EPA 570/9-85-001
NTIS Assession No. PB84 213941

TECHNIQUES FOR THE ASSESSMENT OF THE CARCINOGENIC RISK TO THE U. S. POPULATION DUE TO EXPOSURE FROM SELECTED VOLATILE ORGANIC COMPOUNDS FROM DRINKING WATER VIA THE INGESTION, INHALATION AND DERMAL ROUTES*

C. Richard Cothern, William A. Coniglio and William L. Marcus

William L. Lappenbusch
Chief, Health Effects Branch

Joseph A. Cotruvo
Director, Criteria and Standards Division

OFFICE OF DRINKING WATER (WH-550)
U. S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

Estimates of Population Risk Or Cancer Cases Expected From Existing Concentrations in Public Drinking Water Supplies, Estimates of Cancer Cases Averted And Number of Systems Affected By Drinking Water Regulation and Estimates Of The Uncertainty in These Cases

Benzene
Carbon Tetrachloride
1,2-Dichloroethane
1,1-Dichloroethylene
Perchloroethylene
Trichloroethylene
Vinyl Chloride

JULY 25, 1984

* This report has been reviewed by the Office of Drinking Water, U. S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U. S. Environmental Protection Agency.

EXECUTIVE SUMMARY

The risk of individuals experiencing an adverse health effect after being exposed to a contaminant may be calculated by multiplying the level of personal exposure times a measure of the hazard resulting from the toxic potency of the contaminant. The total number of cases of disease or death expected among an entire society for one lifetime of its members is called the population risk. For this report the present exposure level for drinking water contaminants among the U. S. population is calculated and expressed as a population-concentration figure. Population concentration is the sum of the products of the number of people times their estimated exposure level. Typical units of population concentration found in drinking water contaminants are persons x micrograms/liter. The carcinogenic potency for individual contaminants has been expressed as the number of cases of cancer per person per microgram per liter. This report considers only the carcinogenic risk for specific contaminants and does not attempt to deal with other manifestations of toxicity.

This report develops potency estimates and population risk estimates for nine selected volatile organic contaminants (VOC's) in drinking water. These nine contaminants were selected for study because they offered a range of data of potency and population exposure via drinking water. The compounds include: benzene, carbon tetrachloride, 1,1-dichloro-

ethane, 1,1-dichloroethylene, perchloroethylene, trichloroethylene, and vinyl chloride. The analysis provides a basis for determining the direct effectiveness of national drinking water standards to reduce illness and death from present levels of these chemicals in water supplies. This portion of the analysis could be called risk reduction analysis. It does not consider the benefits the national regulation may have in preventing further deterioration of water supplies and human exposure.

The occurrence estimates were evolved from monitoring data developed in several national drinking water surveys. The surveys included are: the National Organic Reconnaissance Survey the National Organic Monitoring Survey, the National Screening Program for Organics in Drinking Water, the Community Water Supply Survey, the Rural Water Survey and the Groundwater Supply Survey. Taken together, these surveys provided data on over 1,000 groundwater supplies and 1,000 surface water supplies. This data base was drawn from 48,458 groundwater supplies and 11,202 surface water supplies and is shown in Table EX-1.

Drinking water contributes to the three routes of exposure, viz., ingestion, inhalation and dermal. An estimate of the relative importance of these routes for different age groups for VOC's in drinking water is shown in Table EX-2. Due to the chemical characteristics of the VOC's, dermal

TABLE EX-1

TOTAL ESTIMATED POPULATION (IN THOUSANDS) EXPOSED TO THE CONTAMINANT IN DRINKING WATER AT THE INDICATED CONCENTRATION RANGE MICROGRAMS/LITER

Contaminant	Population Concentration (see footnote*) Mean Best Total	Mean Worst Total	Drinking Water Concentration (microgram/liter) <0.5	>0.5-5	>5-10	>10-20	>20-30	>30-40	>40-50	>50-60	>60-70	>70-80	Other
Benzene	1.6×10^7	1.2×10^8	209,166	4,535	123	62	16	6	3	2	1	1	<100
Carbon Tetrachloride	9.5×10^7	1.9×10^8	187,609	24,460	1,389	43	655	0	0	0	0	0	0
1,2-Dichloriethane	3.8×10^7	1.4×10^8	202,187	12,890	0	143	0	0	0	0	0	0	0
1,1-Dichloroethylene	6.7×10^6	2.2×10^8	209,630	2,295	52	0	0	0	0	0	0	0	0
Tetrachloroethylene	4.5×10^7	1.5×10^8	202,989	10,556	434	254	82	0	0	0	105	0	0 >100
Trichloroethylene	1.3×10^8	2.2×10^8	189,288	23,287	426	212	0	738	257	42	0	128	42
Vinyl Chloride	1.5×10^7	1.2×10^8	212,497	1,331	472	0	0	0	0	0	118	0	0

* The population concentration is the sum of the products of the population and the average drinknig water concentration to which that population was exposed. The mean best total assumes that all in the lowest concentration were exposed to 0.0 micrograms/liter, and the mean worst assumes that all were exposed to the maximum or 0.5 micrograms/liter.

TABLE EX-2

COMPARATIVE MODEL OF ABSORBED DOSE FROM
A VOLATILE POLLUTANT (100 MICROGRAMS/LITER) IN DRINKING WATER

	Absorbed Dose (micrograms/day)				
	Formula-fed Infant (4 kg)	Pre-[illegible] (32 kg)	Adult Female (60 kg)	Adult Male[a] (70 kg)	Male[b] (70 kg)
Fluid Ingestion	80	150	200	200	200
Inhalation of enriched indoor air	10	10	70	50	50
Inhalation of enclosed shower air	--	100	300	200	--
Dermal absorption					
Bathing	.02-.06	--	--	--	2.5
Swimming	--	10-300	--	--	--
Total Absorbed Dose micrograms/kg/day	20 micrograms/	10-20 micrograms/	10 micrograms/	7 micrograms/	4 micrograms/

a Showering adult male.
b Adult male that does not take showers.

exposure appears to be very small compared to the ingestion and inhalation exposure potentials, which may be close in magnitude. Respiration exposure is of increased importance due to the enrichment of indoor air from water use.

The carcinogenic potency of these contaminants was estimated from bioassay data developed from animal tests conducted at high dose levels. The biological response at lower levels of exposure was estimated using mathematical models. Figure EX-1 shows a typical extrapolation. The initial starred points are the animal data, and subsequent lines represent the projections from four different models. The error bars on each line represent the upper 95% confidence calculations from the model.

Attempts were made to deal with the factors contributing to uncertainty in these risk estimates, which can be very large. The authors' estimates of contribution to the uncertainty from generic toxicity factors are shown in Table EX-3. Likewise, the contribution to uncertainty from the exposure estimates is shown in Table EX-4. Using this method for determining risk, it is evident that the largest contribution to the uncertainty is from the choice of dose extrapolation model. As can be seen in Figure EX-1, this uncertainty for the region of interest can be five to six orders of magnitude or more. This uncertainty is much larger than all other uncertainties combined and therefore has been used to determine the

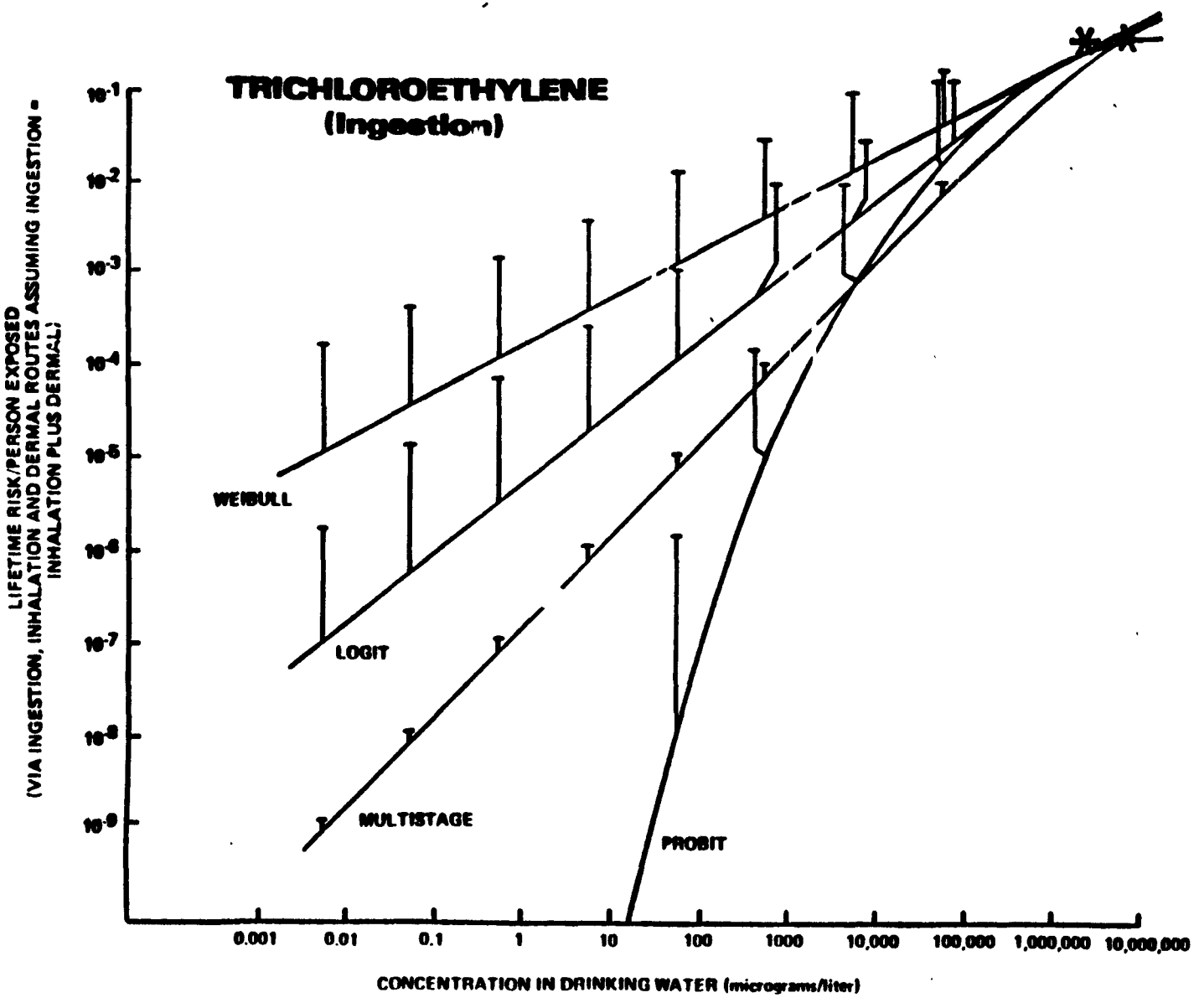

FIGURE FY-1

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multi-stage, probit and Weibull models from those bioassay data for trichloroethane. The bioassay data was obtained by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhala-

TABLE EX-3

HEALTH EFFECTS UNCERTAINTY

	Category	Contribution to Uncertainty*
1.	Choice of endpoint	Less than 10% of the time chose the wrong endpoint--some exception such as chlorinated hydrocarbon, where expect liver cancer (U)
2.	Personnel capabilities	All or nothing (E)
3.	Choice of species, strain, age and sex of animals	All or nothing (E)
4.	Test compound purity (contamination, decay and vehicle contribution)	For the VOC's is generally trivial but could be two orders of magnitude (O)
5.	Inappropriate statistical test methodology	All or nothing (E)
6.	Distribution of animals among doses and number used	One to two orders of magnitude (E)
7.	Selection of dose levels	Up to two orders of magnitude (E)
8.	Lack preliminary tumor change information such as hyperplasia	For VOC's is less than 10% but could for other contaminants be as large as a factor of three (U)
9.	Experimental surroundings	Could be orders of magnitude (O)
10.	Dietary considerations	Factor of two (E)

* U-leads to an underestimate of the risk.
O-leads to an overestimate of the risk.
E-could lead to an overestimate or an underestimate of the risk.

TABLE EX-3 (continued)

11.	GLP's (good laboratory procedures)	One to two orders of magnitude (U)
12.	Time-to-tumor	Could miss the effect or could be an underestimate by a factor of two
13.	Microscopic tissue examination	Could be a factor of two (E)
14.	Diseases in test animals	All or nothing (E)
15.	Statistical noise	Factor of 2 (E)
16.	Outcomes of cancer (if include non-fatal cancers)	Could be two orders of magnitude or more (O)
17.	Conventional choice of p level (e.g., 0.05)	All or nothing (E)
18.	Synergism/Antagonism	Many orders of magnitude (E)
19.	No corresponding tissue in humans	All or nothing (E)
20.	Most sensitive vs. average	Several orders of magnitude (E)
21.	Animal to man	Conservatively two orders of magnitude (E)
22.	Body weight vs. surface	One order of magnitude (E)
23.	Use of upper 95% confidence limit	Up to an order of magnitude (O)
24.	Choice of dose-response model	5-6 orders of magnitude when considering risk levels in the 10^{-4} to 10^{-6}/lifetime range (E)

* U-leads to an underestimate of the risk.
O-leads to an overestimate of the risk.
E-could lead to an overestimate or an underestimate of the risk.

