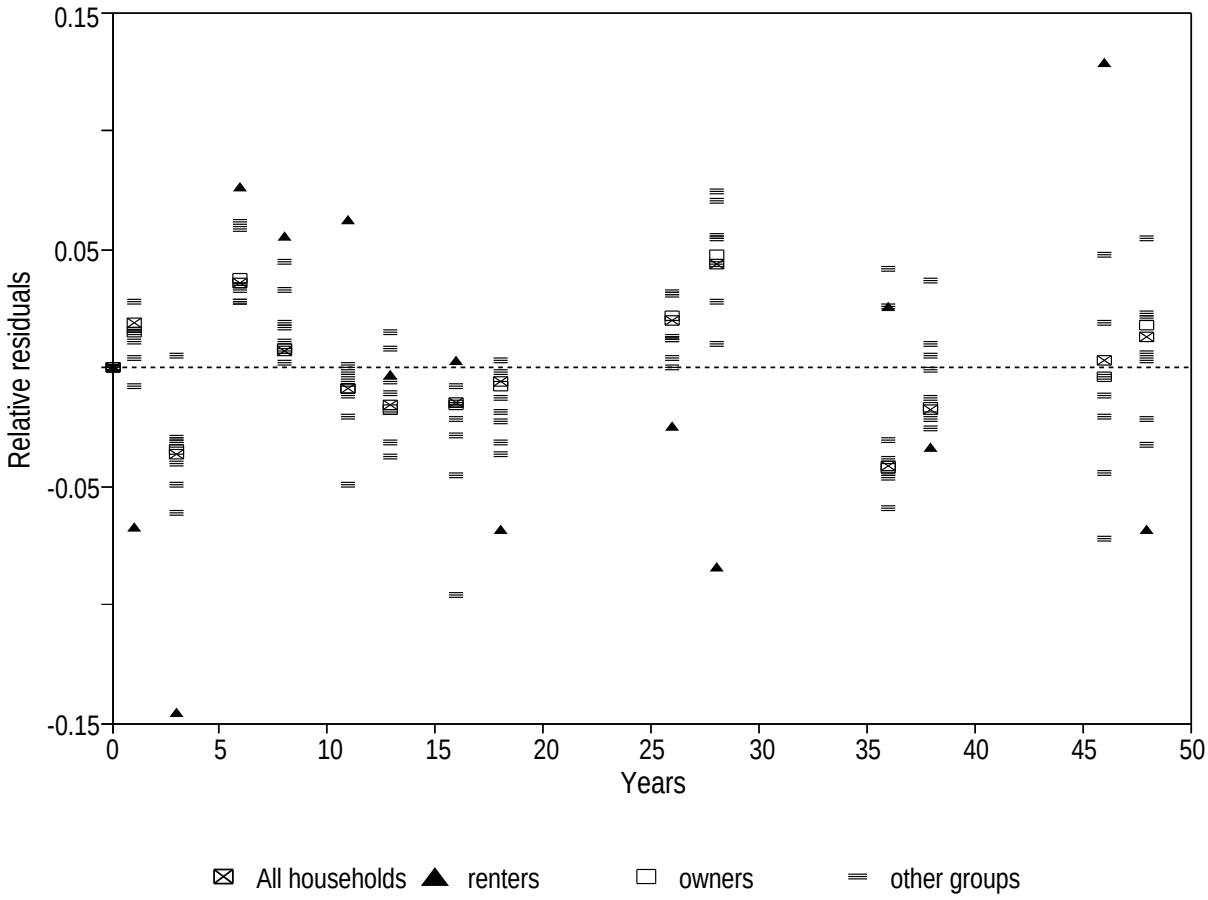

Figure A.2 Relative residuals for all groups

A.3 Showers

A.3.1 Data and distributions

James and Knuiman (1987) have provided some histograms of estimated shower duration and shower flow rates, based on measurements in 2550 homes in Western Australia. The measurements were of total water use in the homes, together with diaries of activities, supplemented by accurate flow measurements over time for appliances in approximately 150 homes. We have digitized the shower duration data with sufficiently high accuracy (estimated total homes in the printed histogram 2494 to 2506, versus actual number 2500) to ensure that the digitization uncertainty is much lower than multinomial variations to be expected in subsamples from a total of 2500. The results of this digitization are given in Table A.5.

Table A.5 Observed shower durations

Shower duration (mins)	Observed number of houses	Shower duration (mins)	Observed number of houses
0–1	5	10–11	135
1–2	12	11–12	105
2–3	61	12–13	56
3–4	166	13–14	46
4–5	318	14–15	32
5–6	394	15–16	26
6–7	372	16–17	13
7–8	287	17–18	20
8–9	239	18–19	15
9–10	190	19–20	8
		20+	50

We found that the cumulative distribution $M(t)$ of shower duration can be very accurately represented by a linear combination of two functions:

$$M(t) = f\Phi(a\ln(t/b)) + (1-f)\frac{\Gamma(A,B/(t+C)) - \Gamma(A,B/C)}{1-\Gamma(A,B/C)} \qquad \text{(A.17)}$$

where Φ is the standard cumulative normal function, and $\Gamma(A,x)$ is the incomplete gamma function given by:

$$\Gamma(A,x) = \frac{1}{\Gamma(A)} \int_{x}^{\infty} t^{A-1} e^{-t} dt \qquad \text{(A.18)}$$

A.3.2 Shower duration and results

We used equation (A.17) to estimate the fraction of houses within each of the observed duration ranges (bins) in Table A.5. The observed number of houses were assumed to be binomial samples from the total number of 2550 houses included (this is negligibly different from the actual multinomial distribution), so that the expected standard deviation of the fraction in each bin could be estimated. The difference between predicted and observed frequencies, expressed as a number of expected standard deviations, was used as a measure of deviation. The sum of the squares of these deviations (approximately a χ_{15}^2 variate) was minimized to obtain the best fits. A plot of the deviations at the best fit value of the parameters appeared to be effectively a standard normal distribution, and the minimized sum had the value 14.8, indicating that no fit could be expected to be better.

The best fit parameters obtained are shown in Table A.6, and a comparison with the observed values is shown in Figure A.3.

Table A.6 Best fit parameter values

Parameter	Value
A	8.616678
B	67.32715
C	1.259006
a	0.91886
b	7.916431
f	0.061406

A.3.3 Number of showers per day

Data were not available describing the number of a times a person showers per day. In the absence of such data, the distribution was estimated to be a triangular distribution with endpoints at 0.25 and 2 showers per day. The peak of the distribution was assumed to be at 1 shower per day.

A.3.4 Important caveats

- The data on shower duration are from a survey in Western Australia during 1981 to 1982. The results may be out of date, and may be inapplicable in other areas (*e.g.* in the U.S.).
- There is no theoretical justification for the shapes of the distributions adopted.
- In deriving the histograms of "data", James and Knuiman (1987) extrapolated from the available measurements. The reliability of these extrapolations is not known.
- The distribution for the number of showers per day is not based on experimental data.

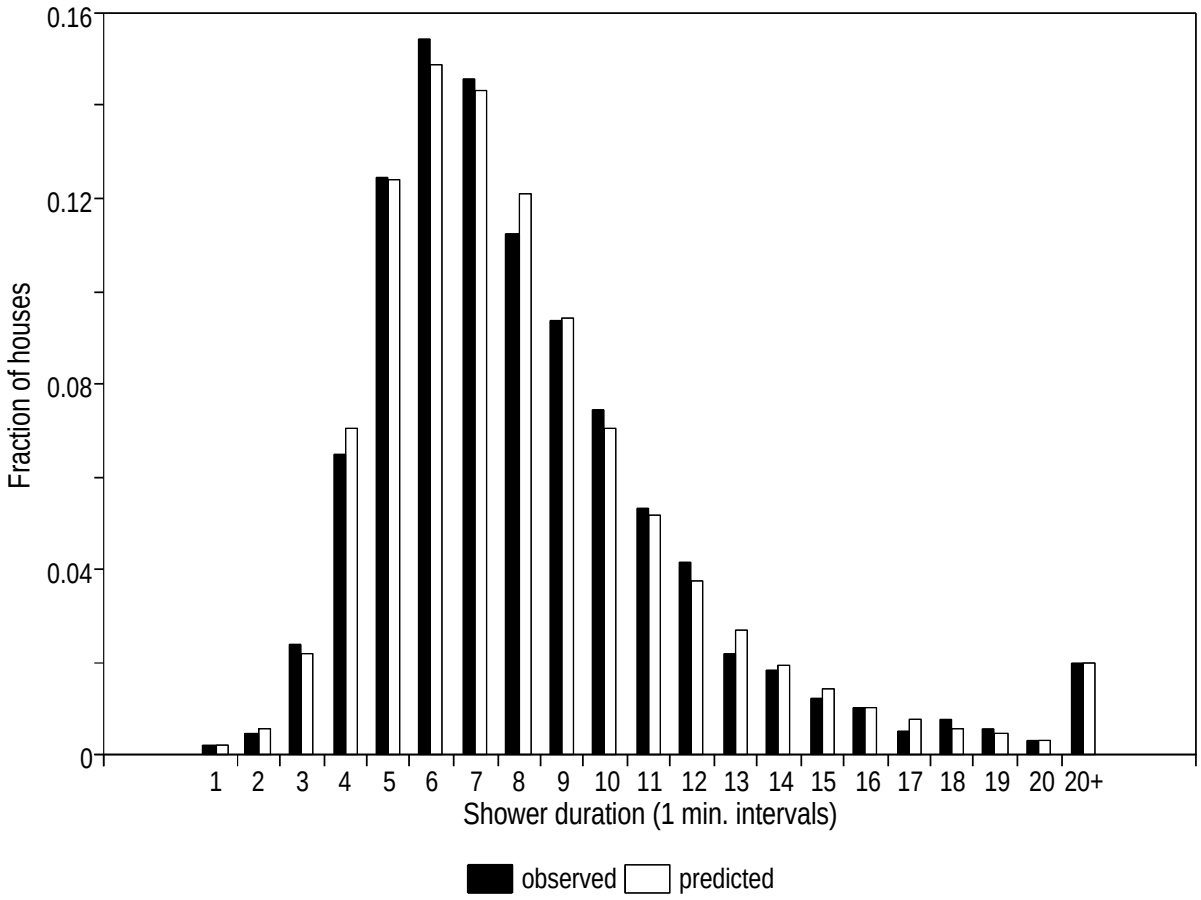

Figure A.3 Comparison of observed and predicted fractions of houses

A.3.5 References for Section A.3

James, I.R., and M.W. Knuiman, 1987. An application of Bayes methodology to the analysis of diary records from a water use study. J. Amer. Stat. Assoc. 82:705–711.

A.4 U.S. Population, 1991

A.4.1 Total population versus age

The distribution by age of the total population varies with calendar year, but it is useful to have such a distribution available to incorporate the major features. Currently those major features are (i) the Gompertz nature of the distribution at large ages, and (ii) the presence of an appreciable baby-boom generation in middle age.

We use the data as given in U.S. Bureau of the Census (1993) for population in 1991 by single years up to age 85, by 5 year intervals from 85 to 100, and the total at age greater than 100. The following functional form provides an adequate fit for the distribution:

$$N(t) = \exp\left(A - Be^{\alpha t} - C(t - t_\eta)e^{-\gamma(t - t_0)^2}\right)$$

where $N(t)$ is the number of persons in the population per year of age, A, B, C, α, γ, t_0 are the fitted parameter, and t is age in years. Fitted values for these parameters are shown in Table A.7. Figure A.4 shows the agreement between data and fitted curve.

Table A.7 Parameters for the population distribution

Parameter	Value
A	8.183238
B	0.004138
α (per year)	0.071444
t_0 (years)	45.02922
C	0.043491
γ (per year)	0.004953

To obtain the values of Table A.7, the logarithm of N as given by equation (A.19) was fitted by least squares to the logarithm of 1991 one-year age group population totals tabulated by the U.S. Bureau of the Census (1993). For ages 85 to 100, only five year age ranges were available, so the sum of one-year values predicted by equation (A.19) were used, with the difference (in logarithms) from the measured values weighted by $1/\sqrt{5}$. A similar approach was taken for the total of all ages above 100 (the sum was carried out to age 120) with the same weighting applied.