TABLE EX-4

SOURCES OF UNCERTAINTY FOR OCCURRENCE, POPULATION CONCENTRATION AND EXPOSURE ESTIMATES USED IN THE ASSESSMENT OF VOLATILE ORGANIC CHEMICALS IN DRINKING WATER

Factors	Impact on estimate of†		
	Occurrence	Population Concentration	Risk
Generation of monitoring data			
Proportion of population sampled	5% (U)		
Representativeness of systems selected			
Geographic distribution, system size and source of water	10% (E)		
Sampling methods			
Site of sample collection	20% (E)		Factor
Time of sample collection	20% (E)	50%	of
Method of sample collection	10% (U)		2
Container type	10% (U)		
Stability during storage	100% (U)		
Sample analysis			
% recovery from sample	.0% (U)		
Compound identification	10% (E)		
Accuracy of quantitative determination	40% (E)		
Assumptions during data analysis			
Lower limits of quantification		Factor of 3-4	Factor of 2 (E)
Oral exposure rates			
Intake rate of water			10% (E)
Pollutant level in consumed water (hot vs. cold)			50% (O)
% absorption for oral intake			10% (O)
Respiratory exposure rates			Factor of 3
Dermal exposure			Negligible

†U = leads to an underestimation of the risk.
O = leads to an overestimation of the risk.
E = could lead to an overestimation or an underestimation of the risk.

range of the population risk estimates.

Table EX-5 is a typical example of the calculation of population risk and the risk reduction analysis. As seen in Table EX-5, the population risk is in the range of less than one to 400 per lifetime. (Fractional population risks resulting from calculations have been rounded off to the designation of less than one). In most cases, the largest contribution to population risk is from the population exposed below the detection limit. As this is an estimate, it may be an artifact of the analysis. A summary of the risk reduction analysis estimates is shown in Table EX-6 for the nine selected VOC's.

The ranges of uncertainty in population risk estimates and risk reduction analysis are quite large. If the actual shape of the dose-response curve were known, the uncertainty in these estimates could be reduced to about two to three orders of magnitude. Data for human exposure, especially at the lowest exposure category, could improve the uncertainty, and data at lower dose levels could also reduce the uncertainty. Neither of these sources of uncertainty can be influenced easily due to the extreme cost involved and because of the unethical nature of human experimentation.

TABLE EX-5

BENZENE-POPULATION RISK ESTIMATES
FOR CURRENT LEVELS OF DRINKING WATER EXPOSURE

Mean Drinking Water Concentration (Micrograms/ Liter)	Number of People Being Served	Total Lifetime Individual Risk For the Mean Concentration*		Lifetime Population Risk
		Low (Probit)	High (Weibull)	
0.25	2.1×10^{8}	$<10^{-10}$	1.7×10^{-6}	<1 - 360
2.75	4.5×10^{6}	$<10^{-10}$	1.3×10^{-5}	<1 - 60
7.5	1.2×10^{5}	$<10^{-10}$	3.1×10^{-5}	<1 - 4
15	6.2×10^{4}	2×10^{-9}	5.1×10^{-5}	<1 - 3
25	1.6×10^{4}	1×10^{-8}	8.9×10^{-5}	<1 - 1
35	6×10^{3}	3×10^{-8}	1.3×10^{-4}	<1 - 1
45	3×10^{3}	4×10^{-8}	1.6×10^{-4}	<1 - 1
55	2×10^{3}	5×10^{-8}	1.7×10^{-4}	<1 - 1
65	1×10^{3}	7×10^{-8}	2.2×10^{-4}	<1 - 1
75	1×10^{3}	1×10^{-7}	2.4×10^{-4}	<1 - 1
100	1×10^{3}	5×10^{-7}	3.2×10^{-4}	<1 - 1
			Total†	400

Benzene-Risk Reduction Analysis
For Limiting Drinking Water Concentration

Maximum Allowable Drinking Water Concentration (Micrograms/Liter)	Approximate Individual Risk Rate for Maximum Concentration	Cummulative Cases Averted°
75	1×10^{-8} - 2×10^{-4}	<1 - 2
35	7×10^{-9} - 1×10^{-4}	<1 - 6
7.5	$<1 \times 10^{-10}$ - 8×10^{-5}	<1 - 10
2.75	$<1 \times 10^{-10}$ - 7×10^{-5}	<1 - 70

† Rounded to one significant figure.

* The total individual risk was determined by assuming that the risk due to inhalation exposure is equal to that due to ingestion.

° Number of cases averted for concentrations shown in the first column.

TABLE EX-6

CASES AVERTED IF A STANDARD IS SET AT THE CONCENTRATION SHOWN

Contaminant	Population Risk†	Micrograms/ Liter	Cases Averted†	Micrograms/ Liter	Cases Averted†	Micrograms/ Liter	Cases Averted†
Benzene	<1 - 400	2.75	<1 - 70	7.5	<1 - 10	75	<1 - 2
Carbon Tetrachloride	DATA	INSUFFICIENT		FOR	ANALYSIS		
1,2-Dichloroethane*	40 - 90,000	2.75	30 - 20,000	7.5	3 - 400	15	3 - 400
1,1-Dichloroethylene	DATA	INSUFFICIENT		FOR	ANALYSIS		
Perchloroethylene	<1 - 10	2.75	<1 - 6	7.5	<1 - 3	65	<1 - 2
Trichloroethylene*	<1 - 70,000	2.75	<1 - 20,000	7.5	<1 - 4,000	100	<1 - 200
Vinyl Chloride*							
(Maltoni-Old)	1 - 50,000	2.75	<1 - 3,000	7.5	<1 - 1,000	65	<1 - 400
(Feron-New)	<1 - 4,000	2.75	<1 - 3,000	7.5	<1 - 2,000	65	<1 - 1,000

† Rounded off to one significant figure.

* See text.

TABLE OF CONTENTS

TABLES

TABLES (continued)

FIGURES

FIGURES (continued)

INTRODUCTION

For over 400 years, we have sought to measure the effects that contaminants, including chemicals, have on living tissue. We have confirmed and reconfirmed one basic truth: all contaminants can cause cellular damage that may be expressed as a loss of organ function or death if the contaminant is present in living tissue at a concentration that is sufficiently high or for a long enough period of time to interfere with normal biochemical function. The amount of a contaminant entering via all specific routes of absorption--the skin, the lung, the intestinal tract--must be considered as exposure. The ability of the contaminant to enter into biochemical reactions essential for life processes is a prime factor in determining if the living organism will exhibit stress, altered behavior, loss of cell or tissue function that is diagnosable as disease or other detrimental effects including death.

Water is the major constituent of living cells. Throughout all recorded history, man has endeavored to assure a constant and convenient supply of water not only to sustain his life, but to enhance social concepts of personal hygiene and provide recreation. A convenient and readily available supply of water has been the driving force behind engineering advances that constructed cisterns, catchments, aquaducts, indoor plumbing, bathtubs, swimming pools, toilets, washing machines, ice makers, showers, hot whirlpools, dishwashers and humidi-

tiers. Hot water and cold water are brought directly to the home for infants and aged, healthy and ailing alike to use and enjoy.

How each hitchhiker in drinking water affects the individuals in it depends, inter alia, upon many factors including: (1) how much of the chemical is absorbed per unit of body weight; (2) how active the contaminant is within the cell and (3) how susceptible the individual is to the chemical action as determined by initial genetics, nutrition and sums of other stresses including chemicals or prior diseases.

Some contaminants in drinking water are cancer-causing agents. Examples of such contaminant include the Volatile Organic Chemicals (VOC's) which are analyzed in succeeding chapters. The particular VOC's being considered here were selected because they appear to be the most important VOC's presently known to be in the national water supply and the dose-response curves for these and other carcinogens do not have a threshold for toxicity (Schneiderman, 1979). To develop drinking water regulatory standards for these contaminants it is necessary to estimate the risk posed by low concentrations in drinking water. This estimation requires extrapolation of dose-response curve from the experimental data at high doses into the unknown region of low concentration. Coupling these estimates with estimates of the number of people exposed, the number of cases of cancer that can be prevented by setting

standards at different levels can be factored into the regulatory process.

Risk estimation for drinking water contaminants requires the knowledge of the level present in the water and the amount of individual exposure resulting from a lifetime behavior pattern. We must look upon the frequency at which various concentrations of the contaminant presently occur in water supplies across the United States, the number of people being served by those water supplies, the resultant exposure to these people and a carcinogenic potency of this substance. Figure 1 shows how this information is used in the risk estimation process.

An estimate of total expected excess cancer in a lifetime for the U. S. from a given contaminant is called the population risk estimate. This is an estimate and not a statement of perfect knowledge. Thus an uncertainty or range of values will be included to describe the estimate. If the frequency distribution of the contributing factors controlling exposure and toxicity were known, the propagation of errors can be determined using standard and well-known methods. However, in the present situation the frequency distributions are not known and thus a range of values will be used to roughly describe the situation.

The factors that are involved in the determination of population risk from drinking water are:

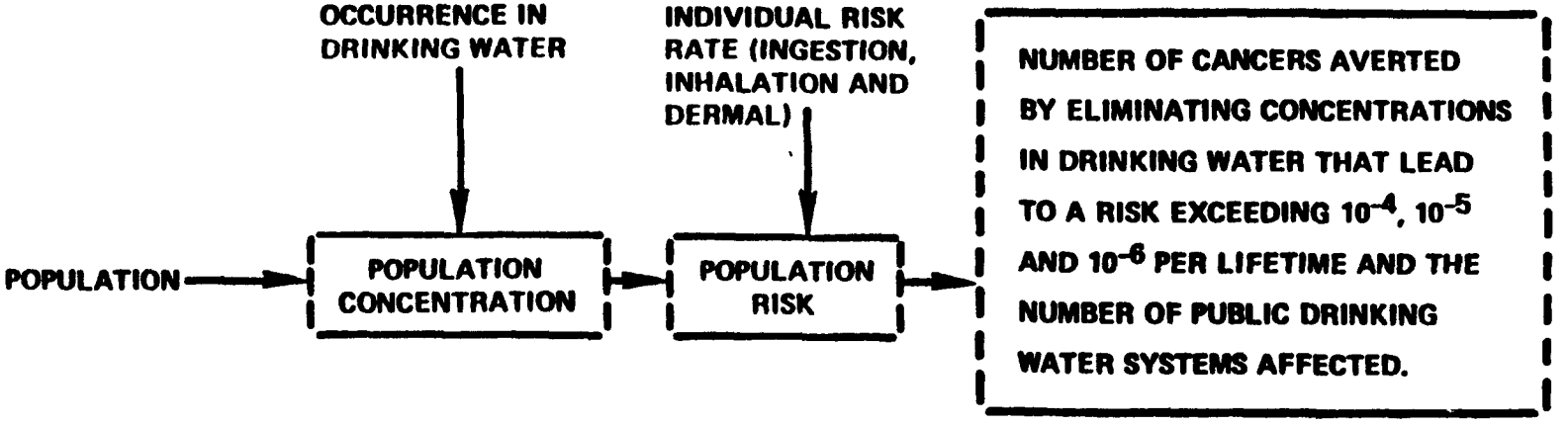

Figure 1. Flow chart showing the input of information to the risk estimation process for contaminants in drinking water.

P_{ic} - The number of people exposed to contaminant i at concentration c

C_{ic} - The amount absorbed from consumption of water containing contaminant i at concentration c

D_{ic} - The amount absorbed from dermal contact of contaminant i at concentration c

I_{ic} - The amount absorbed from inhalation of contaminant i which evolved from water of concentration c

R_{ic} - The individual risk rate for contaminant i at concentration c (this is expressed as the expected number of excess cancers in a lifetime per person per microgram/liter)

The total amount absorbed of contaminant i at concentration c is:

$$A_{ic} = C_{ic} + D_{ic} + I_{ic}$$

The overall population risk for the i^{th} contaminant will be:

$$PR = \text{Summation over n contaminants of } (P_{ic}\ R_{ic}\ A_{ic})$$

There are limits to our knowledge about any subject and the current state of knowledge about the health effects of contaminants in drinking water also has its limits. For some aspects, much is known, for others little or nothing is known, but in all areas there is uncertainty.

It is not possible to measure a quantity with 100% accuracy. There is uncertainty in any measurement. The tools of scientific measurement by their very nature involve both systematic and random error. In addition, several other uncertainties exist such as the effect of age, sex, species, organ

involved and rate of exposure of the test animals or humans. Almost always the dose-response data are known at high levels of exposure with no information being available at the levels of exposure for which a standard is being set.

If data only exist for animals, then uncertainty exists when extrapolated to humans. If there is exposure to more than one contaminant, the lack of information about possible synergistic and antagonistic effects adds to the uncertainty. Some of these factors tend to overestimate risk while others underestimate the risk.