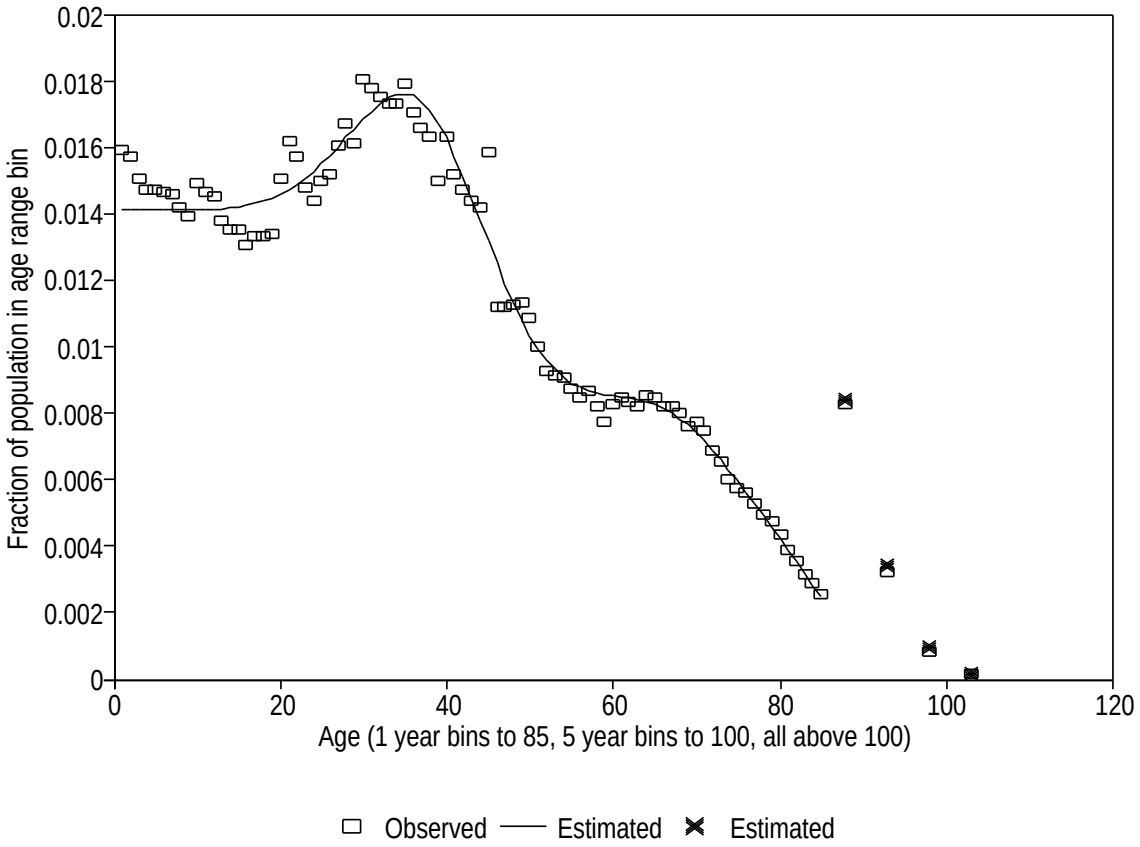

Figure A.4 Population distribution in 1991, observed and fitted

A.4.2 Male fraction of the population

It is sometimes necessary to know how the male fraction in the population varies with age. For 1991, data on this may be obtained from U.S. Bureau of the Census (1993). A reasonable estimate for the male fraction is given by:

$$f_y = a + b\left(0.5 + (1/\pi)\tan^{-\Theta}\left(c\ln\left(t/t_\Theta\right)\right)\right) \qquad \text{(A.20)}$$

where f_m is the male fraction at age t (years), and a, b, c, t_1 are fitted parameters. The parameter values are given in Table A.8, and the fit to the empirical values is shown in Figure A.5.

The fit was obtained by minimizing the sum of squares of the logarithms of the ratio of values given by equation (A.20) to the empirical values.

Table A.8 Parameters for male fraction of population

Parameter	Value
a	0.140646
b	0.377303
c	7.502904
t_1 (years)	83.96292

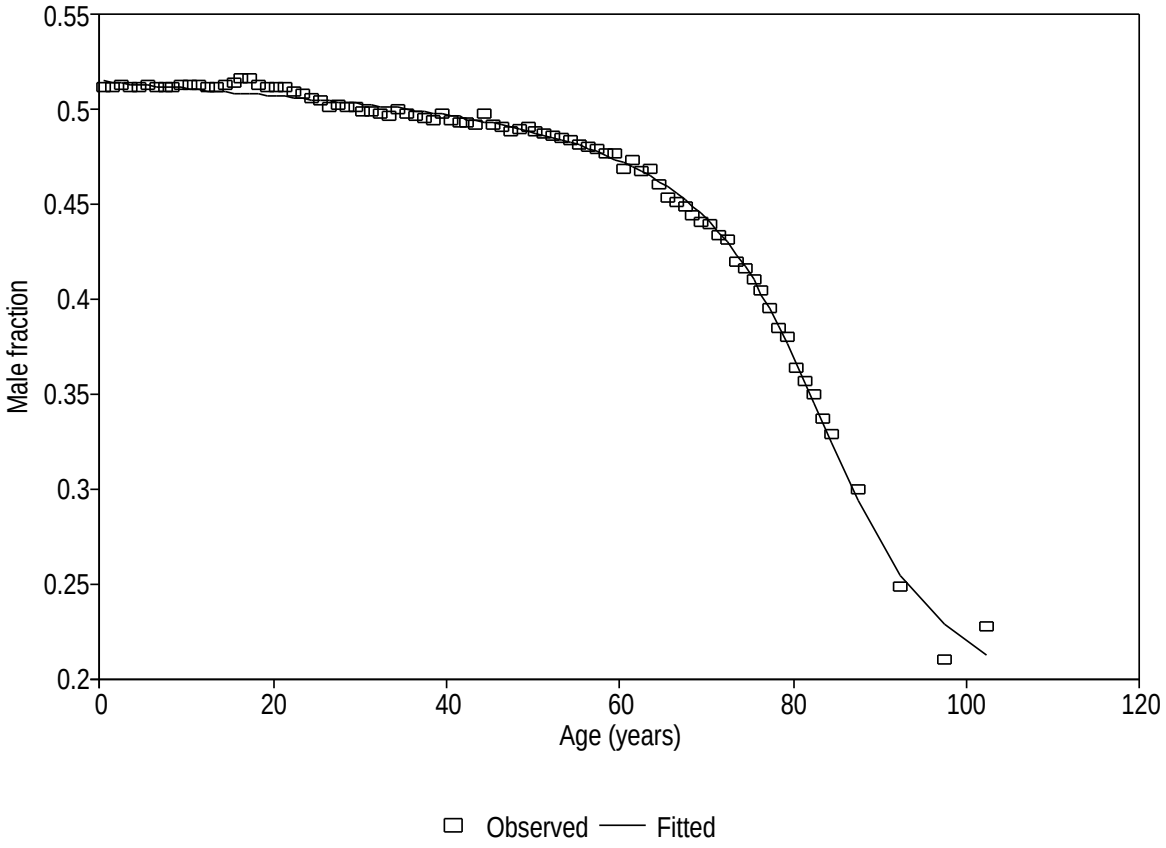

Figure A.5 Male fraction of population, 1991

A.4.3 Important caveats

- The population distribution strictly applies to 1991 only, and is only designed to produce smoothed (over ages) estimates. It is adequate for many risk assessment purposes where only a rough idea of the population distribution is needed.
- The fit to the population distribution is not very good for the changing population distribution in the age range 0 to 20.
- The male-to-female ratio also strictly applies to 1991 only. It is probably adequate for many risk assessment purpose, since it is unlikely to change substantially for some time.
- A better approach for estimating population age distributions would be to explicitly incorporate Census Bureau population projections, to take account of projected birth and death rates. However, this was too complex to do in this context.

A.4.4 References for Section A.4

U.S. Bureau of the Census, 1993. Statistical Abstract of the United States: 1993 (113th edition). Washington, DC.

A.5 Body weight

A.5.1 Data

Body weight data were collected as part of the second National Health and Nutrition Examination Survey, conducted from 1976 to 1980. The distribution used in the risk assessment is based on percentile data reported in *Vital and Health Statistics* (NCHS, 1987). Data are presented for each sex for the age groups of 6-11 months, each year until age 20, — 18-24 years, and every ten years from age 25 until age 75.

A.5.2 Body weight as a function of age

Examination of the data for each age group shows that the body weights are approximately lognormally distributed. The mean (μ) and standard deviation (σ) as functions of time were parameterized using equations of the form:

$$\mu(t) = a_\eta + a_\Theta t + a_\theta t^\theta + b\exp\left(-t(c_\eta + c_\Theta t + c_\theta t^\theta)\right)$$

$$\sigma(t) = A + (B_\eta + B_\Theta t + B_\theta t^\theta + B_\iota t^\iota + B_\iota t^\iota + B_K t^K)\exp\left(-(ct)^\iota\right)$$

where t is the age of the individual in years.

There is no theoretical justification for the selection of a model of this form. This form was chosen because it accurately fits the experimental data. The power in the exponential of the standard deviation fit (σ) has been set to 4. This value is a compromise between the best value for males (the best fit is at an exponent of 7 or 8) and the value for females (the best fit is 2). The best values for the parameters in the two equations above are summarized in Table A.9 below. Figure A.6 and Figure A.7 below show the means and standard deviations of the natural log transformed data and the best fit to the data for males. Figure A.8 and Figure A.9 show the same information for females.

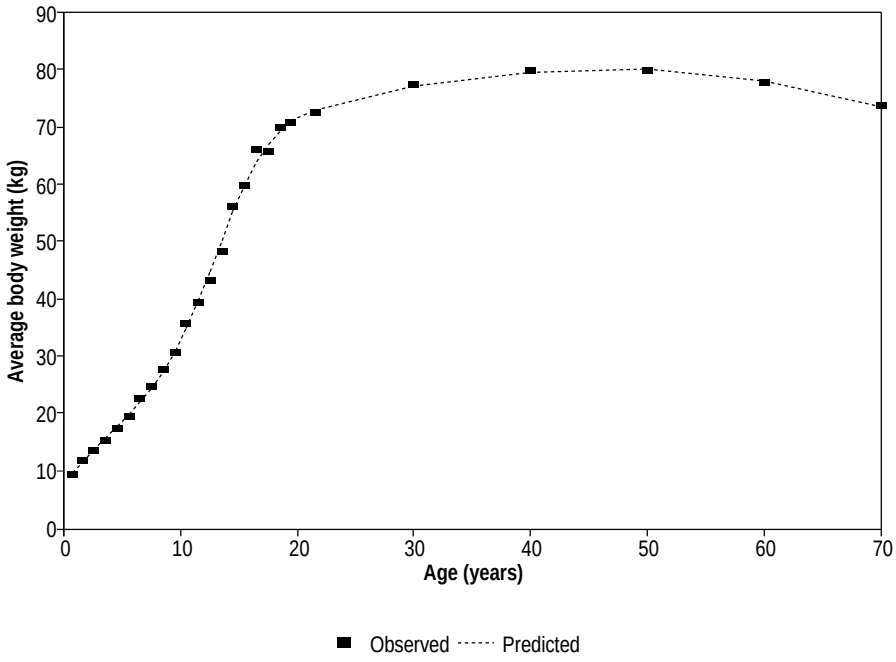

Figure A.6 Average body weight for males as a function of age

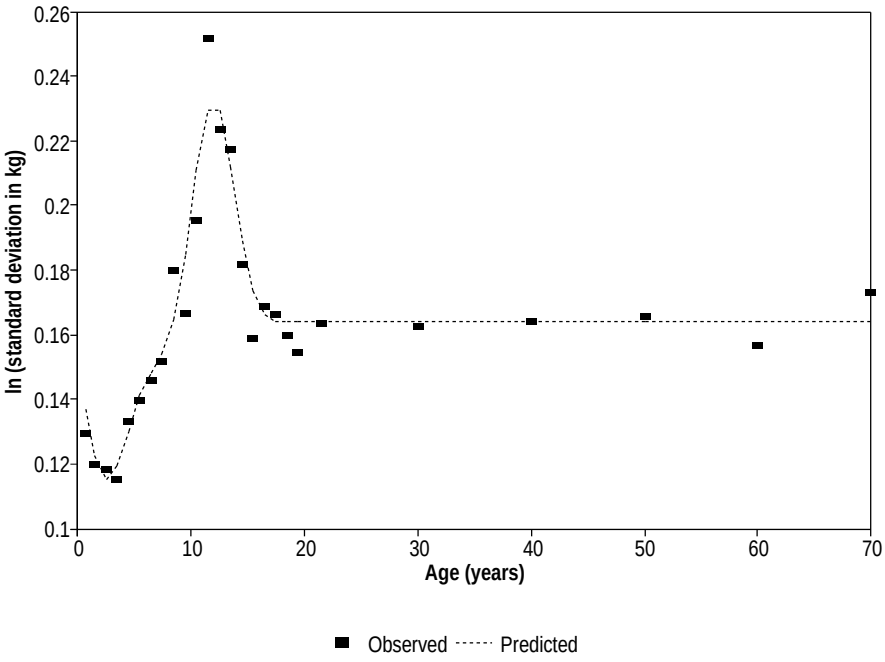

Figure A.7 Standard deviation of the natural log transformed body weight data for males