In order to develop the present risk assessment, a number of assumptions n ed to be made. These assumptions include:

-- The average human ingests 2 liters of drinking water per day and inhales daily an average of 20 m^3 of air

-- The health endpoint for these contaminants is the same for animals and humans

-- 10^{-4} of that per liter in water is transferred to a liter of air

-- 100% of that in water is released to indoor air

-- The only difference between man and the test animal is one of scale

-- Dermal exposure is insignificant compared to oral and inhalation exposure

-- The mathematical expressions used reflect the biological dose-response curves at low doses

-- Everyone is exposed to some level of each contaminant

-- The benign and malignant animal tumors are indicative of cancer

-- Elevated human exposure during infancy does not alter risk

-- National monitoring is representative of the existing exposure profile

-- All of that ingested or inhaled goes to the bloodstream

-- There are no thresholds

The following sections describe the development of individual risk rates, estimates of exposure and population concentrations and calculation of population risk rates, and a detailed discussion of the contributions to the uncertainty.

ment, it was converted to ingestion values by converting the dose in ppm to microgram/m^3. This conversion is accomplished by multiplying the dose in ppm by 1.2 x 10^3 times the ratio (molecular weight of the chemical/molecular weight of air) to give a dose in micrograms/m^3.

Then, to convert the dose in microgram/m^3 to mg/kg/day, multiply by:

$$\frac{20 m^3}{day} \quad \frac{mg}{10^3 \text{ microgram}} \quad \frac{1}{70 \text{ kg}} = 2.86 \times 10^{-4}$$

The resulting human equivalent ingestion exposure bioassay data is then fit with four analytical models; viz, logit, multistage, probit, and Weibull. These four models are chosen somewhat arbitrarily to be generally representative of models currently popularly in use. The mathematical form of these models is shown in Table 1. It is well-known that other models could be used to fit the same data (tor more detail see Van Ryzin, 1980, a and b; Rai and Van Ryzin, 1979 and 1981; Krewski and Van Ryzin, 1981; and Krewski, 1982; and Munro and Krewski, 1981). There are no biologically based criteria for choosing one model over another. The models are in the form of an analytical expression that can be used to mathematically fit the high dose data. Figure 2 shows the shape of the curves for supralinear, sublinear, and linear. The models discussed here are generally sublinear in the low dose range. However, the multistage model is often linear at very low doses. The fits

INDIVIDUAL RISK RATES

The estimates of individual risk rates are based on bioassay data derived from animal experiments. These data are shown in Appendix A along with a description of how procedures are used to convert to a continuous human equivalent exposure. The bioassay data selected in this report to estimate carcinogenic potency are from experiments reviewed and selected by the National Academy of Sciences (1982) and the U. S. Environmental Protection Agency Carcinogen Assessment Group (See, for example, Anderson 1982 and Bayard 1983).

Some of the data was developed in study protocols calling for four or five days/week exposure or, in the case of inhalation experiments for a fraction of a day. Differences between experimental conditions and daily lifetime human exposure were compensated for in mathematical conversion factors as was the fraction of the animal lifetime over which the experimental data were collected. For example, if test animals were exposed five times/week and for fifty-two weeks of the one hundred-four week animal lifetime the dose was multiplied by 5/7 x 52/104 to extrapolate it to a continuous exposure. Extrapolation from animals to humans was accomplished by multiplying animal doses by the cube root of the ratio of the animal weight to human weight.

If the carcinogenic data were from an inhalation experi-

MODEL	PROBABILITY P(d) OF A RESPONSE AT DOSE d	LOW-DOSE BEHAVIOR		
		LINEAR	SUB LINEAR	SUPRA LINEAR
PROBIT	$(2\pi)^{-1/2} \int_{-\infty}^{\alpha+\beta \log d} \exp(-u^2/2)du \quad (\beta > o)$	—	$\beta>o$	—
LOGIT	$\left[1 + \exp(-\alpha - \beta \log d)\right]^{-1} \quad (\beta>o)$	$\beta=1$	$\beta>1$	$\beta<1$
WEIBULL	$1-\exp(-\lambda d^m) \quad (\lambda, m>o)$	$m=1$	$m>1$	$m<1$
ONE-HIT	$1-\exp(-\lambda d) \quad d>o$	$\lambda>o$	—	—
MULTI-STAGE	$1-\exp(-\sum_{j=1}^{k} \beta_j d^j) \quad (\beta_1 \geq o)$	$\beta_1>o$	$\beta_1 = o$	—
MULTI-HIT	$\int_o^{\beta d} \frac{(u^{k-1} e^{-u})}{\Gamma(k)} du$	$k=1$	$k>1$	$k<1$

Table 1. Mathematical form and low dose behavior of selected dose-response models.

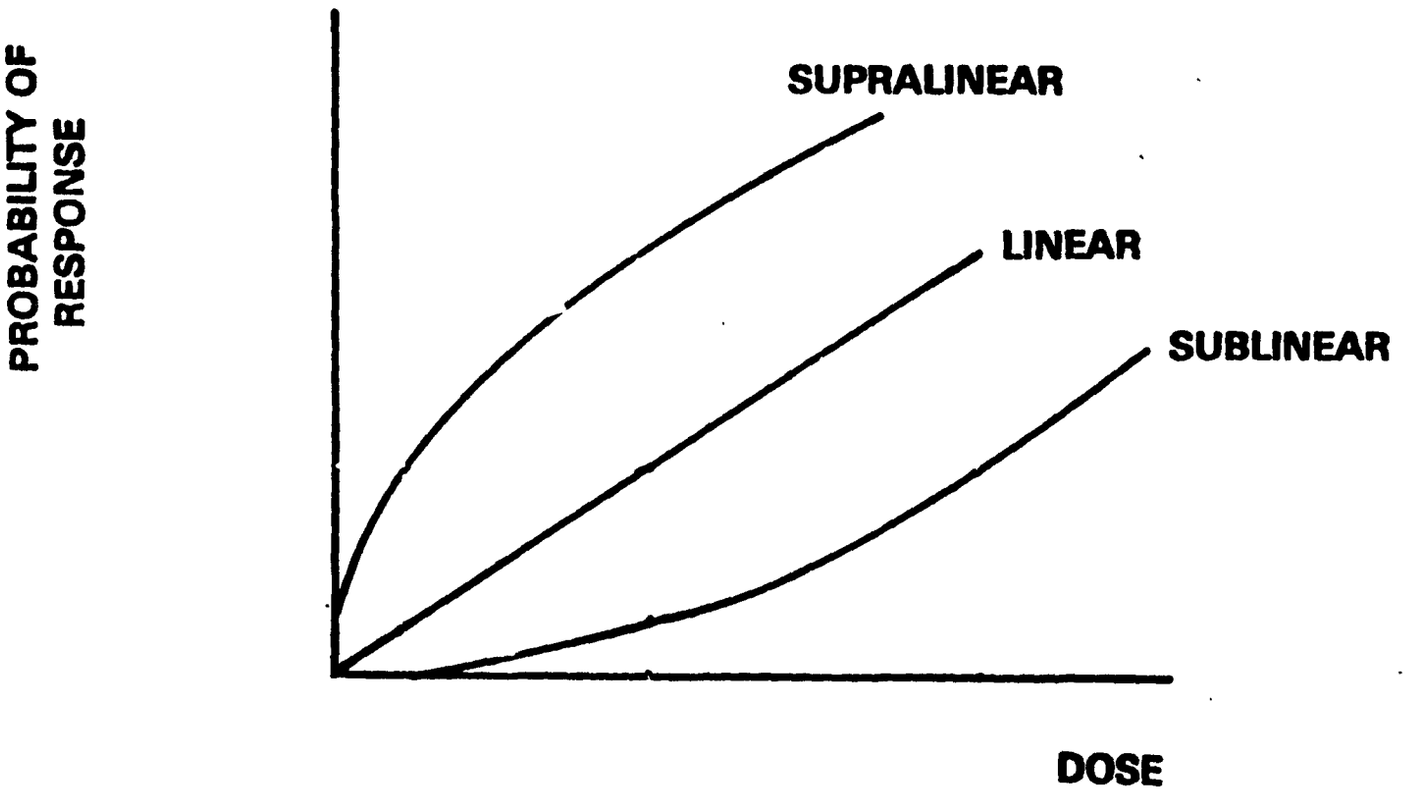

Figure 2. Range of shapes for possible dose-response models.

of these four models to the nine VOC's considered here are plotted on Figures 3-16. The dose axis was changed to water concentration by the following conversion.

The relationship between the mg/kg/day dose and microgram/liter water concentration can be calculated for a 70-kg adult human assuming 2 liters/day is ingested as follows:

$$1 \frac{mg}{kg\ day} \times \frac{day}{2\ liters} \quad 70kg \quad \frac{10^3\ microgram}{1\ mg}$$

$$= 35{,}000 \text{ micrograms/liter}$$

Figure 3 shows the fit to the actual bioassay data for trichloroethylene. Figure 4 shows a log-log plot of the entire region of interest showing the two non-zero data points. The solid lines are the point estimates and the error bars show the upper 95% confidence limits. The computer programs used were GLOBAL from Crump (1982) for the multistage model and Krewski (1982), for the logit, probit and Weibull models.

Figures 5-11 show the model fits to the next seven VOC's. The numerical values used to plot these groups are shown in Appendix B. As can be seen, no one model is consistently the highest or the lowest and the curves span a wide range of risks. Figure 7 for 1,1-dichloroethylene shows only one model because there was only one non-zero data point available from the experimental data. The other models require at least two non-zero data points.

Figure 12 shows the actual bioassay data for carbon tetrachloride. The logit, probit and Weibull models show an increasing risk for decreasing dose. The same behavior is shown on the more complete plots of Figures 13 and 14. This is clearly non-representative of reality the assumptions used.

It should be further noted that the curves shown in Figures 3-14 are not the only models that could be fit to the data. There is no assurance that the actual curve is in the range of those shown. It is possible (although highly unlikely; Van Ryzin, 1984) that the actual curve lies outside these shown. The probit model generally decreases extremely fast relative to dose and the Weibull model generally decreases very slowly relative to dose. This pretty much covers all possibilities.

Figure 4 presents the animal test data and model projections for trichloroethylene. Note that although dose-response curves projected by each model initiate from the same points, they diverge significantly at lower dose levels. At a drinking water concentration of 50 micrograms/liter the Weibull model provides a risk estimate approximating 1×10^{-2} while the probit model provides an estimated risk of 1×10^{-10}. The estimates made using these techniques differ by 8 orders of magnitude, an uncertainty equivalent of not knowing whether one has enough money to buy a cup of coffee or pay the national debt. Figures 5-14 provide information on the other VOC's.

Note that the uncertainty or projections from different models is from 3-5 orders of magnitude at a standardized drinking water concentration of 50 micrograms/liter. None of the models, including the multistage model, consistently provides either the highest or lowest risk estimates for all of the VOC's. Figures 12-14 show that the probit, Weibull and logit models project an increasing cancer risk for decreasing exposure to carbon tetrachloride. This suggests that mathematical extrapolations can be very misleading. Since the response data has not been adjusted for time-of-death (age adjusted) this may be part of the reason. Also, administered dose may not be organ dose (see Van Ryzin and Rai, 1981).

When it enters the body, carbon tetrachloride is converted to an active species by hepatic P450 enzymes. The enzymes are then destroyed by these active species. This phenomenon is called suicide substrate and generally describes the situation where a metabolite kills the enzyme which form it and thus reduces further formation of that species (Dambrauskas, 1970). This could possibly explain the behavior of the dose-response curve shown in Figures 12-14 for carbon tetrachloride.

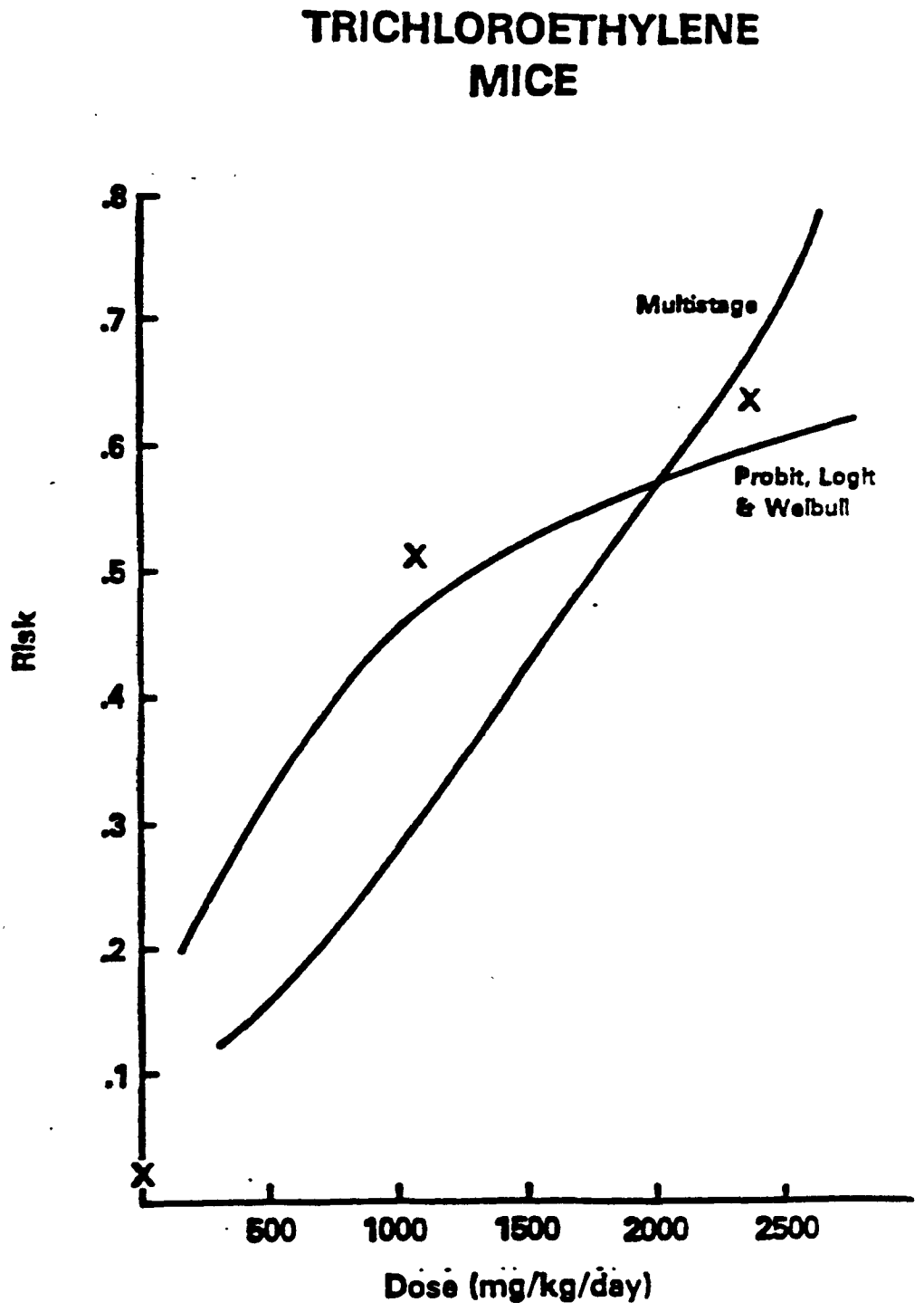

FIGURE 3

Converted dose-response data from experimental exposure to trichloroethylene.