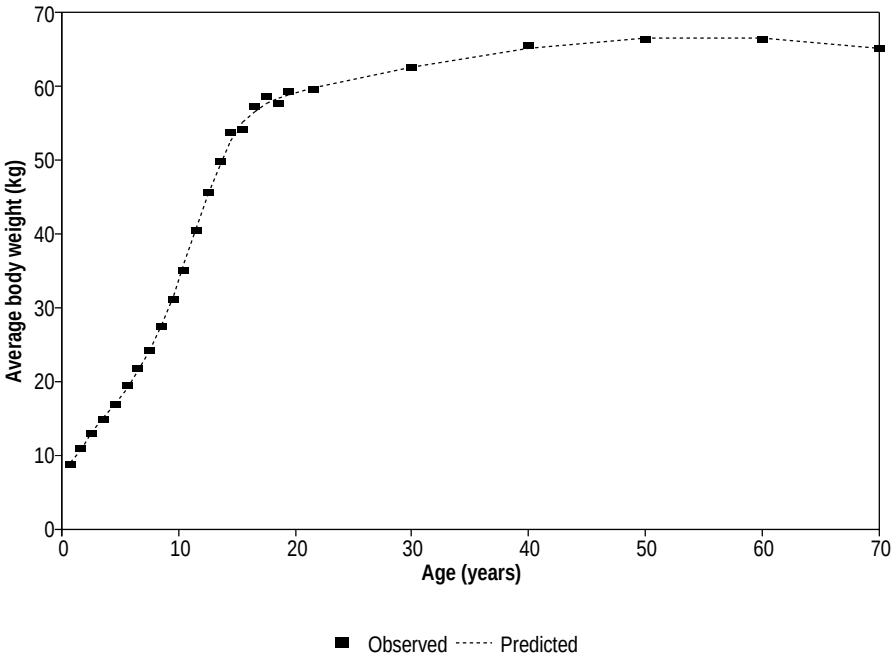

Figure A.8 Average body weight data for females as a function of age

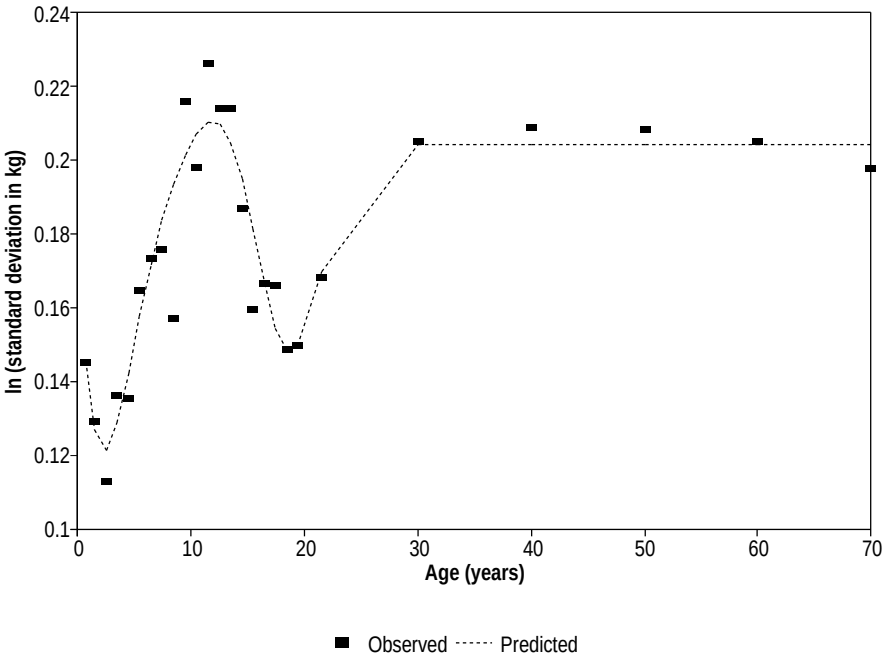

Figure A.9 Standard deviation of the natural log transformed body weight data for females

Table A.9 Best fit values for body weight parameterization

Parameter	Males	Females
a_0	4.06661e+00	3.90383e+00
a_1	1.37603e-02	1.06805e-02
a_2	-1.49397e-04	-9.72316e-05
b	-2.00685e+00	-1.97114e+00
c_0	1.47094e-01	1.84877e-01
c_1	-1.39216e-02	-2.04753e-02
c_2	9.77090e-04	1.59966e-03
A	1.63829e-01	2.03999e-01
B_0	-6.97297e-03	-2.12621e-02
B_1	-2.75757e-02	-6.50109e-02
B_2	4.34040e-05	2.25515e-02
B_3	2.86364e-03	-3.00811e-03
B_4	-5.31960e-04	1.89083e-04
B_5	2.83428e-05	-4.65941e-06
C	1.01888e-01	6.57724e-02

A.5.3 References for Section A.5

National Center for Health Statistics (NCHS), M. F. Najjar and M. Rowland: Anthropometric Reference Data and Prevalence of Overweight, United States, 1976-80. *Vital and Health Statistics*, 1987. Series 11, No. 238. DHHS Pub. No. (PHS) 87-1688. Public Health Service, Washington.

A.6 Water consumption, 1977-78

A.6.1 Data source

Water consumption estimates suitable for estimating distributions have been made by Ershow and Cantor (1989), based on the 1977–1978 Nationwide Food Consumption Survey (NFCS) of the U.S. Department of Agriculture. The study design of the NFCS allowed a weighting to obtain population estimates, and Ershow and Cantor provide such estimates for mean and standard deviation for the parameters they estimate. In addition, Ershow and Cantor provide percentiles of the distribution of raw values in the survey data. These distributions allow some estimation of the shapes of the distributions, but they cannot be used directly because the weighting factors have been omitted.

A.6.2 Tapwater intake per unit body weight

Ershow and Cantor (1989, Table 14) provide estimates for weighted mean and standard deviation, and unweighted percentiles, for each sex and for age groups [0,0.5), [0.5–1), [1–4), [4,7), [7,11), [11,15), [15,20), [20,45), [45,65), [65,75), [75+], of the U.S. population. It is mentioned in the text that the values and distributions for infants under the age of one may be misleading, because of technical difficulties in evaluating the tapwater component of such infants' intake.

Examination of the (unweighted) distributions indicates that they are practically indistinguishable from lognormal in each age group (although the distribution for the whole population is not lognormal). However, we chose to use the given weighted means and SD estimates (rather than the unweighted percentiles) as the basis for what follows, in order to correctly obtain weighted population estimates.

We require a smooth interpolation function for the parameters of a lognormal variability distribution for tapwater intake per unit body weight at each age. The functional form chosen was:

$$A + (B + Ct + Dt^{\theta}) \exp(-\lambda t^{\varrho}) \qquad \text{(A.22)}$$

for both mean (μ) and standard deviation (σ) of the logarithm of tapwater intake per unit body weight (measured in g/kg-day), where A, B, C, D, λ, and k are fitted constants, with $k = 1$ for the standard deviation. This form was chosen simply to give a good fit to the available data, and to look sufficiently smooth.

The values of A, B, C, D, λ, and k were obtained by a fitting procedure, using all the age ranges above one-year. First, the observed standard deviation (s) for each age range was slightly modified to take account of the variation of the mean value of tapwater intake across each age range included in the data. The observed mean values (m) were assumed to apply to the center of each age range involved (80 for the oldest group), and interpolated linearly between those central ages to estimate values (m_l and m_u) at the lower and upper end points of each age range. The corrected value s_c was obtained as:

$$s_{\ddot{Y}} = \sqrt{s^{\theta} - (m_{..} - m_{s})^{\theta}/12} \qquad \text{(A.23)}$$

by assuming a uniform distribution of values for *m* within each age range. For the highest age range, the estimated correction was taken to be identical to that for the next lower age group (since there was little variation with age at large ages). This procedure resulted in very small corrections to the *s*.

The observed estimate of arithmetic mean (*m*) and corrected standard deviation (s_c) were fitted to the functional forms (A.22) for μ and σ by minimizing the weighted sums of squared deviations:

$$\sum \frac{(m - \xi)^{\theta}}{\sigma_{\ddot{Y}}^{\theta}} + \frac{(s_{\ddot{Y}}^{\theta} - \xi^{\theta} t)^{\theta}}{\sigma_{,}^{\theta}} \qquad \text{(A.24)}$$

where:

$$\xi = e^{\mu + \sigma^2/2\theta}$$
$$t = e^{\sigma^2} - \Theta \qquad \text{(A.25)}$$

and the expected variances (approximate for s^2) used as weights are given by:

$$\sigma_{\ddot{Y}}^{\theta} = \xi^{\theta} t/n$$
$$\sigma_{,}^{\theta} = (\xi^{\theta} t)^{\theta}(e^{\theta\sigma^2} + 2e^{\theta\sigma^2} + 3e^{\theta\sigma^2} - 4)/n \qquad \text{(A.26)}$$

where *n* is the (unweighted) number of observations for the age range.

Results for the parameters *A*, *B*, *C*, *D*, λ, and *k* are shown in Table A.10 for males, and Table A.11 for females. Comparison of these parametric fits with the original data are shown in Figure A.10 and Figure A.11.

Table A.10 Water consumption per unit body weight, Males

Parameter	μ	σ
A	2.978778	0.410696
B	0.53693	0.218693
C	0.327883	-0.02207
D	-0.03348	0.002701
λ	0.187475	0.087092
k	0.91258	1

Table A.11 Water consumption per unit body weight, Females

Parameter	μ	σ
A	3.019445	0.289118
B	0.610601	0.270106
C	0.071572	0.006551
D	-0.01294	0.000456
λ	0.026765	0.043997
k	1.441246	1

A.6.3 Important caveats for tapwater intake per unit body weight

- There is no theoretical justification for the functional forms (A.22). They merely provide a reasonable fit to the data.
- For ages less than one-year, the estimates obtained, while reasonable, bear no relation to the measurements. The measurements are themselves suspect in this age range.
- The data were obtained in 1977–1978. Since then, the consumption of carbonated soft drinks has increased by approximately 50%, from approximately 300 gram/person-day to 450 gram/person-day (based on estimates of deliveries to market and resident population, U.S. Bureau of the Census, 1993). Since total water intake is unlikely to have changed much, most of this change has probably been in tapwater intake. Since the change amounts to around 2 g/kg-day for adults, or 10% of the total tapwater intake, there could have been a significant change in mean tapwater intake.
- The estimates obtained by Ershow and Cantor (1989) for components of the intake of water do not agree very well with those obtained by evaluation of apparent market

consumption (U.S. Bureau of the Census, 1993). The differences are considerable for soft drinks, beer and wine. Such comparisons are difficult, however, because of potentially different definitions. The Census Bureau estimates for any given year have themselves varied from year to year for no apparent reason.

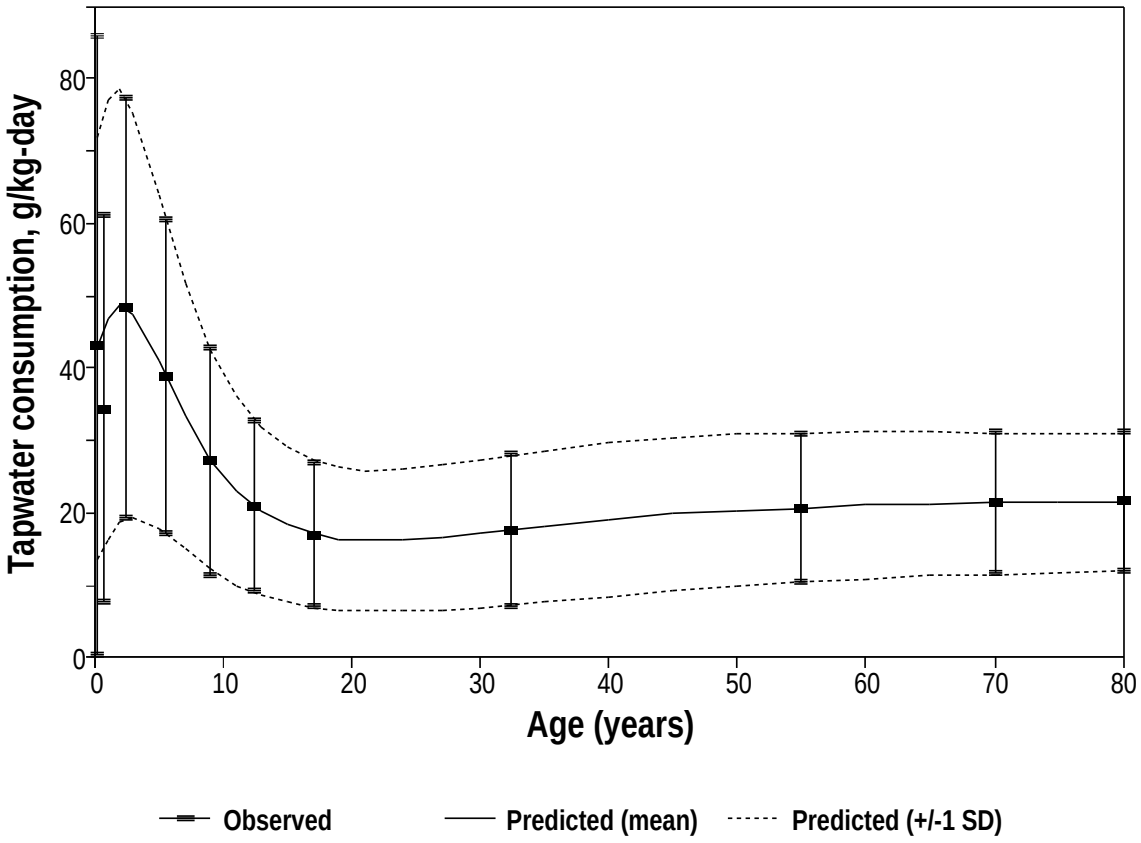

Figure A.10 Water consumption per unit bodyweight, arithmetic mean and standard deviation, for males, as a function of age

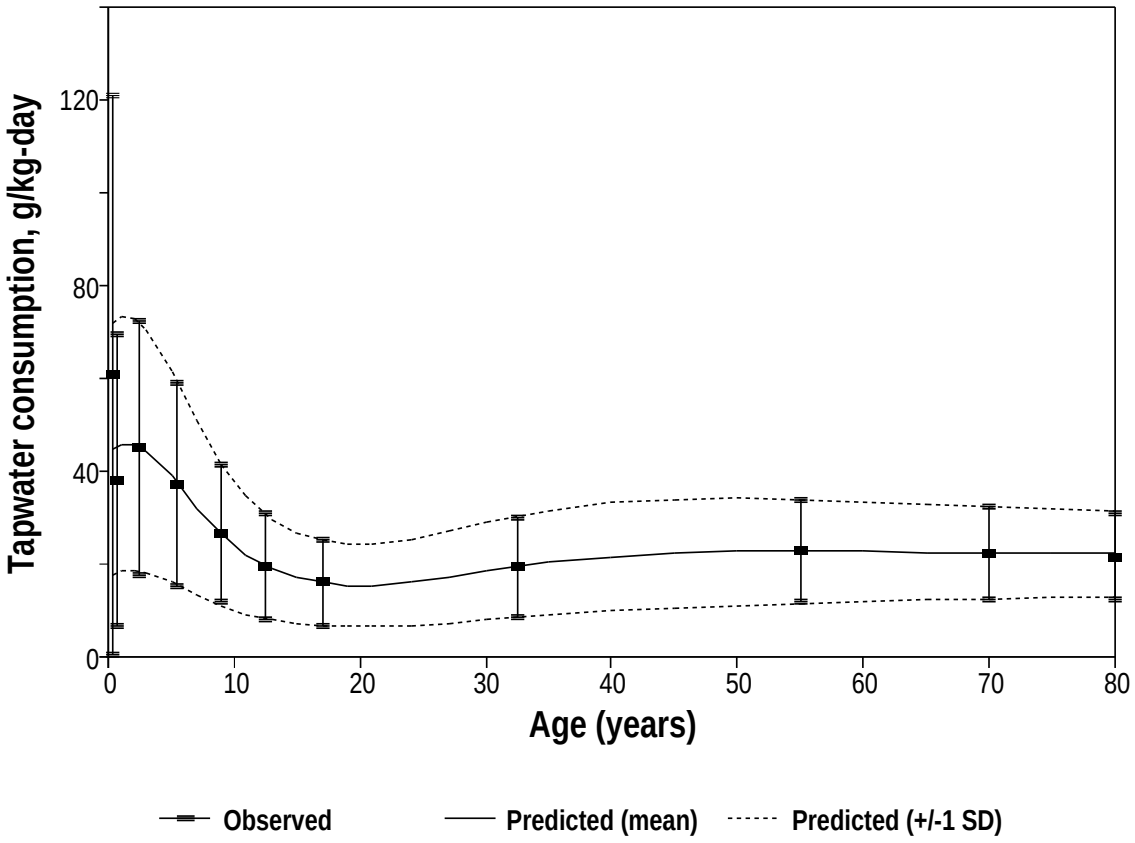

Figure A.11 Water consumption per unit bodyweight, arithmetic mean and standard deviation, for females, as a function of age

A.6.4 Fraction of water consumed at home

No studies describing the fraction of water ingested from the domestic water supply were available. We assumed a uniform distribution of fractions ranging from 0.5 to 1.0 with a central estimate of 0.75. The value 1.0 is used for this fraction when calculating the best estimate under U.S. EPA guidance.

A.6.5 References for Section A.6

Ershow, A.G., and K.P. Cantor, 1989. Total water and tapwater intake in the United States: population-based estimates of quantities and sources. Life Sciences Research Office, Federation of American Societies for Experimental Biology, MD 20814. May 1989. Prepared under National Cancer Institute Order #263-MD-810264.

U.S. Bureau of the Census, 1993. Statistical Abstract of the United States: 1993 (113th edition). Washington, DC.

A.7 Life expectancy

The life expectancy for an individual was estimated from the age of the individual based on a fit to data in the *Statistical Abstract of the United States: 1993* (U.S. Bureau of the Census, 1993). A graph of average remaining life as a function of age is approximately linear until age 50, at which point it can be more closely approximated by an exponential. Life expectancy for individuals in the risk assessment was calculated as:

$$T_{rem} = 72.807 - 0.931\,T \quad \text{for } T < 50$$
$$T_{rem} = 27.323\,e^{-0.0152\,(T - 50)} \quad \text{for } T \geq 50 \qquad \text{(A.27)}$$

where T is the age of the individual and T_{rem} is the remaining life expectancy for the individual. The value of T in Equation (A.27) above is simply T_{rem} plus the original age of the individual. Figure A.12 below is a plot of the life expectancy data and the equations above.

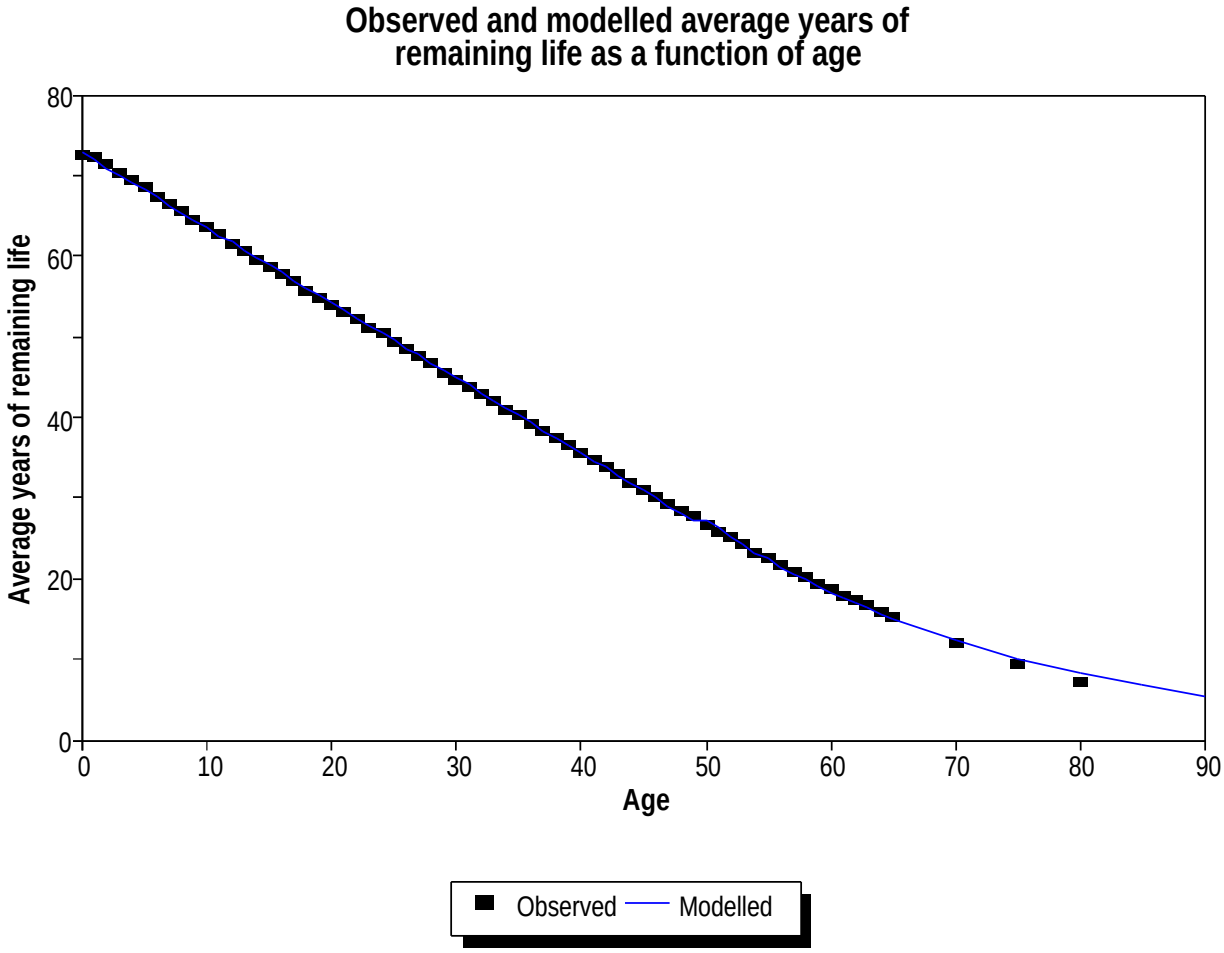

Figure A.12 Average remaining life expectancy as a function of age

A.8 Surface area

The surface areas used to calculate dermal exposures to 1,1-DCE while showering were the 50th percentile surface areas for males and females (Anderson *et al.*, 1985). Table A.1 below lists the surface areas for individuals as a function of age.

Table A.1 Surface area as a function of age

Age (years)	Surface Area (m^2)		Age (years)	Surface Area (m^2)	
	Males	Females		Males	Females
<1	0.465	0.436	10	1.18	1.17
1	0.532	0.505	11	1.23	1.30
2	0.603	0.579	12	1.34	1.40
3	0.664	0.649	13	1.47	1.48
4	0.731	0.706	14	1.61	1.55
5	0.793	0.779	15	1.70	1.57
6	0.866	0.843	16	1.76	1.60
7	0.936	0.917	17	1.80	1.63
8	1.00	1.00	18+	1.94	1.69
9	1.07	1.06			

Appendix B: Obtaining distributions for carcinogenic potencies derived from studies on laboratory animals

Carcinogenic potency estimates are generally based either on data from experiments in laboratory animals, or on results from epidemiology. The analyses of the different types of data to obtain distributions are different for the two cases. However, the underlying philosophy (using likelihood methods) is similar, so similar that the results from both types of data can be consistently combined if both are available.

This appendix starts by describing the current method for obtaining an estimate of carcinogenic potency from a single experiment on one species/strain/sex of laboratory animal. The method is then extended to allow calculation of a probability distribution for the carcinogenic potency in that laboratory animal; and modified slightly to make it self-consistent. In this modified form, the method may then be used to combine different experiments on laboratory animals.

The results from laboratory animals are extrapolated to humans. We describe the current method for performing this extrapolation, and describe the empirical data available on such inter-species extrapolation. From the empirical data, we propose a probabilistic extrapolation method that may be combined with the distributions obtained from animal experiments to obtain a probabilistic estimate for the carcinogenic potency in humans.

It should be noted that all the methods described here make an assumption that the models used for analysis of either animal data or epidemiology are correct, in the sense that they correctly extrapolate the dose-response curve to unmeasured (and possibly unmeasurable) regions (*e.g.* to low doses, or to greater age, and so forth). More precisely, there are various combinations of conditions under which the methods described in this Appendix would give accurate results. For example, if:

1 the standard multistage dose-response formula accurately represents the dose-response curve for laboratory strains of mice and rats in the experimentally measurable region (greater than about 10% response); and

2 a simple one-hit dose-response model accurately extrapolates from high doses (or exposures), corresponding to epidemiologically measurable relative risks, to extremely low doses (or exposures) *in humans*; and

3 the observed correlations between potency parameters (Sections B.6 and B.7) are biologically meaningful, and not an accident of experimental designs; and

4 various other assumptions used in the analyses (*e.g.* that the results of animal bioassays are binomial samples, that all the animals receive the same dose) are correct;

then the methodology derived here would be accurate. Other combinations of equally implausible conditions may also suffice. However, it has to be realized that there are some currently unpredictable sources of uncertainty — the extrapolation to low doses being perhaps the best example — that probably still dominate. The uncertainties that are derived using the methods to be described have to be considered incomplete; the true uncertainty is

almost certainly larger in some respects. However, the assumptions going into the selection of models is generally held to be conservative, in the sense that the estimates of human potency obtained from the model will be overestimates. Thus, any uncertainties due to selection of the wrong models are likely to be biased — the true potencies are likely to be lower rather than higher.

B.1 The current approach

The method starts with the current approach, so this will be summarized (see Anderson *et al.*, 1983). It is assumed that the dose-response relationship for tumor incidence at the end of a lifetime bioassay can be represented by:

$$p(d) = 1 - \exp\left(-\sum_{j=1}^{m} q_j d^j \right) \qquad \text{with} \quad q_j \geq 0 \quad \text{for all } j \tag{B.1}$$

In this formula the terms are:

$p(d)$	Lifetime probability of tumor at a lifetime average dose rate d, and
q_j	$0 \leq j \leq m$, the $m + 1$ parameters of the model.

The end-of-life tumor data may be characterized by an array of values:

$$\begin{array}{lll} d_1 & r_1 & n_1 \\ d_2 & r_2 & n_2 \\ \ldots & \ldots & \ldots \\ d_k & r_k & n_k \end{array} \tag{B.2}$$

where r_i of the n_i animals in the dose group receiving dose d_i are observed to have a tumor at the end of their lifetime, for $i = 1,2,\ldots k$. Here k is the total number of dose groups, and in the standard method the value of m is chosen to be $k - 1$, so that there are the same number of parameters as dose groups. The method of fitting the parameters is maximum likelihood. At this point, the choice of type of tumor to be analyzed is not specified: it is assumed that a single such array of values can be selected for each bioassay.