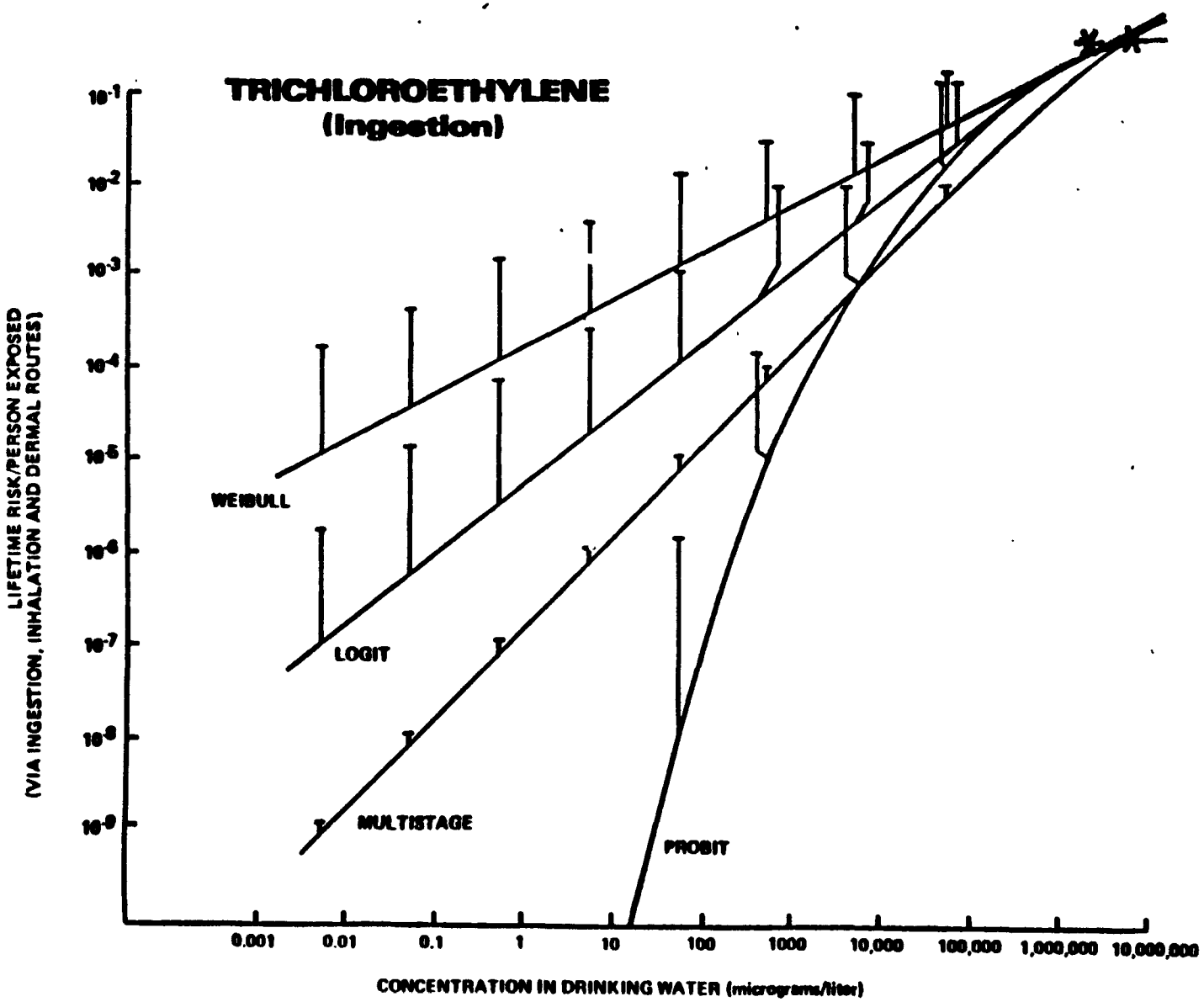

FIGURE 4

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multistage, probit and Weibull models from those bioassay data for trichloroethylene. The bioassay data was obtained by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to ingestion exposure. The error introduced by this simplifying assumption is less than the widths of the lines shown. The lines are the point estimates and the error bars indicate the upper 95% confidence limit. The lower 95% confidence limit is the

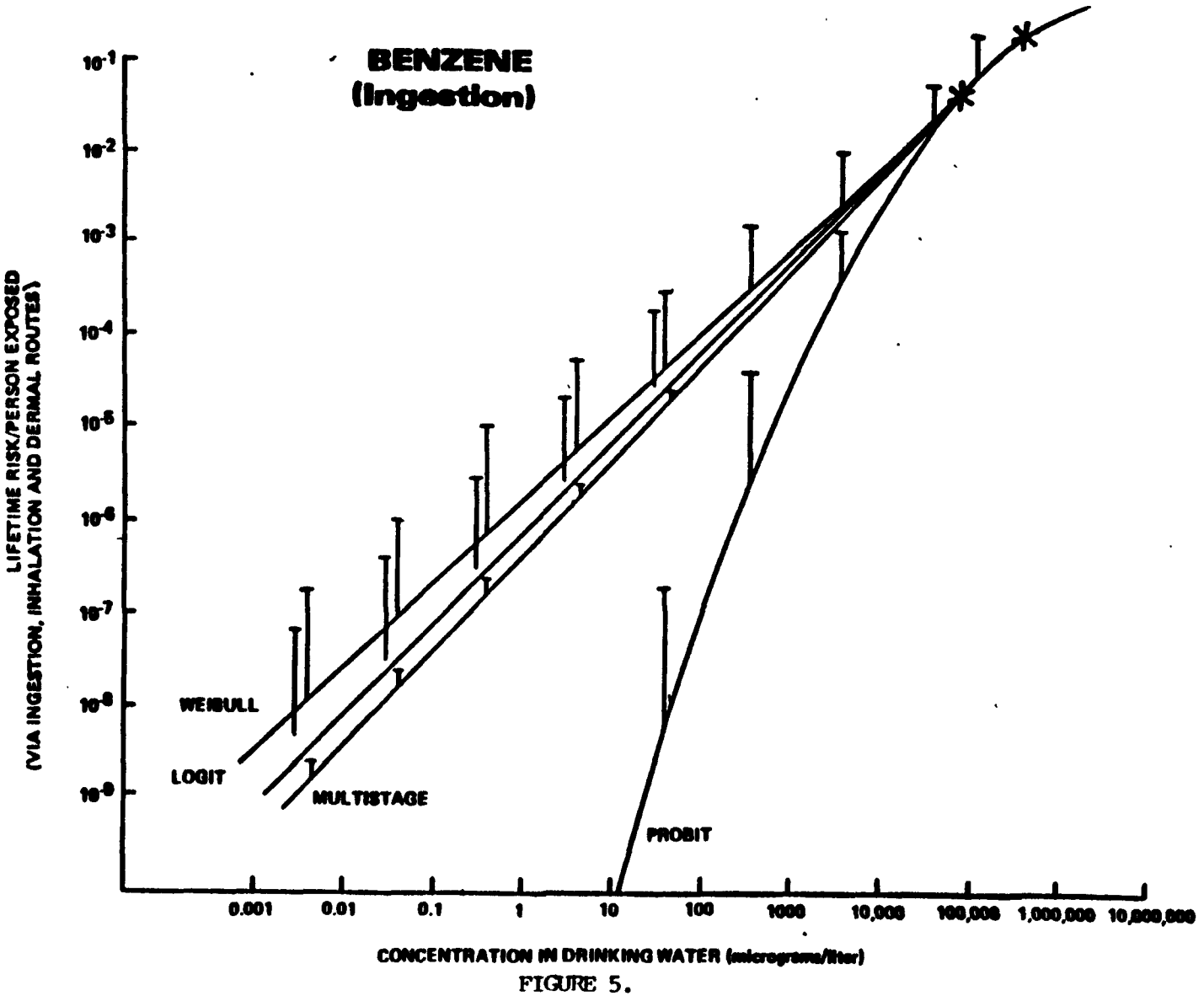

FIGURE 5.

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multistage, probit and Weibull models from those bioassay data for benzene. The bioassay data was obtained by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to ingestion exposure. The error introduced by this simplifying assumption is less than the width of the lines shown. The lines are the point estimates and the error bars indicate the upper 95% confidence limit. The lower 95% confidence limit

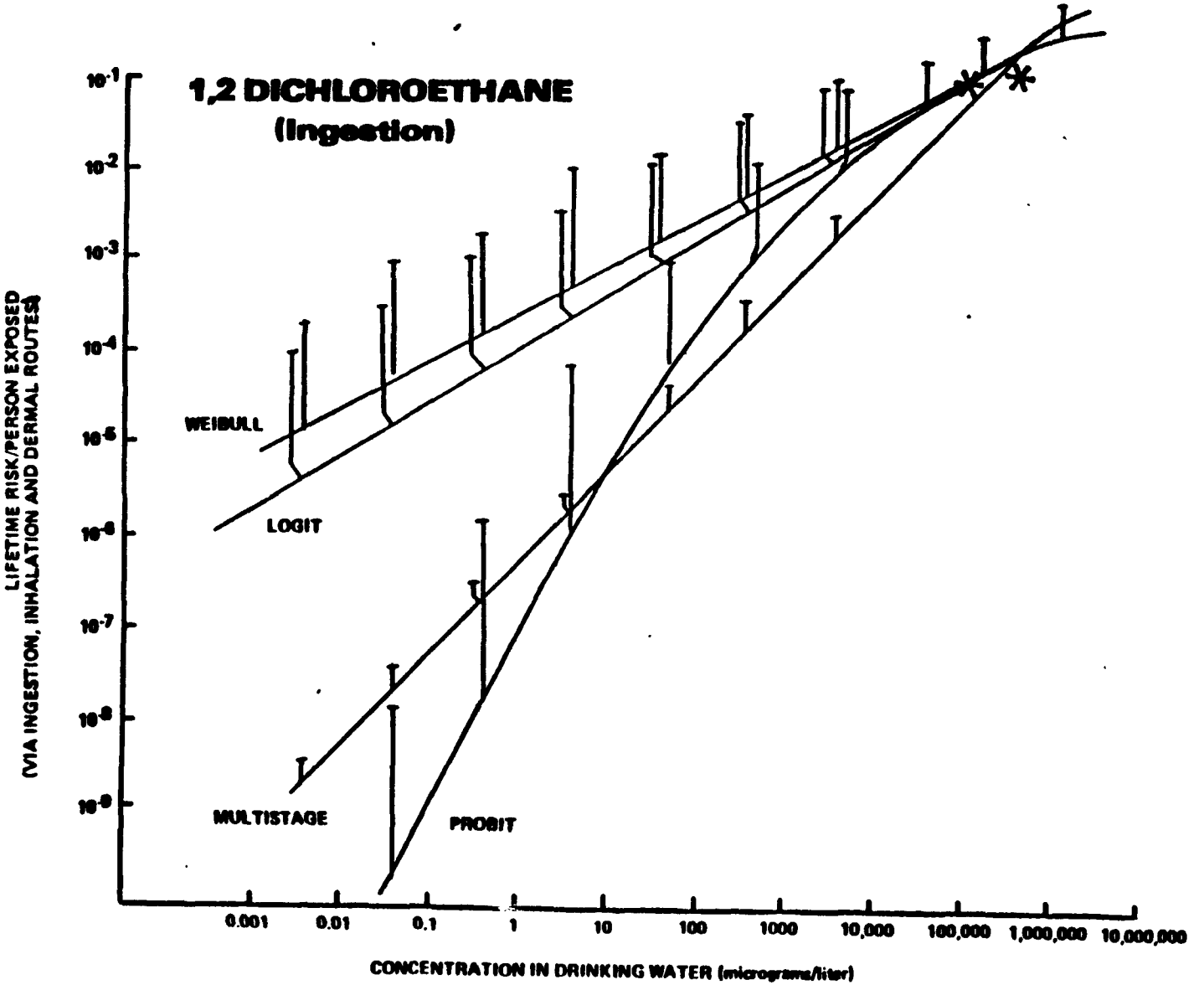

FIGURE 6

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multistage, probit and Weibull models from those bioassay data for dichloroethane. The bioassay data was obtained by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to ingestion exposure. The error introduced by this simplifying assumption is less than the widths of the lines shown. The lines are the point estimates and the error bars indicate the upper 95% confidence limit. The lower 95% confidence limit is the horizontal axis.