The parameters q_j of this model are estimated by fitting the dose-response relationship to the results using maximum likelihood. With all sorts of assumptions about the similarity of the animals, their dosing, and their treatment, the observed number r_i out of n_i at dose rate d_i may be taken to be a binomial sample with probability $p(d_i)$. Then the likelihood for the set of results given in the scheme in (B.2) is given by:

$$L = \prod_{x \in \Theta} \binom{n_x}{r_x} p_x^{r_x} (1 - p_x)^{n_x - r_x} \quad \text{where} \quad p_x = p(d_x) \tag{B.3}$$

so that the loglikelihood is, apart from trivial constants that may be omitted:

$$J = \sum_{x \in \Theta} r_x \ln p_x + (n_x - r_x) \ln(1 - p_x) \tag{B.4}$$

where J should be considered a function of the parameters q_j.

The value of the maximum J_0 of this function is obtained by maximizing with respect to all the parameters q_j, subject to the restriction $q_j \geq 0$, and then a single value q_1^* is derived by applying the asymptotic theory of maximum likelihood to obtain an approximate 95% upper confidence limit on q_1. Asymptotically the function $f(q_s)$ defined by:

$$f(q_s) = 2(J_0 - J'(q_s)) \tag{B.5}$$

should be χ^2 distributed with one degree of freedom, where:

J_0	is the maximum value of the loglikelihood function, and
$J'(q_s)$	is the value of the loglikelihood when re-maximized with parameter number s held fixed at the value q_s.

Equivalently, $h = \sqrt{f}$ should be normally distributed with zero mean and unit variance (the sign of the square root is the same as the sign of the deviation of q_s from its maximum likelihood value).

Specifically, q_1^* is obtained as the larger solution of:

$$f(q_0) = 2(J_0 - J'(q_0)) = \chi^2_{1,1-2\alpha} \tag{B.6}$$

where $\chi^2_{1,1-2\alpha}$ is the (1-2α) point of the χ^2 distribution with one degree of freedom and $\alpha = 0.05$. This q_1^* is the basis for the values commonly known by some such term as the "carcinogenic potency slope factor," and published by EPA in IRIS or the HEAST (an extra factor for interspecies extrapolation is applied to obtain the published values). There are several computer programs available to evaluate q_1^* and similar estimators of potency (*e.g.* Crump, Howe, and Landingham, 1987; Crouch, 1992).

B.2 A distribution from the current approach

The confidence limit on q_1 defined in Section B.1, equation (B.6), is applied with $\alpha = 0.05$ to obtain the point estimate q_1^* in the standard approach. However, simply varying α and obtaining both solutions of equation (B.6) immediately gives a cumulative distribution for q_1, and the differential distribution is straightforward to obtain by differentiation. This cumulative distribution may be expressed as:

$$\Pr(q_\Theta) = \Phi\left(\pm\sqrt{2(J_\eta - J'(q_\Theta))}\right) \qquad \text{(B.7)}$$

where Φ is the cumulative standard normal function (B.6) and the sign of the square root is chosen to be the same as the sign of the deviation of q_1 from its maximum likelihood value. Such a definition of the cumulative distribution for q_1 is perhaps the closest approximation that can be obtained to the current approach.

An example of the distribution obtained by this approach may be seen in Figure B.1, based on the results obtained in an experiment with acrylonitrile in the drinking water of Sprague-Dawley rats summarized in Table B.1 (corresponding to Table 13-35A in EPA, 1983). Both cumulative and differential distributions are plotted, with scales on the left and right of the figure respectively.

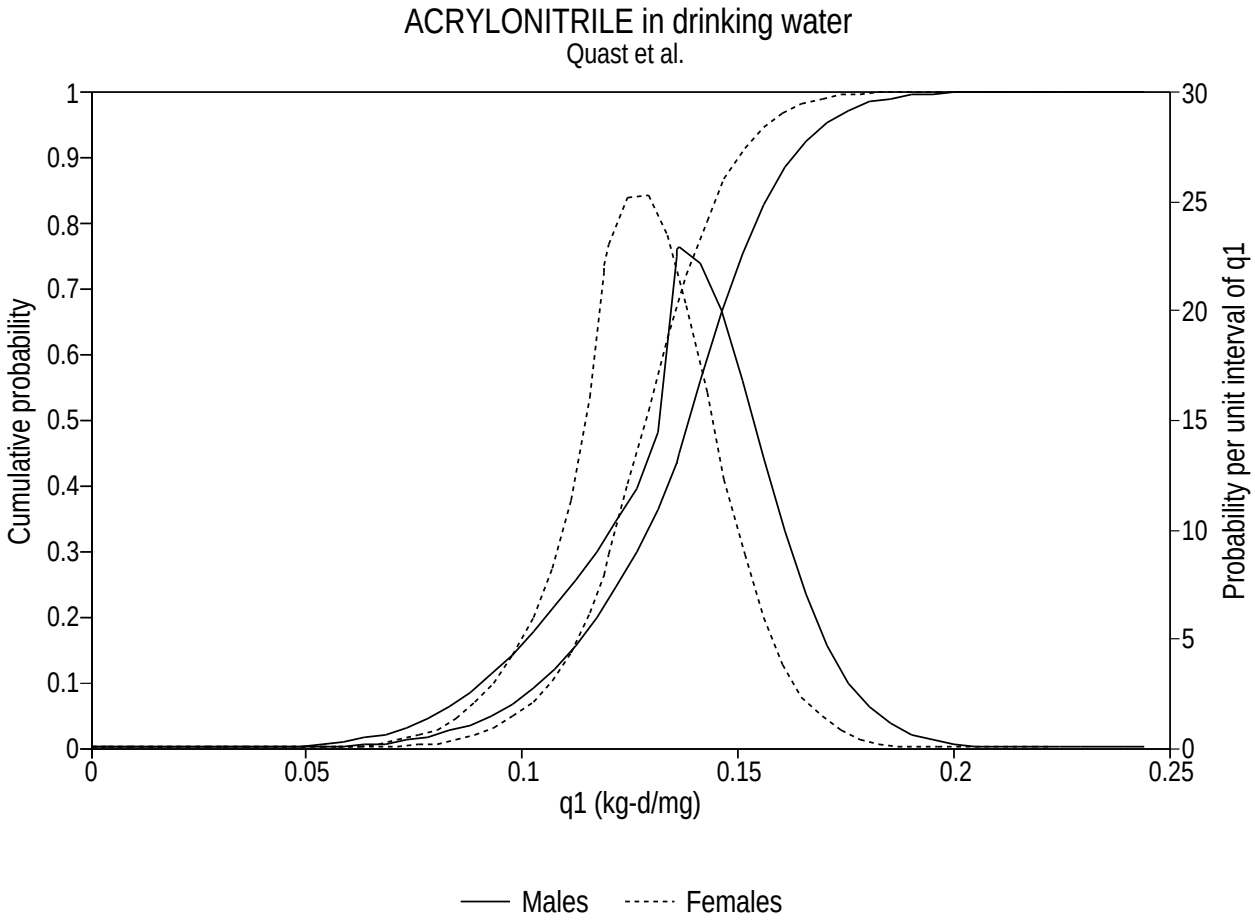

Figure B.1 Cumulative and differential distribution for q_1 for rats (see Table B.1)

(It should be noted that the differential distribution is continuous, although it may have kinks in it where its derivative exhibits finite jumps.)

Table B.1 End of life data for rats fed AN in their drinking water (Quast *et al.*, 1980; as cited in EPA, 1983)

(multiple tumor types and sites, excluding mammary gland tumors in females)

	Males			**Females**	
Average dose mg/kg-d	**Number with tumor**	**Number at risk**	**Average dose mg/kd-d**	**Number with tumor**	**Number at risk**
0	4	80	0	3	80
3.42	18	47	4.36	24	48
8.53	36	48	10.76	37	48
21.18	45	48	24.97	45	48

While this procedure may be the closest possible to the current approach, it has substantial disadvantages, some of which are associated with the arbitrariness of the current approach. The main disadvantages that we have identified are:

- Fixing the number of parameters to include in the dose-response curve leads to substantial variation of the distributions (and also the 95th percentile) with inclusion or exclusion of dose groups. Such variations are due entirely to the arbitrary selection of the maximum power in the dose-response curve to be one less than the number of dose groups.

- In many cases, the density function vanishes for the whole range of values below the 50th percentile (or some other percentile) with a delta function of strength 50% (or other value) at zero. This is unreasonable. What the data is indicating is indifference to (equal likelihood for) any values less than the 50th percentile.

- With this definition for the density function, it is difficult to consistently add in other experimental or epidemiological results in a way that is mathematically symmetrical and that agrees with probability theory.

- The density function can have some rather strange and unnecessary properties — for example, square root singularities — that do not seem to be a reflection of any real characteristics of the data.

B.3 Using the likelihood function to obtain a distribution

The problems discussed in Section B.2 can be overcome simply and consistently by:

- Adopting the likelihood principal — by treating the likelihood function as an empirical distribution function.

- Removing the unphysical restriction on the highest power allowed in the model. Some restriction is still necessary, but this should be based on a physical idea that the dose-response curve cannot plausibly be expected to increase faster than, say, the 7^{th} power of dose.

The first suggestion may be philosophical anathema for some statisticians, but for the small samples involved here is probably as good a procedure as the current approach, and is quite practical. It has the minor disadvantage that it requires a numerical integration to obtain percentiles of the distribution, but the whole point of this method is the use of the complete distribution.

The second suggestion removes the unphysical behavior of the distribution under addition or removal of results at single doses. It can be implemented in a straightforward fashion, since it corresponds to the current approach with a large number of parameters allowed to be non-zero. Except in degenerate cases, the actual number of parameters that are non-zero is limited to the number of dose-groups automatically, by the imposition of the positivity constraints in equation (B.1), and the degenerate cases are taken care of by the distributional approach (the density function is flat — *i.e.* the likelihood is constant — over the regions of degeneracy).

B.4 Adding in extra information

There are at least two ways in which extra information may be incorporated into the analysis described so far. All that is involved is a modification of the likelihood function, since this defines the distribution.

The first obvious modification is to add in some probability for a chemical to not be a carcinogen at all at low enough doses. In the procedure described in Section B.2, a (potentially high) probability for non-carcinogenicity arises automatically, since the cumulative distribution necessarily starts at the (approximate) value p of the significance for the linear term in the dose-response curve — *i.e.* there is a probability of p that the potency is zero, and p can be large (it is quite typically 0.5). In the modification of Section B.3, this possibility of zero potency is entirely eliminated, since the likelihood function can only indicate indifference between small values of potency (zero has no preferred role). This approach thus ignores the weight-of-evidence that may be available as to whether the chemical under study is a carcinogen at all at low doses. However, it is straightforward to augment the distribution obtained in Section B.3 to include a probability of zero potency — all that is mathematically required is to add a delta function of the required strength at the origin. The size of the probability for zero potency has to be determined by some other method, for

example the best judgement of toxicologists; or perhaps, by the *p*-value for the dose-response curve as a whole (*i.e.* a measure of the objective probability that the experiment is demonstrating an increased tumor response at any dose). This last approach will be used here, using the smallest *p*-value derived in any experiment.

An alternative, and very general, method of modification of Section B.3 is to use a different dose-response curve in the same type of analysis. As an example, a dose-response curve that incorporates any of the known pharmacokinetics of the chemical under test may be substituted. The principles of analysis remain identical, although the implementation of such principles may become difficult. We have not at this point incorporated any other than the multi-stage formula as a dose-response curve.

B.5 Multiple animal experiments

The current approach for using the results of multiple experiments is *ad hoc*. Typically, if there are several experiments (*e.g.* for different sexes or different strains) giving somewhat different estimates for q_1^*, those values will be averaged together in some way (typically by finding a weighted geometric mean) and the result used as the published estimate. Such an approach corresponds to an assumption that the different experiments are measuring the same parameter (*i.e.* that q_1 is the same for the different sexes, strains, or whatever), with experimental and/or other errors, so that some form of averaging is appropriate.

The most direct extension of this to the approach described in Section B.3 is to make the assumption that such different experiments are measuring the same parameter, and the difference in parameter values obtained in the different experiments is simply due to the binomial randomness of the experiments (*i.e.* the size of the experimental groups). Then the loglikelihood functions for the different experiments may simply be added to obtain the loglikelihood function for the combination of experiments.

Alternatively, it may be thought that there is some additional difference between the experiments — they are measuring the same parameter, but there is uncertainty in addition to that built into the analysis of each experiment. Then the correct procedure is to add the loglikelihoods and add also the loglikelihood for the additional uncertainty, to obtain the loglikelihood for the combined experiments.

B.6 Empirical observations of interspecies ratios (animal data only)

Various authors have examined the problem of interspecies comparisons. The following discussion subsumes the analyses of Crouch and Wilson (1979), Crouch (1983), Gaylor and Chen (1986), Chen and Gaylor (1987) and Metzger *et al.* (1989), since all these analyses used the NTP dataset, or the dataset published by Gold *et al.* (1984).