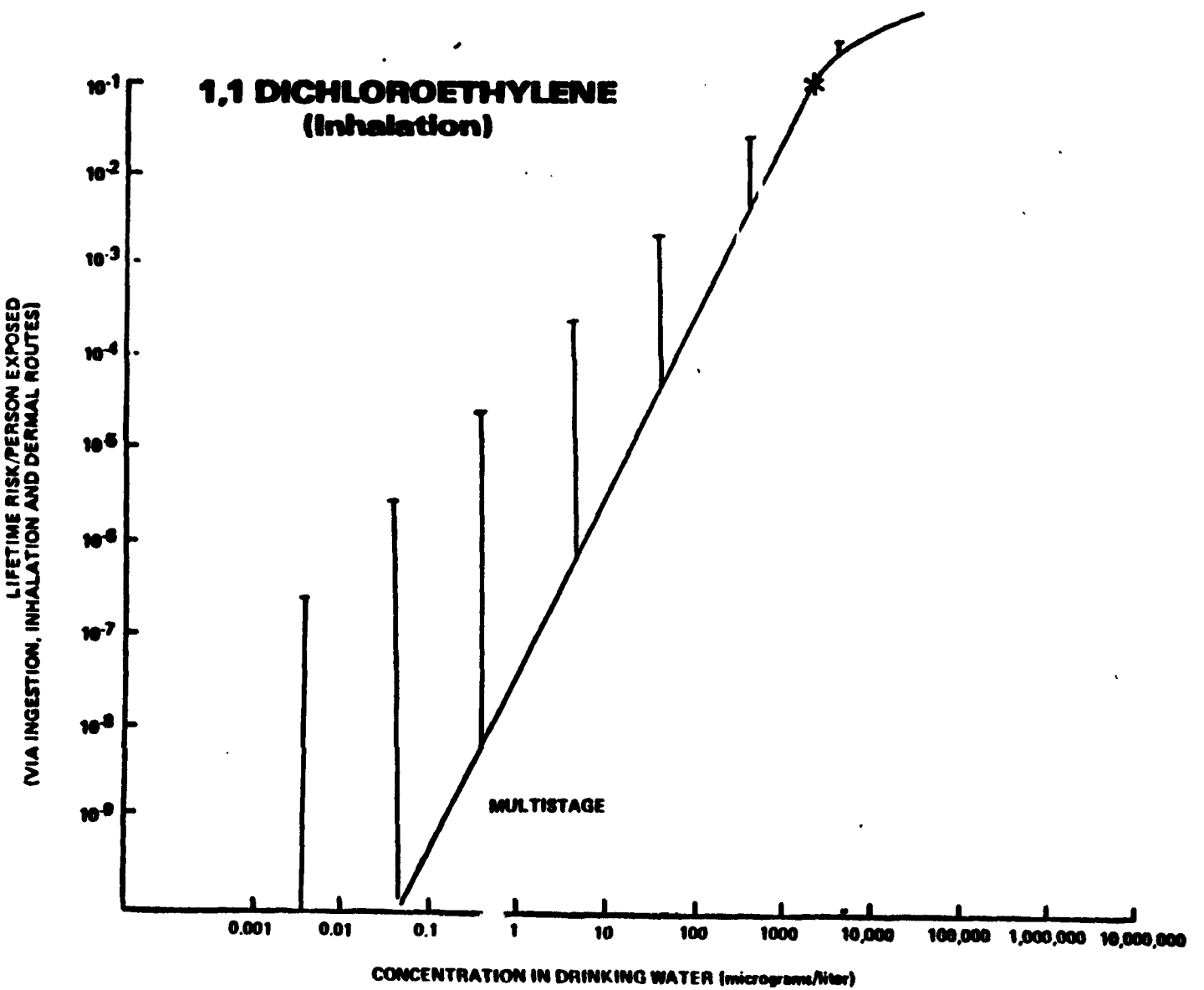

FIGURE 7

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multistage, probit and Weibull models from those bioassay data for dichloroethylene. The bioassay data was obtained by inhalation exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to ingestion

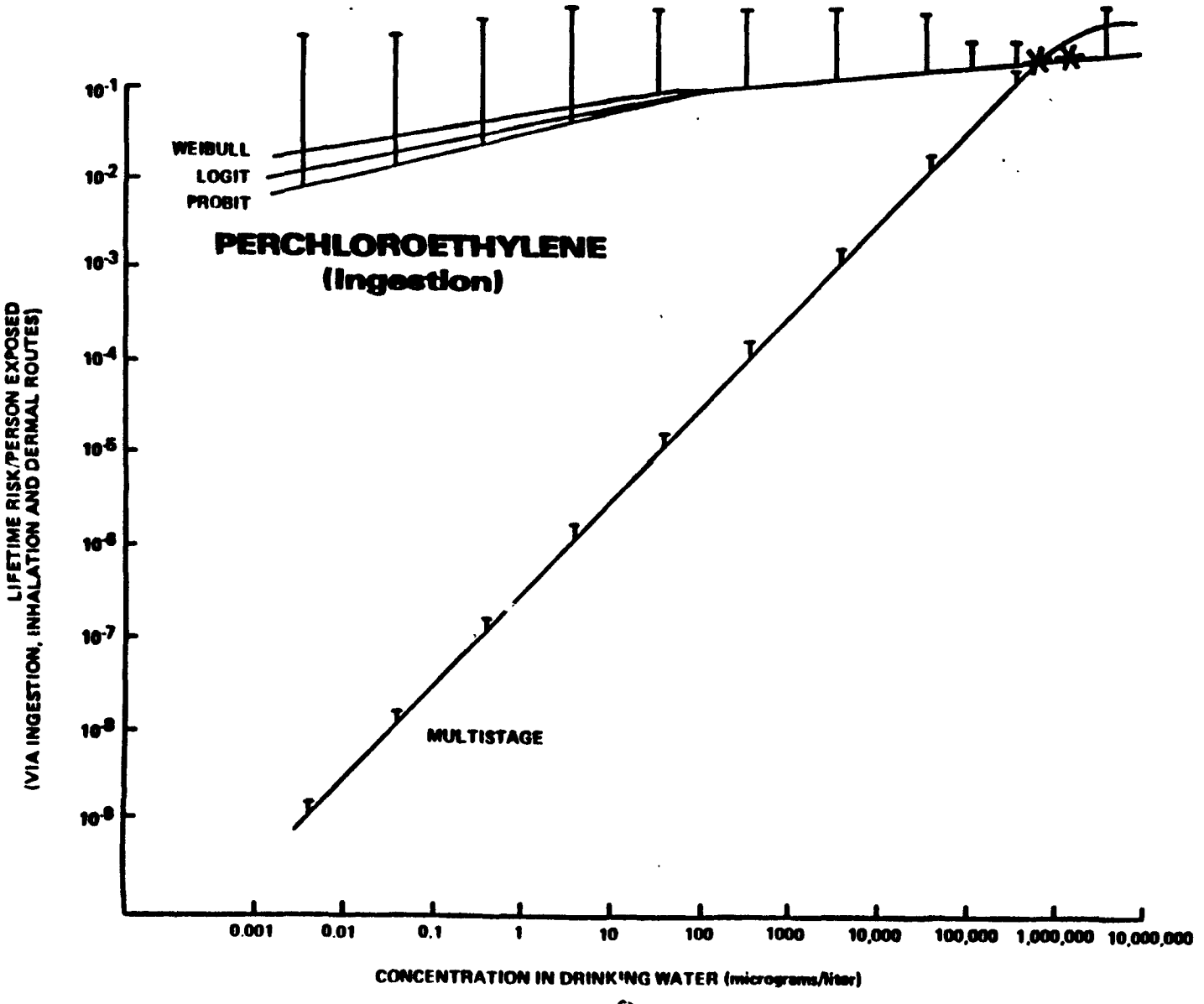

FIGURE 8

Bioassay data (starred points in the upper right hand corner) and model extrapolations for the logit, multistage, probit and Weibull models from those bioassay data for perchloroethylene. The bioassay data was obtained by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to ingestion exposure. The error introduced by this simplifying assumption is less than the widths of the lines shown. The lines are the point estimates of the

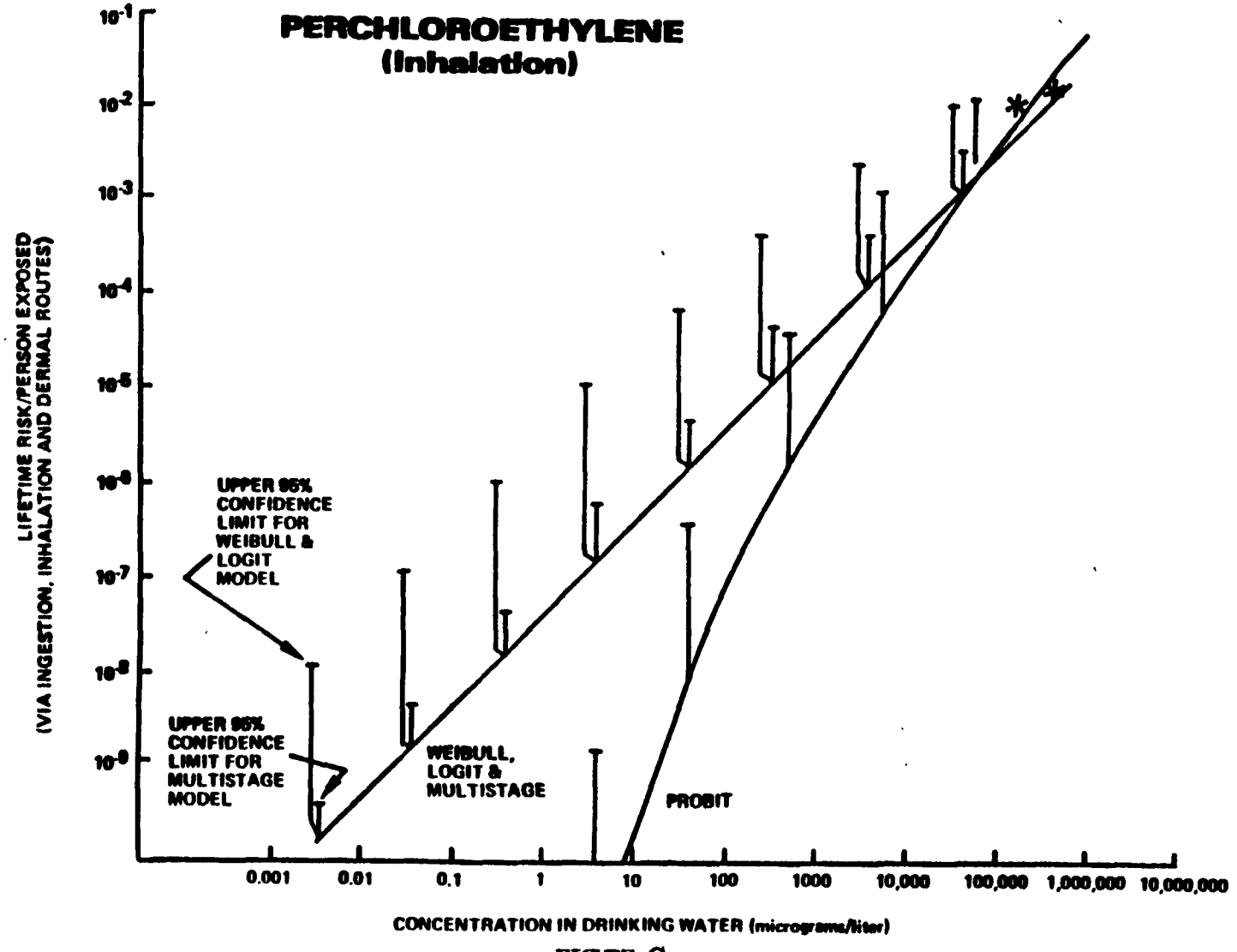

FIGURE 9

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multistage, probit and Weibull models from those bioassay data for perchloroethylene. The bioassay data was

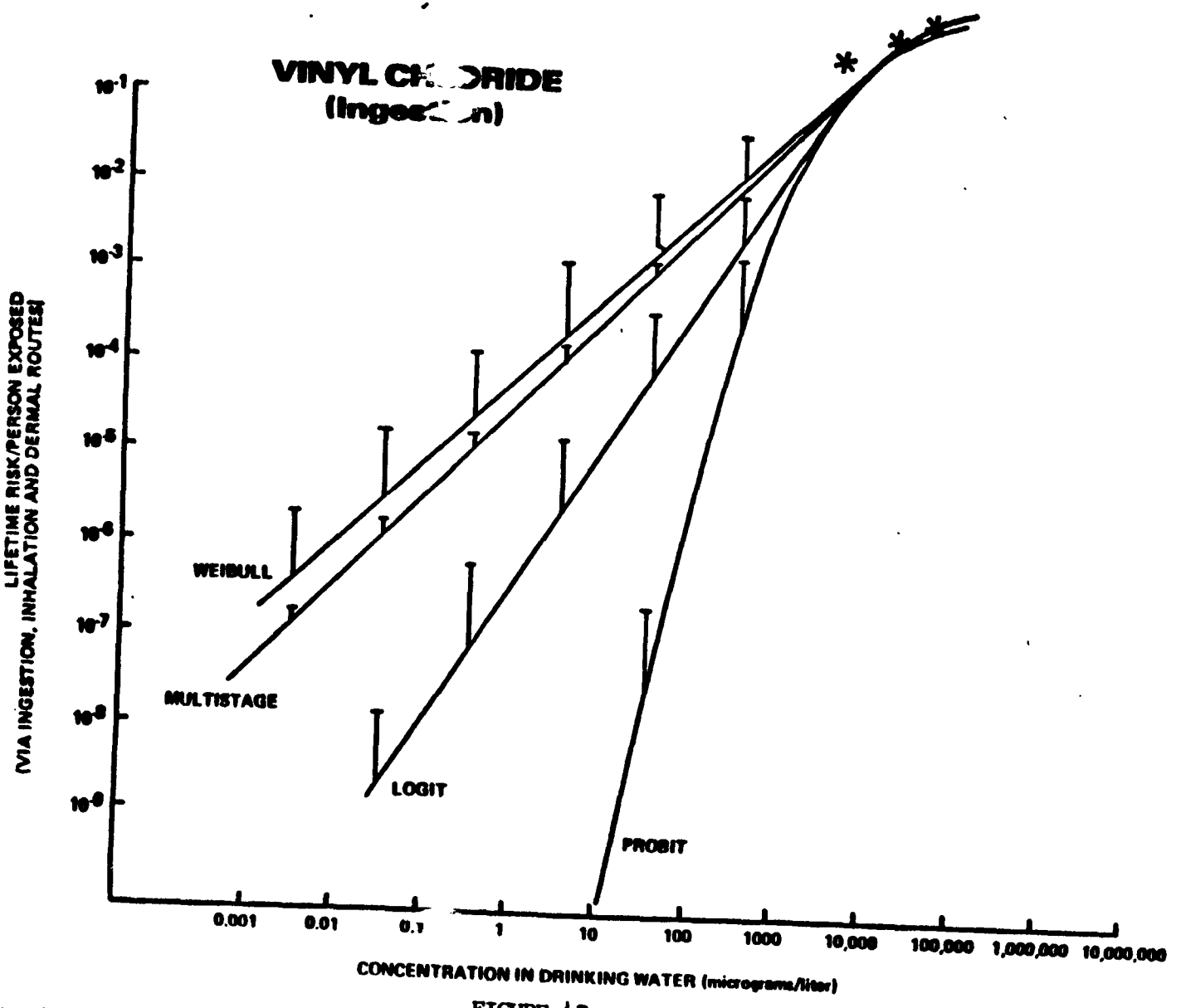

FIGURE 10

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit multi-stage, probit and Weibull models from those bioassay data for vinyl chloride. The bioassay data was obtained by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to ingestion exposure. The error introduced by this simplifying assumption is less than the widths of the lines shown. The lines are the point estimates and the error bars indicate the upper 95% confidence limit. The lower 95% confidence limit is the horizontal axis.

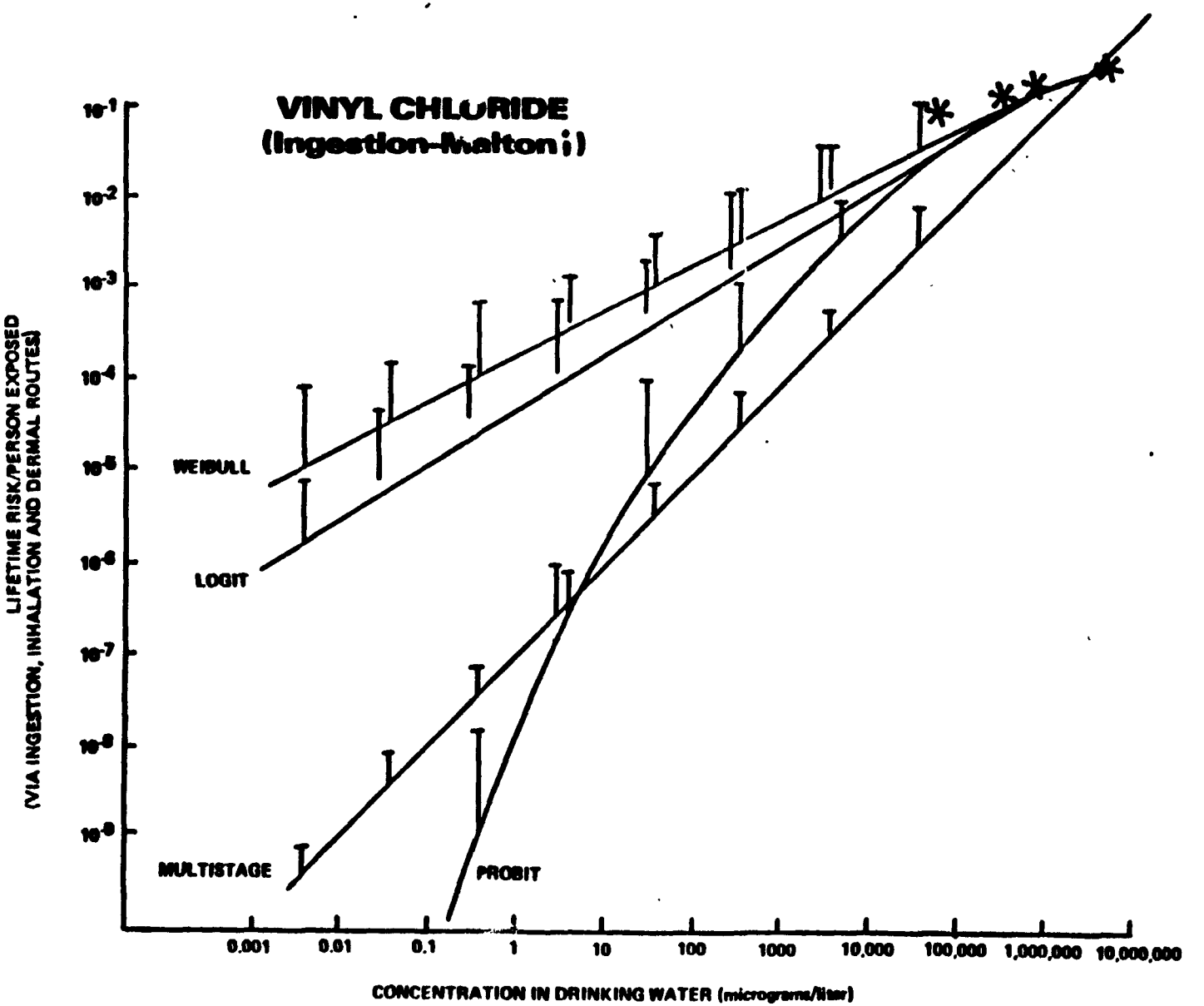

FIGURE 11

Bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multi-stage, probit and Weibull models from those bioassay data for vinyl chloride. The bioassay data was obtained by ingestion maltoni exposure. The lifeti[illegible]

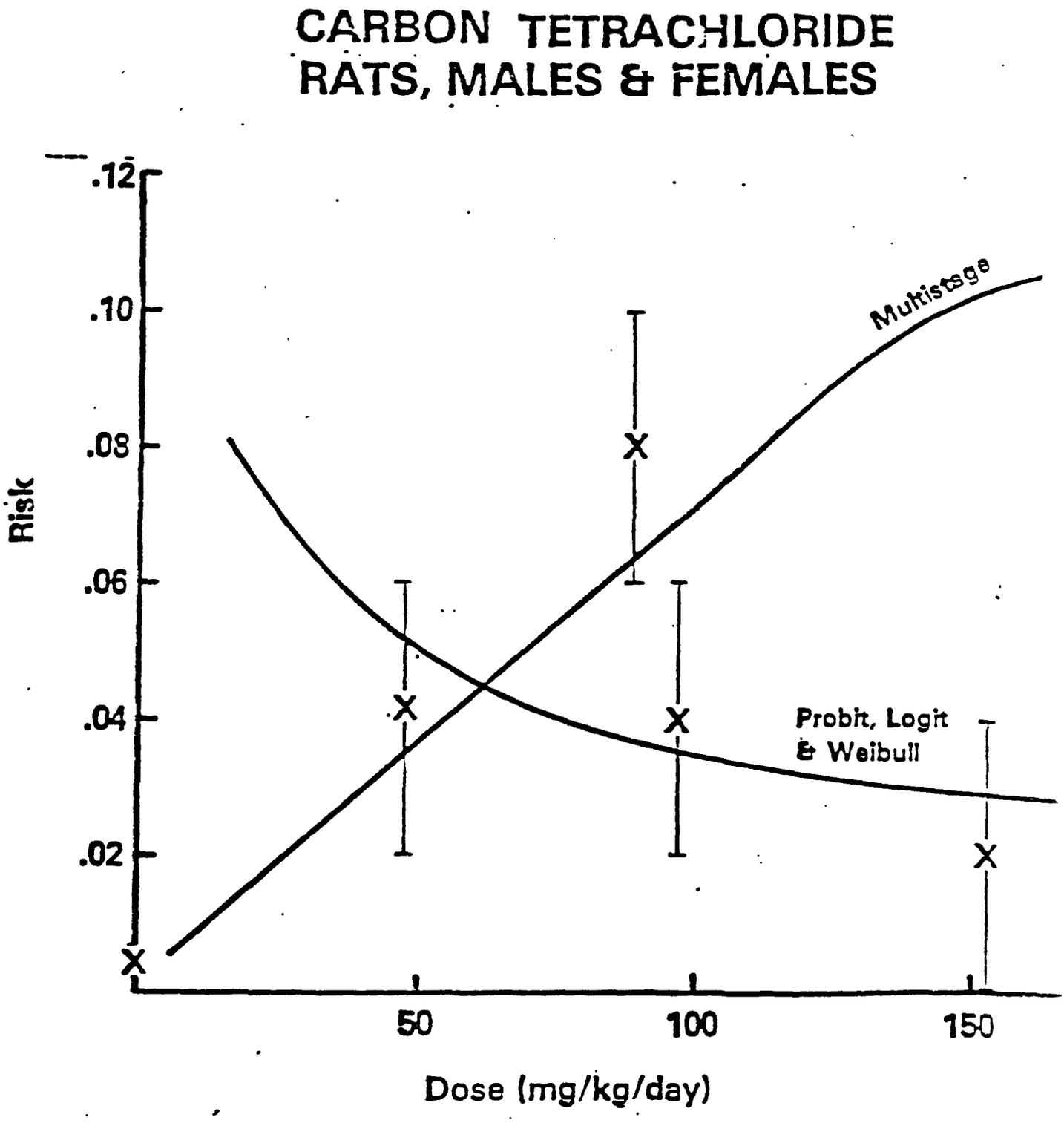

FIGURE 12. Converted dose-response data from experimental exposure to carbon tetrachloride.