The summary data in the Gold *et al.* (1992) database (kindly supplied in computer readable form by Lois Gold) were analyzed (this incorporates Gold *et al.*, 1984, 1986, 1987, 1990). Equation (B.1) was applied to 4499 experiments on 1136 chemicals, and comparisons between mice, rats and hamsters made. Some results are shown in the following figures, which show scatterplots of q_1^* (the upper 95% confidence limit on q_1 in equation (B.1))

derived from experiments in different species, together with the distribution of the ratios of values of q_1^* for two species.[1] These plots show the empirical distributions plotted on a scale on which samples from a lognormal distribution should be approximately a straight line. This plot was obtained by ordering the samples in increasing size and then plotting (Cunnane, 1978):

$$\Phi^{-\Theta}((i-3/8)/(N+1/4)) \qquad \text{(B.8)}$$

versus the logarithm of the i^{th} sample value, where:

N is the total number samples, and
Φ^{-1} is the inverse cumulative normal function.

From these analyses, it is possible to derive statistics on these interspecies ratios, and to reject the standard interspecies factor equal to the 1/3 power of body weight (similarly for a proposed factor equal to the 1/4 power of body weight) — Figure B.8. Summarizing the results:

- There is a very large variation from chemical to chemical in the ratio of carcinogenic potencies, or potency slopes, between species,
- The distribution of interspecies ratios appears to be approximately lognormal,
- The GSD (uncertainty factor) of this distribution is approximately 8.2 = exp(2.1) for all comparisons, and
- The median (or any other parameter) of the distribution does not agree with any simple allometric scaling law.

[1] All the figures shown have been computed with an extrapolation to a full lifetime corresponding to an $(age)^3$ power law added as a multiplier in the exponent in equation (B.1), rather than the $(age)^2$ power law used by Gold *et al.* (1992). Only experiments that were significant at $p \leq 0.05$ have been included — no conclusions are changed by altering this inclusion criterion between $p \leq 0.05$ and $p \leq 0.001$.

B.7 Empirical observations of interspecies ratios (animal-human data)

Allen, Crump, Shipp *et al.* (Allen *et al.*, 1987; Allen *et al.*, 1988; Crump *et al.*, 1989) examined animal-human comparisons. Figure B.9 and Figure B.10 show results from their "Analysis 0," which is closest to standard method. They used a "Risk-Related-Dose" (RRD) which is equivalent (within a constant factor) to the inverse of q_1*. On Figure B.9, most uncertainty estimates are very large — those shown as reaching the edges, top or bottom of this figure extending to zero or infinity.

Examination of these figures indicates that we can write:

$$\ln(\mathrm{RRD}_{\text{animal}}) = \ln(\mathrm{RRD}_{\text{human}}) + \ln A + \epsilon \qquad \text{(B.9)}$$

where the median interspecies factor is A = 1.2 ×/÷ 1.8 at 1 standard error (and so is not significantly different from unity), and ε is an uncertainty term, with mean zero and standard deviation 2.4, implying an uncertainty factor of 11. The data are too sparse to consider this a good estimate of the standard deviation of the laboratory animal to human distribution, but it is indistinguishable from between-species uncertainties found for laboratory animals. The median value A corresponds to rough equality for RRDs (and hence q_1*) between humans and animals with dosing on a per-unit-body weight basis.

B.8 Empirically based distribution for human potency

The results of Sections B.6 and B.7 may now be combined to give an empirical lognormal distribution for animal to human extrapolation. The data shown in section B.7 indicate that the least biased estimate for the median of this extrapolation is approximately 1.2, with an uncertainty of a factor of 1.8, when doses are measured on a per-unit-body weight basis (and Section B.6 indicates that there is no simple allometric scaling law that can improve on this estimate). The data shown in Section B.6 indicate that there is also an additional uncertainty of a factor of exp(2.1) = 8.2, while the data of Section B.7 agree with this. Combining the uncertainty of the median with the remaining uncertainty leads to a total uncertainty of approximately a factor of exp(2.2).

Thus, to obtain a distribution for human potency requires first an estimate of carcinogenicity in test animals, using the methods of sections B.3 through B.5, followed by a convolution with the empirical interspecies distribution, which is lognormal with median of approximately 1.0 (the best estimate is 1.2, but we suggest the use of 1.0, since this is well within the error estimate) and has a geometric standard deviation of a factor of exp(2.2) = 9.0. In general, the uncertainty in the interspecies extrapolation greatly dominates the uncertainty obtained in sections B.3 through B.5 (which is essentially simply a measurement error); except that section B.5 indicates how to add in additional information that may result in an appreciable probability for a human potency of zero.

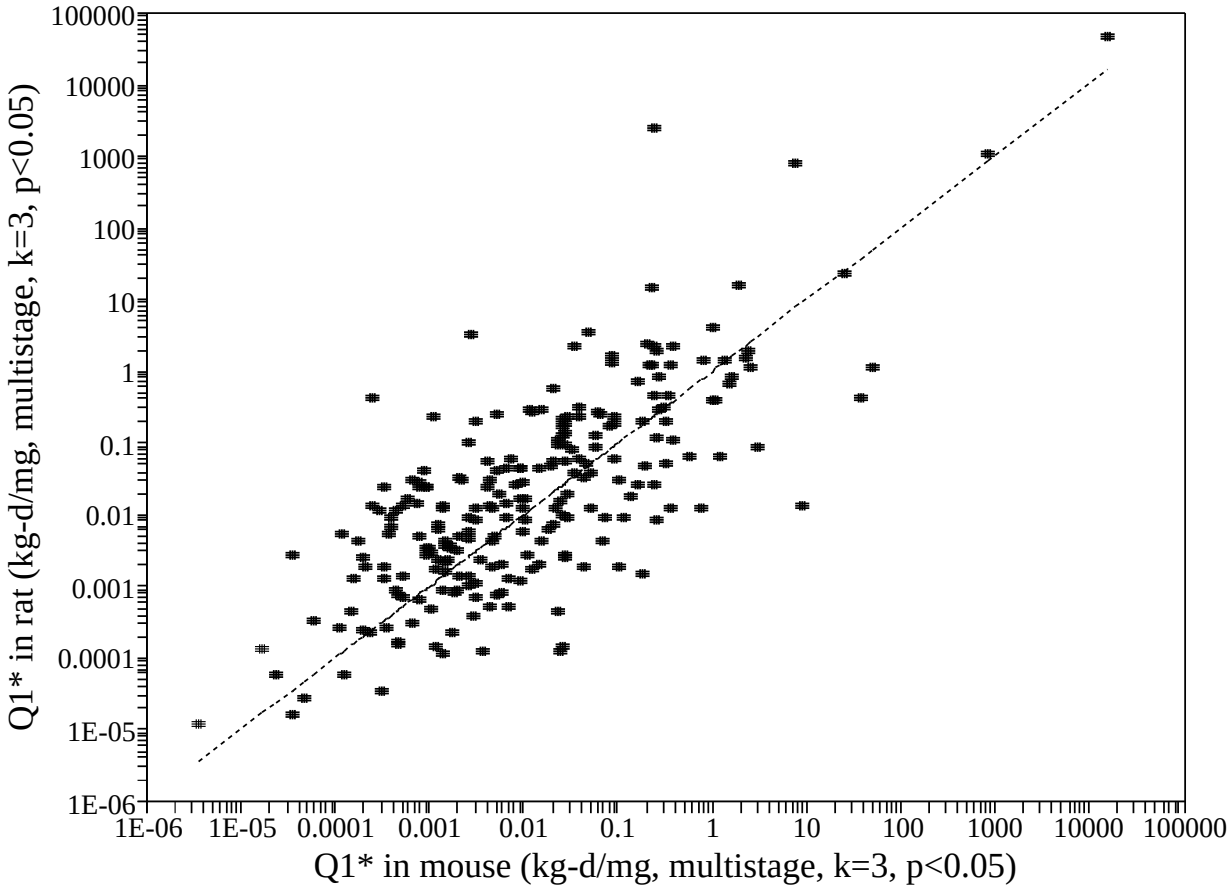

Figure B.2 Scatterplot of q_1* in mice and rats

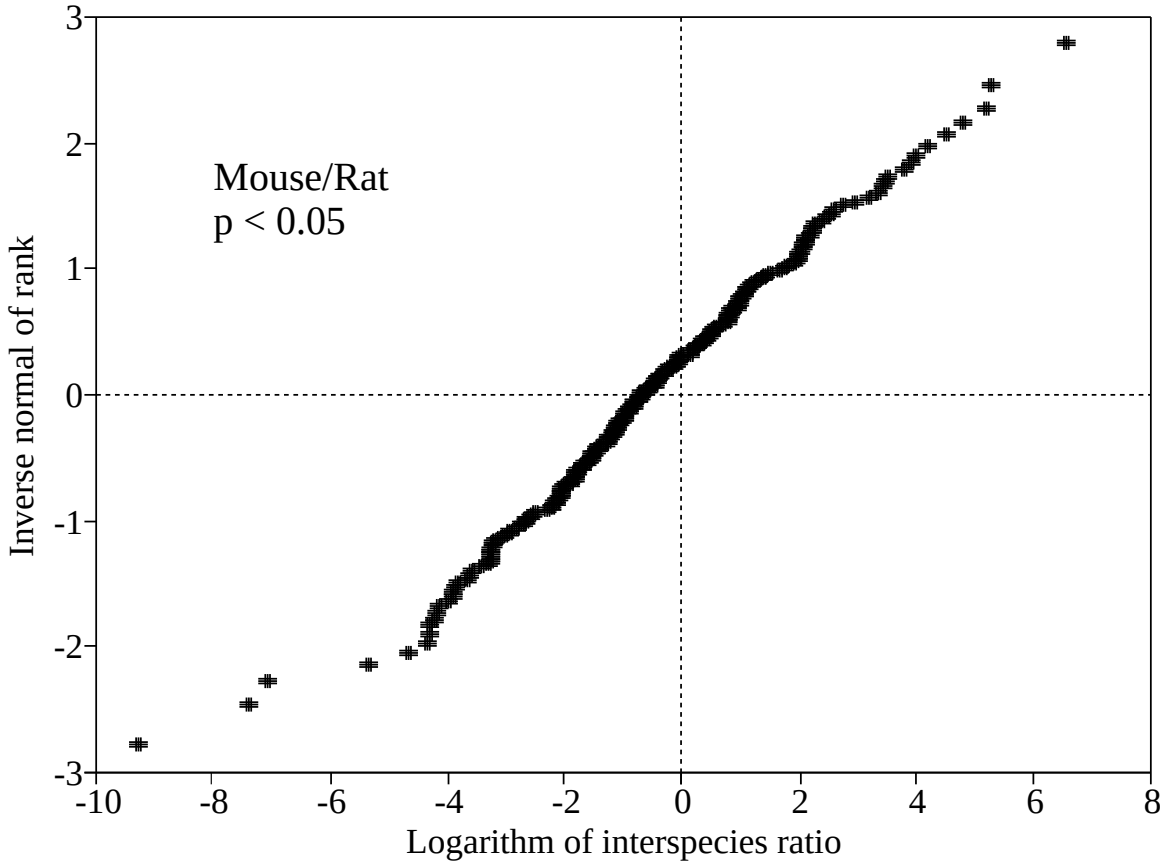

Figure B.3 Interspecies ratios of q_1* (Mouse/Rat)

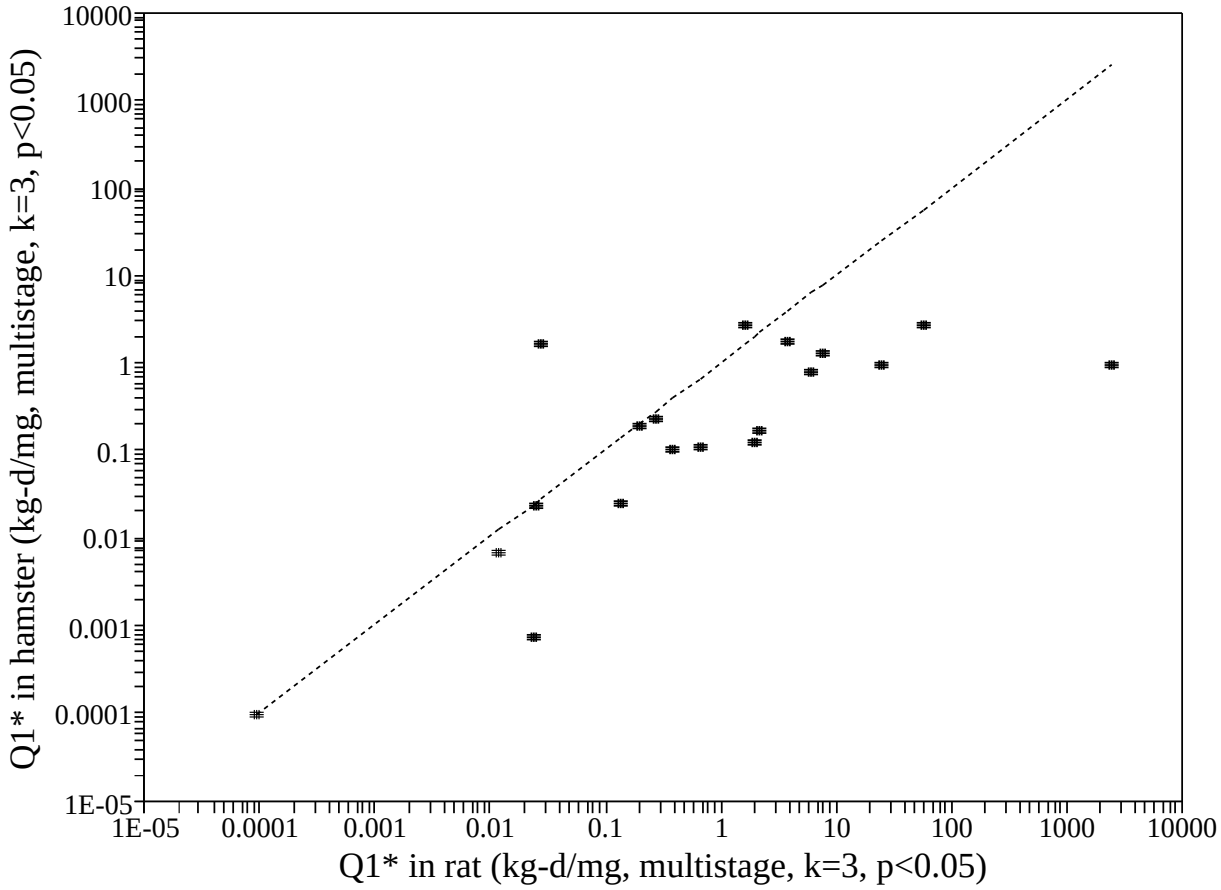

Figure B.4 Scatterplot of q_1* in rats and hamsters

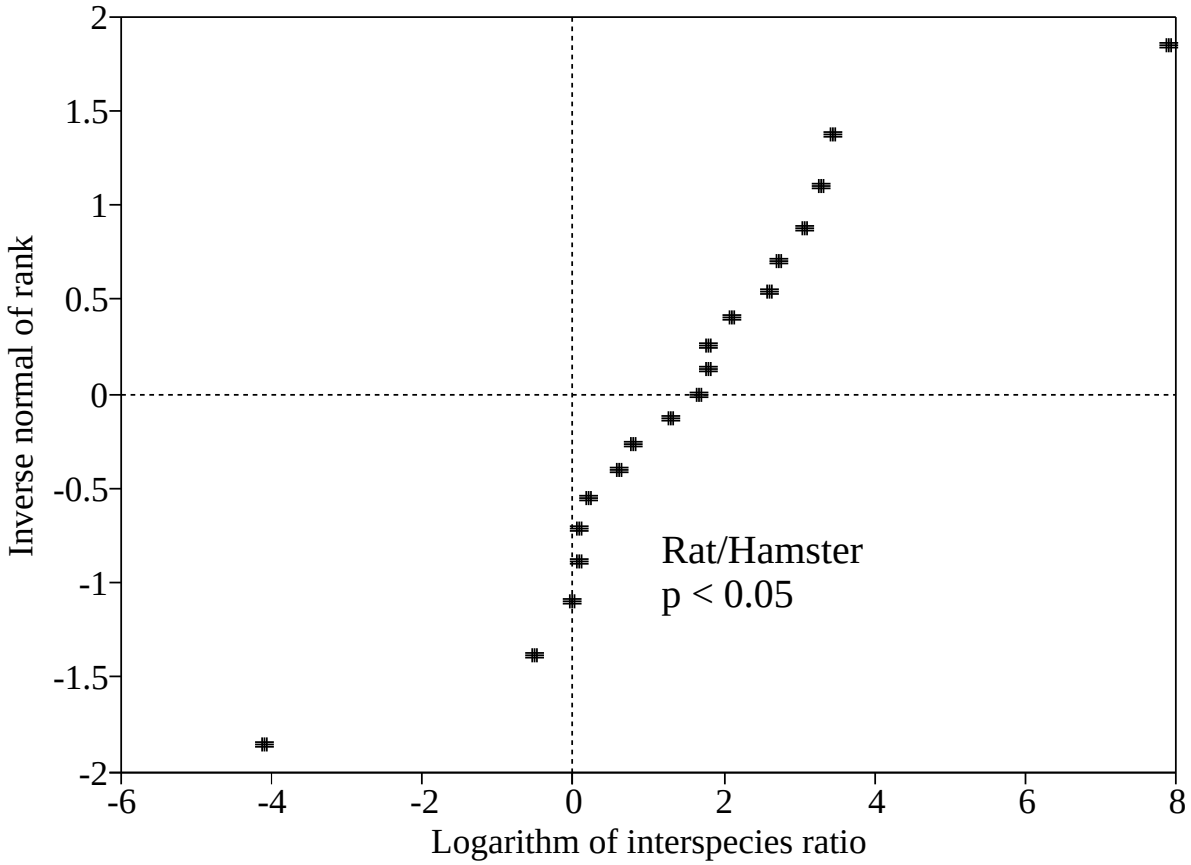

Figure B.5 Interspecies ratios of q_1* (Rat/Hamster)

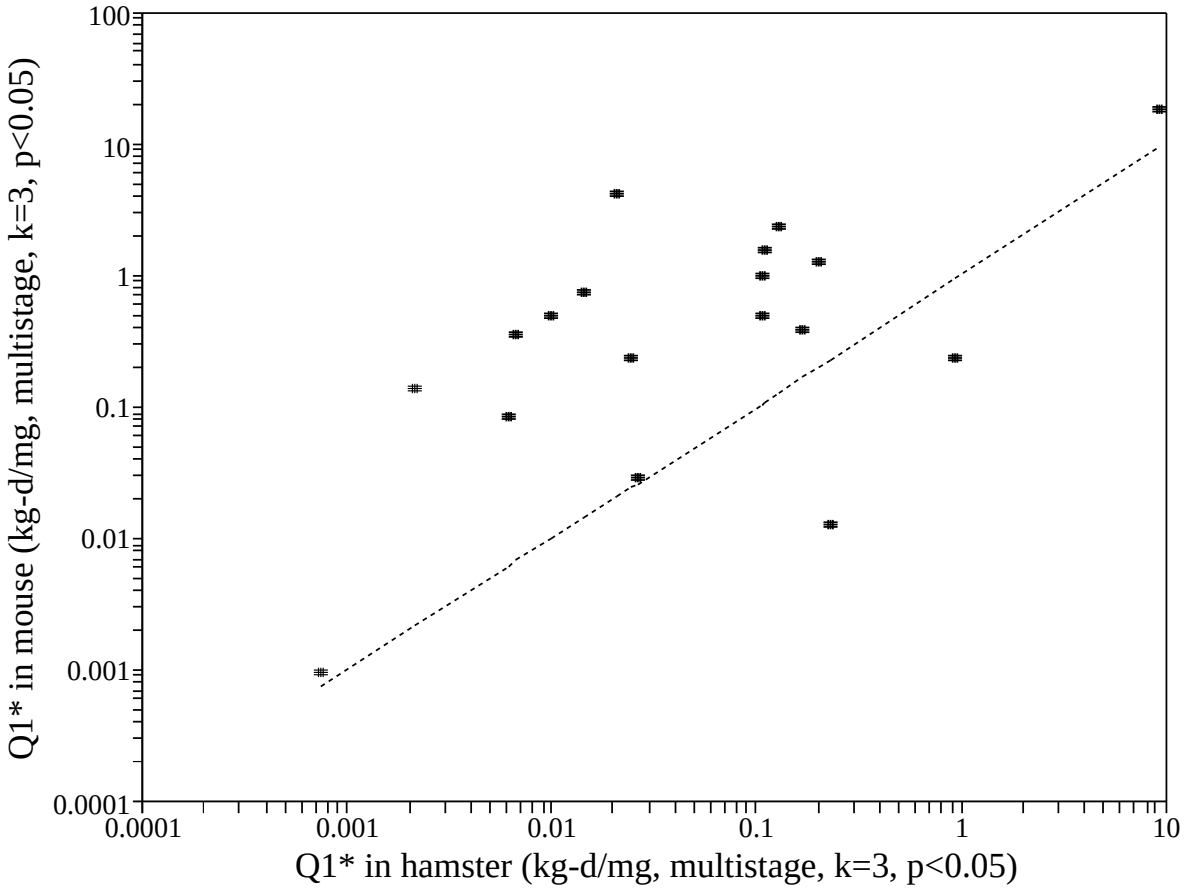

Figure B.6 Scatterplot of q_1* for hamsters and mice

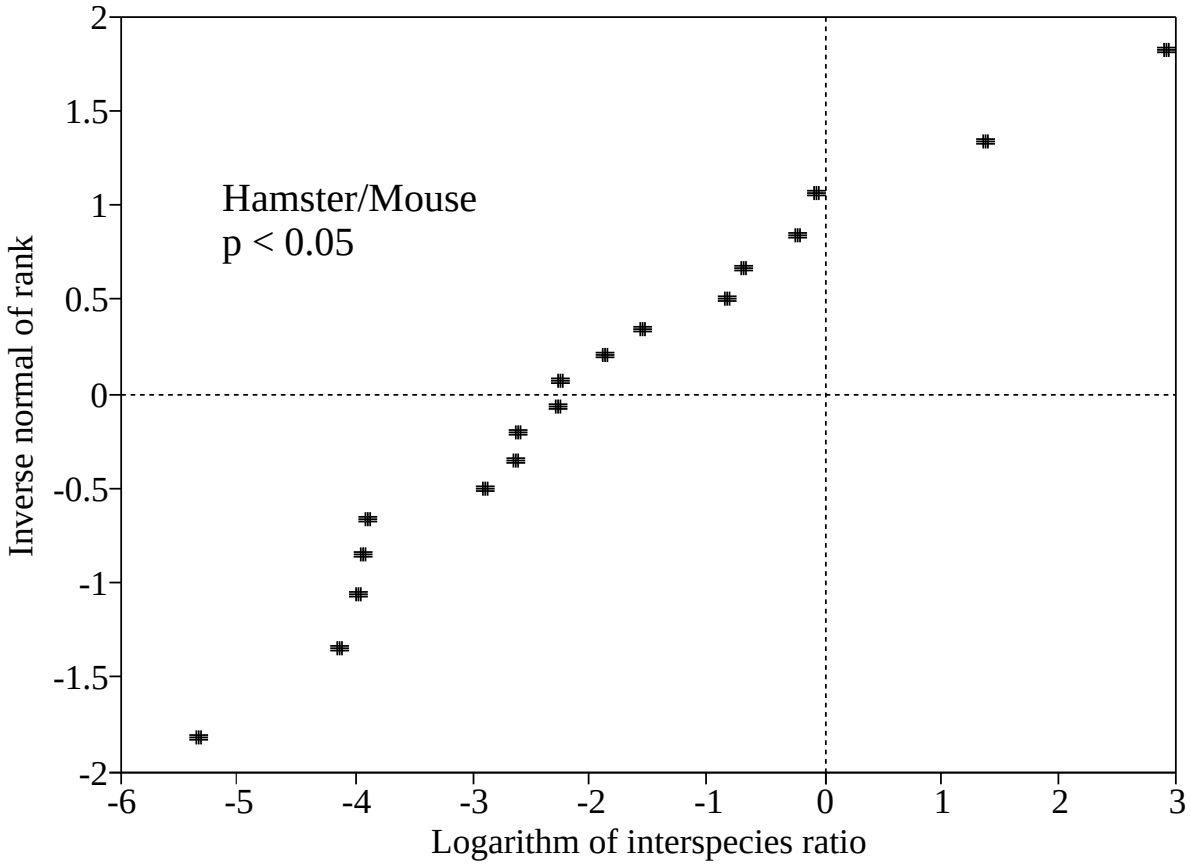

Figure B.7 Interspecies ratios of q_1* (Hamster/Mouse)