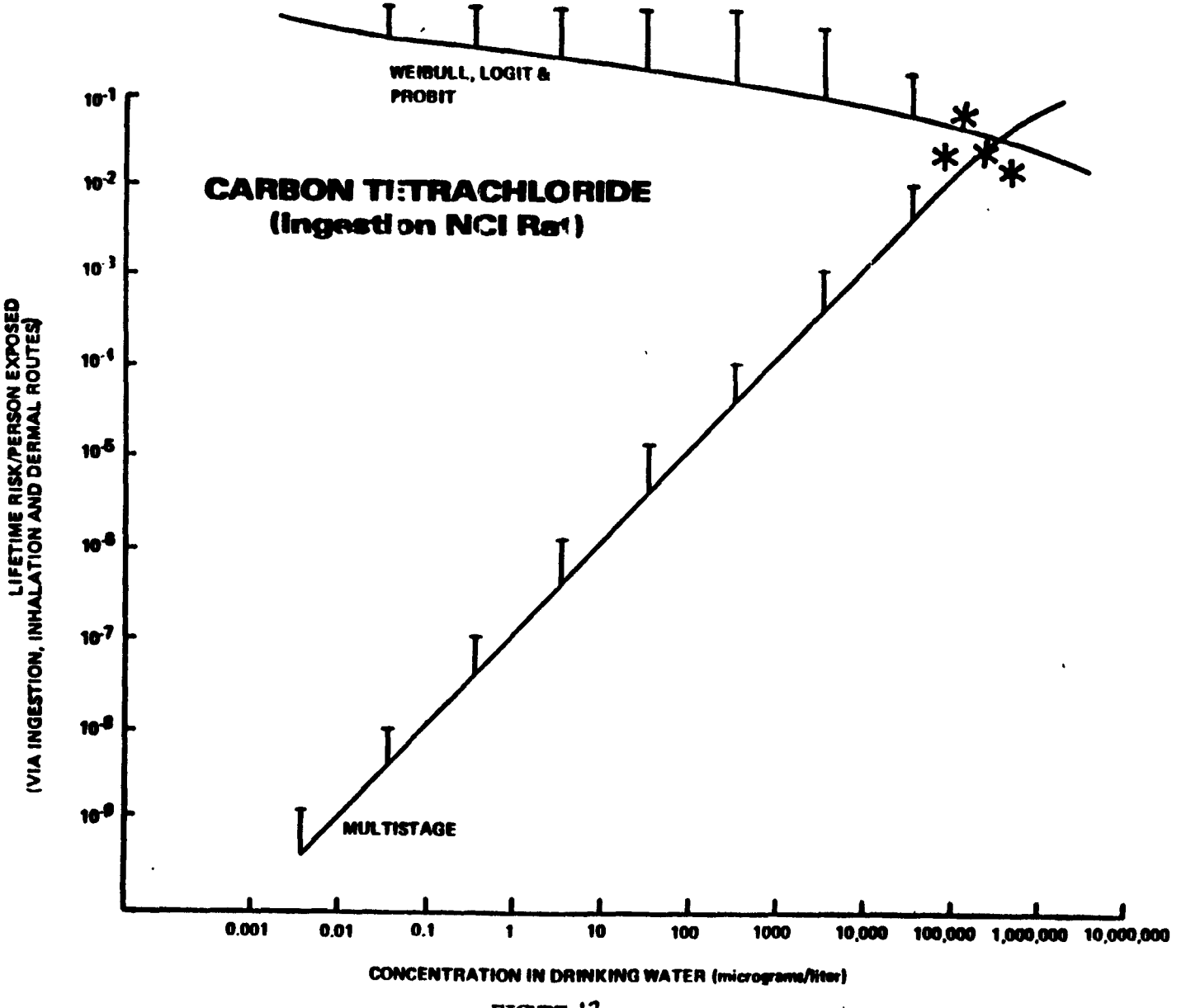

FIGURE 13

The bioassay data (starred points in the upper right hand corner) and model extrapolations for the logit, multi-stage, probit and Weibull models from those bioassay data for carbon tetrachloride. The bioassay data was obtain by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to in-

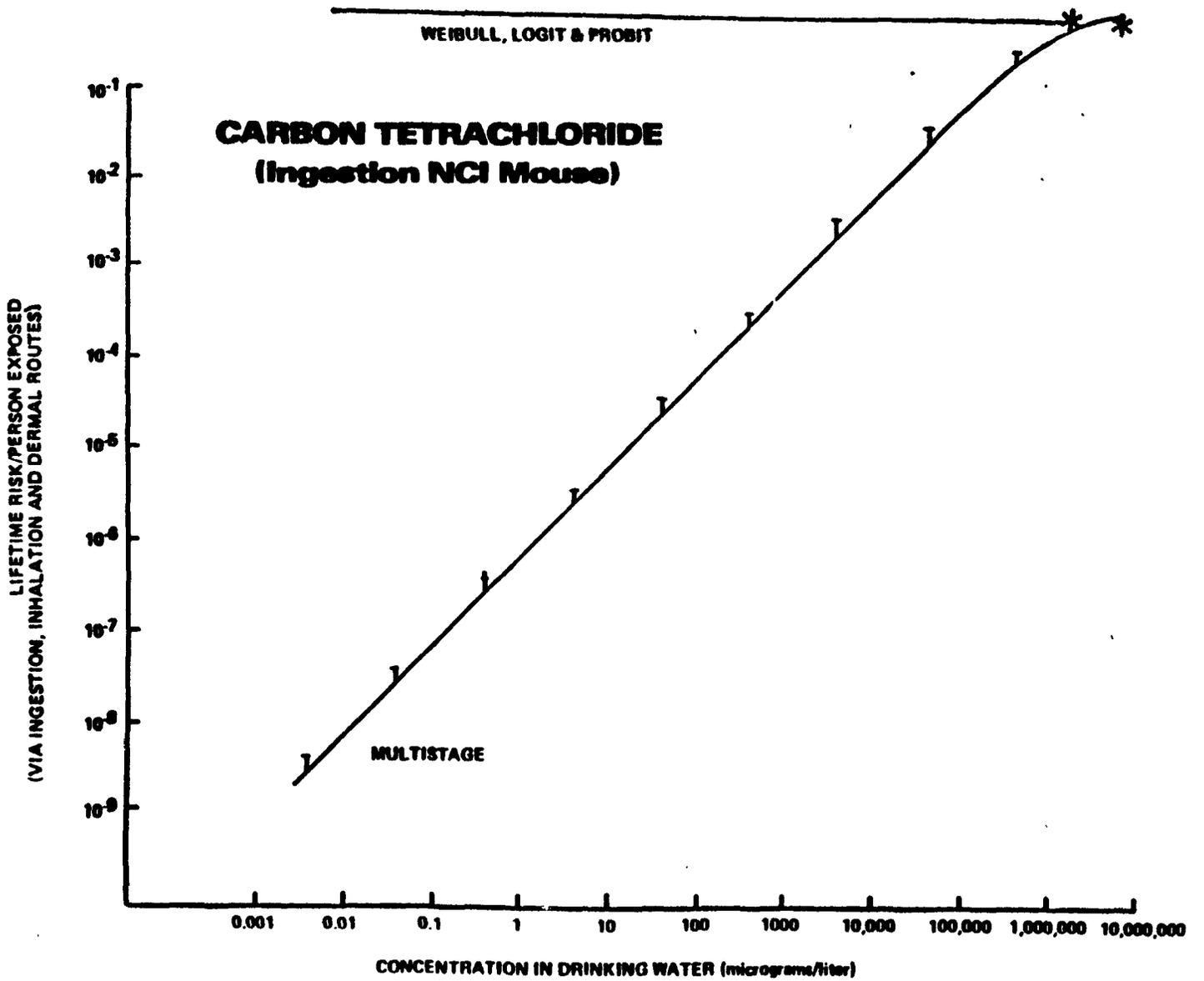

FIGURE 14

The bioassay data (starred points in upper right hand corner) and model extrapolations for the logit, multi-stage, probit and Weibull models from those bioassay data for carbon tetrachloride. The bioassay data was obtained by ingestion exposure. The lifetime risk/person exposed was calculated assuming that the risk due to inhalation and dermal exposure is equal to that due to ingestion exposure. The error introduced by this simplifying assumption is less than the widths of the lines shown. The lines are the point estimates and the error bars indicate the upper 95% confidence limit. The lower 95% confidence limit is the horizontal axis.

OCCURRENCE, EXPOSURE AND POPULATION-CONCENTRATION

The average adult in the United States uses 48-56 gallons or water per day (Watson, 1967, Linaweaver, 1969). This means that approximately 255 gallons of water per day would flow through a household inhabited by two (2) children and two (2) adults (Bailey, 1969). Under these conditions, 1 milligram/liter or one ppm concentration of a contaminant in drinking water means that nearly 1 gram of each substance is passing through the home. The remaining quantity of each contaminant is present in water used to launder clothes, clean the dishes and house and remove biological wastes from the home (Figure 15).

The quantity of any pollutant absorbed by each individual each day is the result of many personal choices and several factors over which we have very little direct control. Where one works and lives, what one eats and drinks, all have a profound influence on the exposure level to pollutants. Our personal preferences also make a difference in the magnitude of our exposure to the pollutants entering the home through drinking water. Table 2 provides an example of the potential differences in exposure of family members from drinking water pollutants. This table has been prepared for a theoretical contaminant (whose characteristics are similar to those of trichloroethylene) which is one hundred percent absorbed for all oral intake, fifty percent for respiratory intake and is

FIGURE 17

DISTRIBUTION OF A CONTAMINANT (1MG/L) ACCORDING TO HOUSEHOLD WATER USE PATTERN

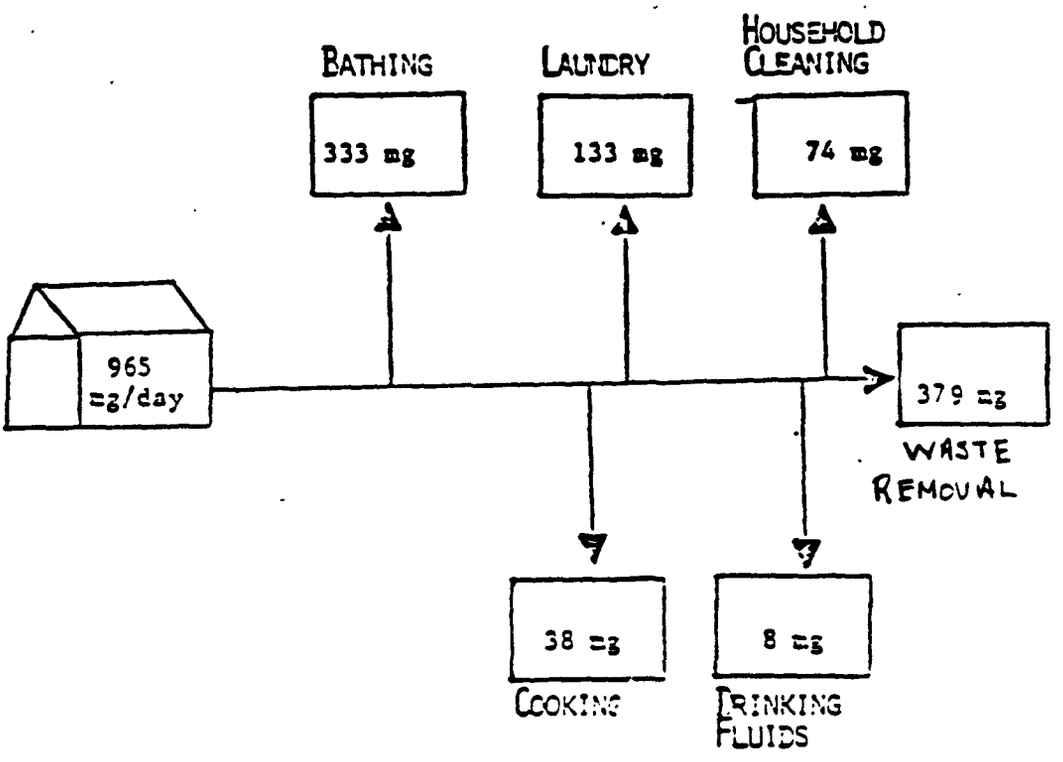

TABLE 2

COMPARATIVE MODEL OF ABSORBED DOSE FROM
A VOLATILE POLLUTANT (100 MICROGRAMS/LITER) IN DRINKING WATER

	Absorbed Dose (micrograms/day)				
	Formula-fed Infant (4 kg)	Pre-teen (32 kg)	Adult Female (60 kg)	Adult Male[a] (70 kg)	Male[b] (70 kg)
Fluid Ingestion	80	150	200	200	200
Inhalation of enriched indoor air	10	10	70	50	50
Inhalation of enclosed shower air	--	100	300	200	--
Dermal absorption					
Bathing	.02-.06	--	--	--	2.5
Swimming	--	10-300	--	--	--
Total Absorbed Dose micrograms/kg/day	20 micrograms/	10-20 micrograms/	10 micrograms/	7 micrograms/	4 micrograms/

a Showering adult male.

absorbed in direct proportion to water flux across the skin. These simplifying assumptions allow a relative comparison between the exposure to family members of different ages and activity patterns. Calculations using standard respiration and fluid intake show that drinking water contaminants can be responsible for higher levels of exposure per unit of body weight among children than adults and that respiratory exposure for volatile chemicals during and immediately after showering can be of the same magnitude as those created by direct ingestion. Prolonged swimming with its associated oral and dermal absorption also has the potential for significantly increasing exposure to drinking water contaminants.

The analysis which follows presents the distribution of concentration of selected volatile organic chemicals in a profile of U. S. drinking water supplies and provides a basis for estimating the resultant levels of human exposure.

There are approximately 60,000 public water supplies in the United States serving roughly 214 million people. Three major federally sponsored monitoring surveys have been conducted nation-wide to obtain information on the concentration of selected volatile organic chemicals appearing as pollutants in the potable water being provided to the nation's cities and towns. During these surveys, water supplies serving as few as 25 people were sampled. The data from these surveys have been assembled into a national profile that has been projected to

all public water supplies. No information is available from which to estimate conditions in private wells or systems not covered by the definition of a community water supply (serving more than 25 people on 15 connections).