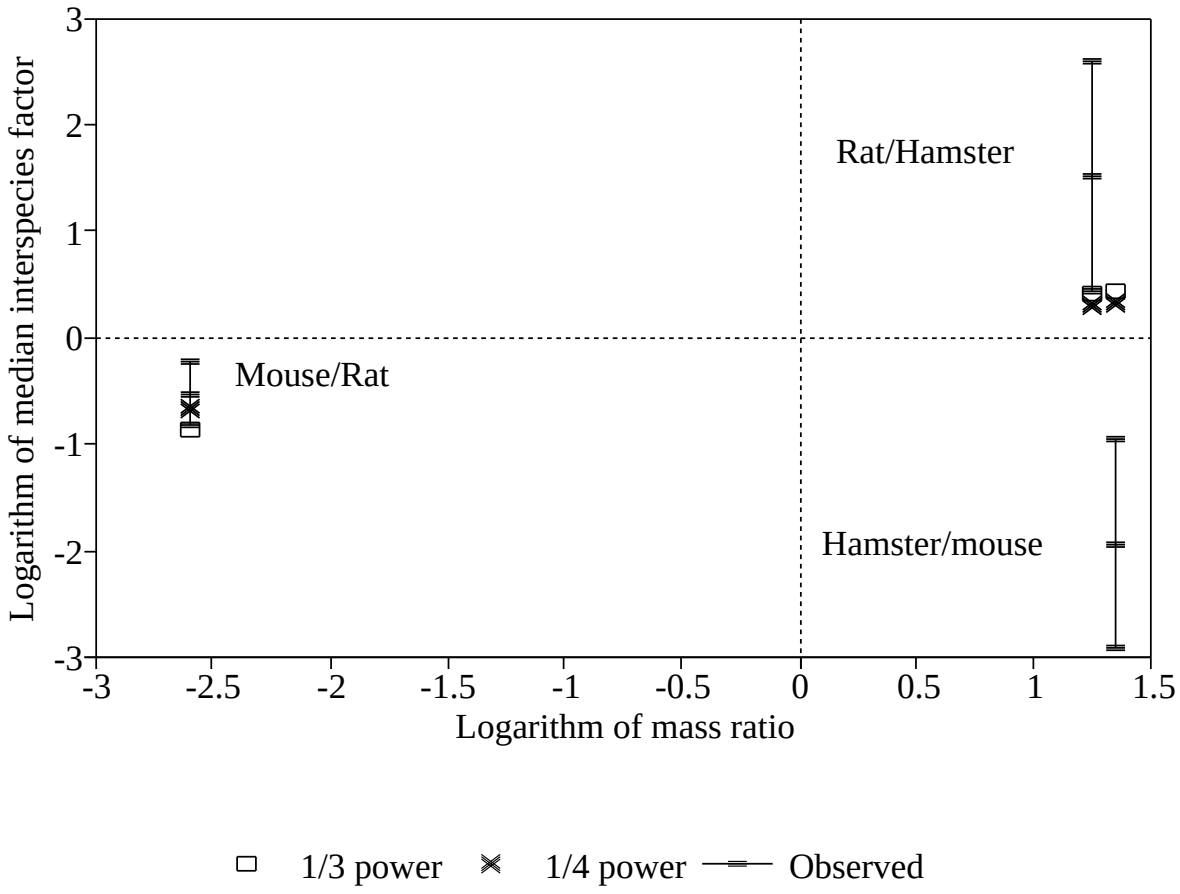

Figure B.8 Empirical estimates of geometric mean interspecies factors for q_1* (Error bars are ±2 SE)

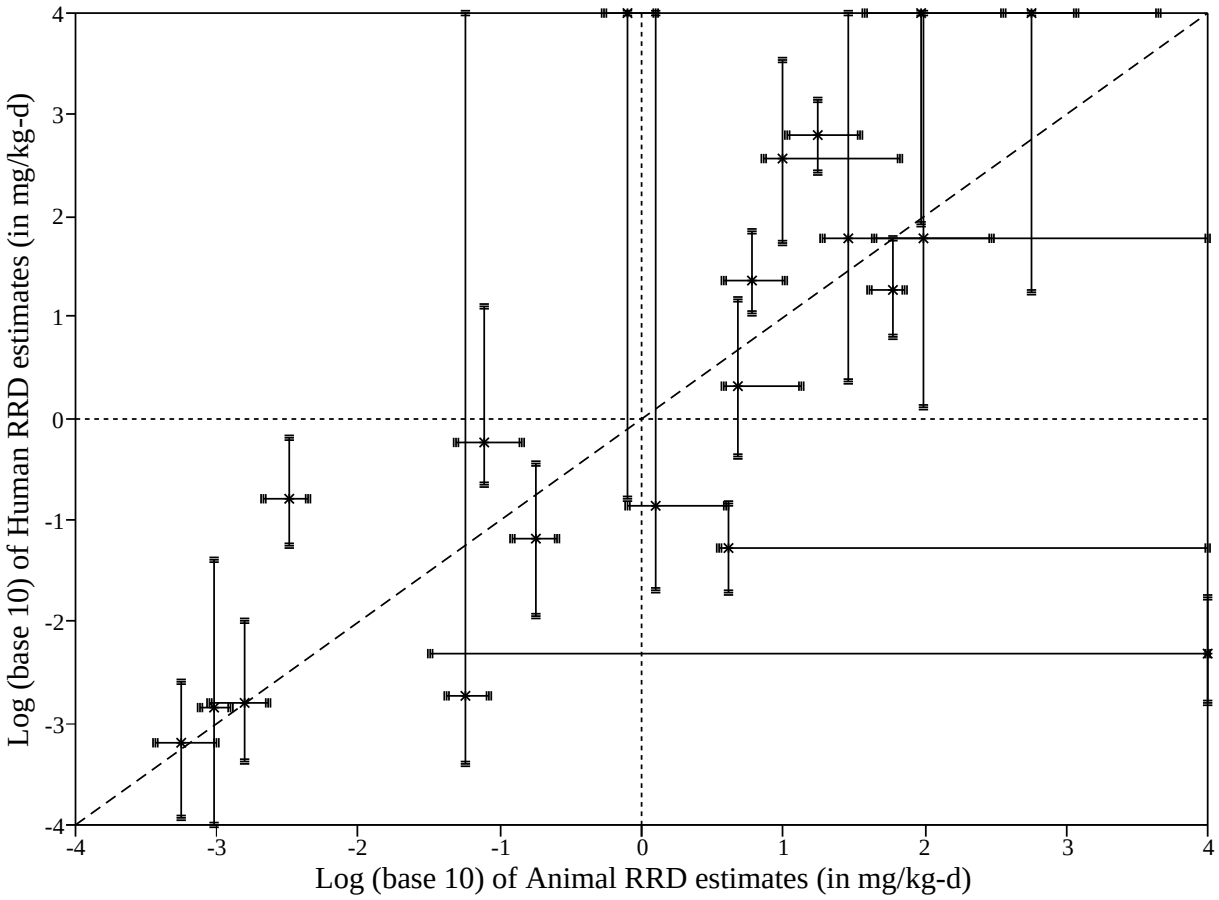

Figure B.9 Allen, Crump, Shipp, *et al.* base case

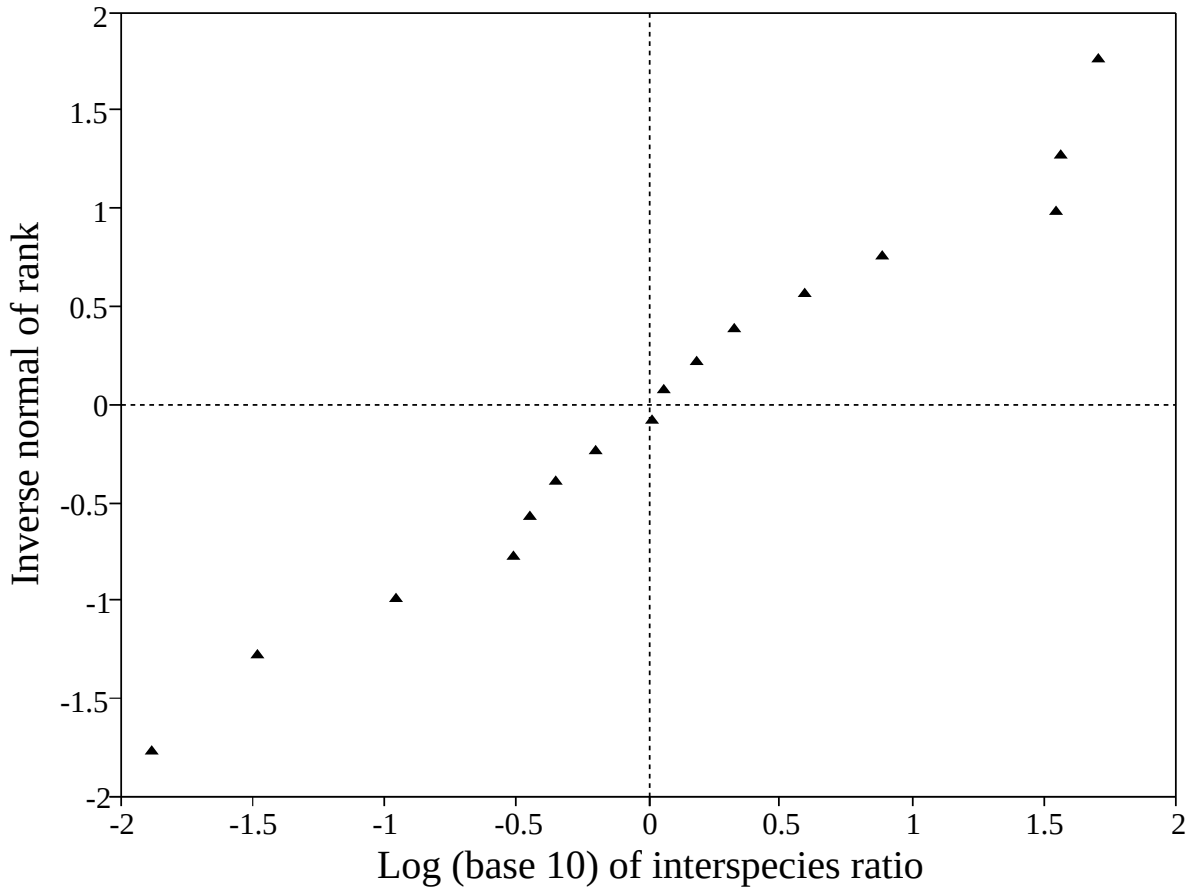

Figure B.10 Distribution of uncertainty component of animal to human comparison

B.9 References for Appendix B

Allen, B.C., A.M. Shipp, K.S. Crump, B. Kilian, M.L. Hogg, J. Tudor, and B. Keller, 1987. Investigation of Cancer Risk Assessment Methods (3 vols. plus summary report). Clement Associates Inc. (January, 1987) and OHER, ORS, EPA/600/6-87/007 (September, 1987).

Allen, B.C., K.S. Crump, and A.M. Shipp, 1988. Correlation between carcinogenic potency of chemicals in animals and humans. Risk Analysis 8:531–544.

Anderson, E.L., and the Carcinogen Assessment Group of the U.S. Environmental Protection Agency, 1983. Quantitative approaches in use to assess cancer risk. Risk Analysis 3:277–295.

Chen, J.J., and D.W. Gaylor, 1987. Carcinogenic Risk Assessment: Comparison of estimated safe doses for rats and mice. Environ. Health Perspect. 72:305–309.

Crouch, E., and R. Wilson, 1979. Interspecies comparison of carcinogenic potency. J. Toxicol. Environ. Health 5:1095–1118.

Crouch, E.A.C., 1983. Uncertainties in interspecies extrapolations of carcinogenicity. Environ. Health Perspect. 50:321–327.

Crouch, E.A.C., 1992. MSTAGE 2.01 — a program to fit end-of-life carcinogenesis bioassay data to the multistage formula. Available on request from the author.

Crump, K.S., Howe, R.B., Van Landingham, C., 1987. TOX_RISK — Toxicology Risk Assessment Program. Available from Clement Associates, K.S. Crump Division, 1201 Gaines Street, Ruston, LA 71270.

Crump, K., B. Allen, and A. Shipp, 1989. Choice of dose measure for extrapolating carcinogenic risk from animals to humans: an empirical investigation of 23 chemicals. Health Physics 57 (Sup 1):287–393.

Cunnane, C., 1978. Unbiased plotting positions — a review. J. Hydrology 37:205–222

EPA, 1983. Health assessment document for acrylonitrile. Office of Health and Environmental Assessment, Washington, DC 20460. EPA-600/8-82-007F. Final Report, October 1983.

Gaylor, D.W., and J.J. Chen, 1986. Relative potency of chemical carcinogens in rodents. Risk Analysis 6:382–290.

Gold, L.S., C.B. Sawyer, R. Magaw, G.M. Backman, M. de Veciana, R. Levinson, N.K. Hooper, W.R. Havender, L. Bernstein, R. Peto, M.C. Pike, and B.N. Ames, 1984. A Carcinogenic Potency Database of the standardized results of animal bioassays. Environ. Health Perspect. 58:9–319.

Gold, L.S., M. de Veciana, G.M. Backman, R. Magaw, P. Lopipero, M. Smith, M. Blumenthal, R. Levinson, L. Bernstein, and B.N. Ames, 1986. Chronological supplement to the Carcinogenic Potency Database: standardized results of animal bioassays published through December 1982. Environ. Health Perspect. 67:161–200.

Gold, L.S., T.H. Slone, G.M. Backman, R. Magaw, M. DaCosta, P. Lopipero, M. Blumenthal, R. Levinson, L. Bernstein, and B.N. Ames, 1987. Second chronological supplement to the Carcinogenic Potency Database: standardized results of animal bioassays published through December 1984 and by the National Toxicology Program through May 1986. Environ. Health Perspect. 74:237–329.

Gold, L.S., T.H. Slone, G.M. Backman, S. Eisenberg, M. DaCosta, M. Wong, N.B. Manley, L. Rohrbach, and B.N. Ames, 1990. Third chronological supplement to the Carcinogenic Potency Database: standardized results of animal bioassays published through December 1986 and by the National Toxicology Program through June 1987. Environ. Health Perspect. 84:215–285.

Gold, L.S., N.B. Manley, T.H. Slone, G.B. Garfinkel, L. Rohrbach, and B.N. Ames, 1992. The fifth plot of the Carcinogenic Potency Database: Results of animal bioassays published in the general literature through 1988 and by the National Toxicology Program through 1989. Environ. Health Perspect. 100.

Metzger, B., E. Crouch, and R. Wilson, 1989. On the relationship between carcinogenicity and acute toxicity. Risk Analysis 9:169–177.