The risk analysis being presented in this paper deals with a selected list of volatile organic chemicals which represents a subset of the VOC's for which monitoring data are availalbe (Table 5). In general the level of volatile chemicals found in surface water is lower than that in groundwater because of dilution and volatilization in the atmosphere during transport while the absence of these processes may cause groundwater levels to be very high. See Appendix C for extrapolation of the occurrence data (Letkiewicz, 1983). A description of the monitoring data assembled into an analysis of chemical occurrence and the subsequent national projections for each chemical follow.

In compiling the data, existing data were averaged if they were for the same site. No raw water data were used. Only data for finished and distribution water samples were used. The data were tested for biases by intercomparing the survey data. Analytical functions were chosen that best fit the existing data.

The following discussion for specific contaminants relies on information and technical details from Letkiewicz, et al.

TABLE 3

OCCURRENCE OF SELECTED DRINKING WATER IMPURITIES

	Ground Waters					Surface Waters				
Organics	N	% POS	Median micro-grams/ liter	3rd Quantile micro-grams/ liter	Max micro-grams/ liter	N	% POS	Median micro-grams/ liter	3rd Quantile micro-grams/ liter	Max
Benzene	1223	1.5	1.26	7.1	15.0	223	13	0.2	0.2	1.4
1,2-Dichloroethylene	984	3.0	1.2	2.0	34.7	251	2.4	0.5	1.4	2.3
1,1-Dichloroethane	939	2.7	0.3	0.9	6.3	103	2.0	0.4	0.5	0.5
1,2-Dichloroethane	1001	1.2	0.6	0.9	1.1	301	8.3	0.9	1.2	19.0
Trichloroethylene	994	7.0	1.0	4.1	160.0	296	16.5	0.6	0.9	32.0
Carbon Tetrachloride	1466	3.0	0.5	1.5	16	311	16	0.9	1.16	30
Tetrachloroethylene	1457	7.1	0.8	1.8	69	296	8.1	0.4	0.7	3.2
Vinyl Chloride	984	0.9	5.9	8.9	66	150	2.7	0.3	2.0	2.6

N = the number of systems.

(1983). A general overview of the amount of each contaminant released to air, land and water is shown in Table 4. The tables in Appendix C list the estimated concentrations of drinking water according to size of drinking water supply. To generalize the analysis, the level 0.5 micrograms/liter was chosen as the lower limit of quantitation. The level of 0.5 micrograms/liter was selected for these analyses because it was a relatively common minimum quantifiable concentration across all surveys.

TABLE 4

ESTIMATED AMOUNT OF EACH VOC RELEASED TO AIR, LAND AND WATER
(DATA FROM LETKIEWICZ, 1983)

	Amount Released to:		
Contaminant	Air	Land	Water
Benzene	111,000-288,000	770	2,600-13,500
Carbon Tetrachloride	43,000	1,200	25
1,2-Dichloroethane	11,900	100	250
1,1-Dichloroethylene	2,000	0	2
Perchloroethylene	153,000	104,000	400
Trichloroethylene	103,000	7,500	190

Benzene

Approximately 90% of the benzene used in 1978 was a feedstock in the production of three chemicals (ethylbenzene, cumene and cyclohexane). Other more minor uses included the production of nitrobenzene, chlorobenzenes, alkylbenzenes, maleic anhydride, and biphenyl, and use as a solvent and pesticide ingredient.

Total releases of benzene from production, indirect production, and use processes in 1978 were estimated to be 111,000-288,000 kkg to air, 770 kkg to land, and 2,600-13,500 kkg to water. Major benzene releases to air resulted from incomplete combustion of gasoline. Major releases to water occurred through oil spillage and solvent use. Benzene entering water is expected to be volatilized and to undergo degradation in the atmosphere.

Four Federal surveys were used to estimate the levels of benzene in the nation's public drinking water supplies: the National Organics Monitoring Survey (NOMS), the National Screening Program for Organics in Drinking Water (NSP), the 1978 Community Water Supply Survey (CWSS) and the Groundwater Supply Survey (GWSS). Additional state data are also reported, but were not used in developing the national estimates since these data came from only a few states and were not well-characterized with respect to water type and system size

sampled.

The data obtained from the four Federal studies was combined and sorted by source and size category in order to develop estimates of the number of systems nationally in each source/size category containing benzene within various concentration range. The methodology used is presented in Appendix D. The national estimate of systems was used to calculate the number of persons exposed to public drinking water containing benzene levels in those ranges.

Using the combined Federal survey data and the delta distribution model for projecting national occurrence, it was estimated that about 98.7% of the groundwater systems (Table C-1) of all sizes contain either no benzene at low levels and it was not possible to estimate how many are free of benzene contamination. Of the estimated 635 systems estimated to have levels higher than 0.5 micrograms/liter, 55 (0.1% of total groundwater systems) are estimated to have concentrations >10 micrograms/liter, and 2 systems (<0.01%) to have levels >50 micrograms/liter. It is estimated from the Federal data that no groundwater systems will have concentrations >80 micrograms/liter. It should be noted, however, that the state data indicate that some groundwater supplies may have benzene present at substantially higher levels.

For surface water supplies (Table C-2), it is estimated

that about 97% will have either no benzene present or levels <0.5 micrograms/liter. It is estimated that 301 surface water systems have levels between 0.5 and 5 micrograms/liter (2.7% of total surface water systems); none are estimated to have benzene above 5 micrograms/liter.

It is important to note that some of the Federal data used in computing the national estimates are from samples held for a prolonged period of time prior to analysis, with possible biodegradation of benzene. Therefore, these projections of national occurrence may underestimate actual contaminant levels.

Using the combined Federal survey data for surface water and groundwater supplies, it was estimated that 209,590,000 persons (97.7% of the population served by public drinking water systems) are receiving water either free of benzene combination, or having levels less than 0.5 micrograms/liter. Of the 4,829,000 persons (2.3%) receiving water containing benzene levels $\geq$ 0.5 micrograms/liter, an estimated 155,000 (<0.1%) are exposed to levels >5 micrograms/liter. About 3,200 (<0.1%) are estimated to be exposed to levels >30 micrograms/liter. Of the approximately 4.7 million people estimated to be exposed to levels ranging from 0.5 to 5 micrograms/liter, 81% obtain water from surface water supplies. However, all exposure to benzene in drinking water at levels about 5 micrograms/liter is projected to be from groundwater sources.

Carbon Tetrachloride

Production of carbon tetrachloride in 1981 amounted to 329,000 kkg. In addition to its direct manufacture, carbon tetrachloride may be produced indirectly during the production or breakdown of other chlorinated chemicals.

The major use of carbon tetrachloride is the production of chlorofluorocarbons 11 and 12, which was estimated to account for 87% of its use in 1978. Although chlorofluorocarbon production has declined since 1978, their production still accounted for 91% of carbon tetrachloride consumption in 1981. It is also used in grain fumigant formulations and has numerous other minor uses. Fumigant use was estimated to account for the majority of the environmental releases of carbon tetrachloride in 1978. Total releases of carbon tetrachloride from production and use processes in 1978 were estimated to be 43,000 kkg to air, 1,200 kkg to land, and 25 kkg to water. Carbon tetrachloride is quite stable in the environment. Quantities of carbon tetrachloride entering surface water are expected to be volatilized and to undergo eventual degradation in the stratosphere.

Six Federal surveys were used to estimate levels of carbon tetrachloride in the nation's public drinking water supplies: the National Organics Reconnaissance Survey (NORS), the National Organic Monitoring Survey (NOMS), the National Screen-

ing Program for Organics in Drinking Water (NSP), the 1978 Community Water Supply Survey (CWSS), the Rural Water Survey (RWS), and the Groundwater Supply Survey (GWSS). Additional state and miscellaneous data are also reported, but were not used in developing the national estimates for several reasons. These data came from only a few states and were not well-characterized with respect to water type and system size sampled. Since states do not monitor water supplies routinely, the reported data are believed to represent sites of recognized contamination and could, therefore, bias the national estimates.

The data obtained from the six Federal studies was combined and sorted by source and size category in order to develop estimates of the number of systems nationally in each source/size category containing carbon tetrachloride within various concentration ranges. The methodology used is provided in Appendix D. The national estimate of systems was used to calculate the number of persons exposed to public drinking water containing carbon tetrachloride levels in those ranges.

There was an apparent relationship between system size (based on population served) and the frequency of occurrence of carbon tetrachloride at levels $\geq$ 0.5 micrograms/liter in the combined survey data for both groundwater and surface water. Generally, a higher frequency of occurrence of carbon tetrachloride was found in the large size categories than in the small and medium size categories.

Using the combined survey data and the multinomial approach for projecting national occurrence, it was estimated that about 99% of the groundwater systems (Table C-3) of all sizes contain either no carbon tetrachloride or levels less than 0.5 micrograms/liter. It is not possible, however, to estimate how many of these systems actually contain carbon tetrachloride at low levels and how many are free of carbon tetrachloride contamination. Of the 355 systems estimated to have levels higher than 0.5 micrograms/liter, 110 (0.2% of total groundwater systems) are projected to have concentrations >5 micrograms/liter; none are estimated to have levels >20 micrograms/liter. Although a greater percentage of the very large systems are expected to be contaminated, in absolute numbers more of the smaller groundwater systems are expected to have contamination because of the very large number of small groundwater systems in the United States.

For surface water supplies (Table C-4), it is estimated that about 95% will have either no carbon tetrachloride present or levels <0.5 micrograms/ liter. It is estimated that 575 surface water systems have levels $\geq$0.5 micrograms/liter (5.1% of total surface water systems); 28 supplies (0.2%) are estimated to have levels >5 micrograms/liter. Only 9 (<0.1%) are expected to have levels above 20 micrograms/liter and none are estimated to have carbon tetrachloride above 30 micrograms/ liter.

Using combined data from surface water and groundwater supplies, it was estimated that 187,609,000 persons (87.5% of the population served by public drinking water systems) are receiving water with no carbon tetrachloride or levels at less than 0.5 micrograms/liter. Of the 26,810,000 persons (12.5%) receiving water containing carbon tetrachloride levels $\geq$0.5 micrograms/liter, an estimated 2,087,000 (1.0%) are exposed to levels >5 micrograms/liter. About 655,000 (0.3%) are estimated to be exposed to levels >20 micrograms/liter; no exposure above 30 micrograms/liter is expected. Of the approximately 27 million people estimated to be exposed to levels $\geq$0.5 micrograms/ liter, 92% obtain water from surface water supplies. All exposure to carbon tetrachloride in drinking water at levels above 20 micrograms/liter is projected to be from surface water sources.

1,2-Dichloroethane

1,2-Dichloroethane (CH_2ClCH_2Cl), also known as ethylene dichloride, is a colorless, flammable liquid with a pleasant odor. It has a high vapor pressure (63.8 mm Hg at 20° C) and a low water solubility (0.87 g/100 g at 20° C).

Production of 1,2-dichloroethane in 1980 amounted to 5,050,000 kkg. In addition to its direct manufacture, 1,2-dichloroethane may be produced indirectly during the production or breakdown of other chlorinated chemicals or during the

chlorination of water and wastewater. The production of 1,2-dichloroethane was estimated to account for approximately half of the environmental releases of 1,2-dichloroethane in 1978. The majority of 1,2-dichloroethane production facilities occur in Louisiana and Texas.

The major use of 1,2-dichloroethane is as a feedstock in vinyl chloride monomer production, which was estimated to account for 85% of its use in 1979. Other applications are numerous and include use as a feedstock for other chlorinated chemicals and gasoline lead scavenging. Total releases of 1,2-dichloroethane from production and use processes in 1978 were estimated to be 11,900 kkg to air, 100 kkg to land, and 250 k g to water. 1,2-Dichloroethane entering water is expected to be volatilized and to undergo degradation in the troposphere.

Five Federal surveys were used to estimate levels of 1,2-dichloroethane in public drinking water supplies in the United States: the National Organics Reconnaissance Survey (NORS), the Program for Organics in Drinking Water (NSP), the 1978 Community Water Supply Survey (CWSS), and the Groundwater Supply Survey (GWSS). Additional state data were reported for comparison, but were not used in developing the projections for several reasons. These data came from only a few states and were not adequately characterized as to water type and system size. Since states do not monitor water supplies

routinely, the reported data are believed to represent sites of suspected contamination.

The data obtained from the five Federal studies were combined and sorted by source and size category to allow for projections of the number of systems in each source/size category containing a specified level of 1,2-dichlorethane and the number of persons exposed to public drinking water containing 1,2-dichloroethane at a specified level. The methodology used is presented in Appendix D.

In general, a relationship was seen between groundwater system size and the likelihood of 1,2-dichloroethane contamination. Of the groundwater systems serving 5 3,300 persons, 0.6% contained detectable 1,2-dichloroethane, while 1.3% of the systems of intermediate size (3,301-10,000 persons) and 3.6% of the systems of large and very large size (>10,000 persons) contained detectable 1,2-dichloroethane.

Insufficient positive data were obtained for 1,2-dichloroethane in both surface water (Table C-5) and groundwater (Table C-6) systems to allow for statistical projections of the number of systems in each source/size category containing a specified level of 1,2-dichloroethane. However, since all 1,2-dichloroethane measurements in groundwater and surface water were below 20 micrograms/liter, it was estimated that there is effectively no 1,2-dichloroethane contamination at levels above 20 micro-

grams/liter. In addition, most of the systems in the United States are projected to have levels of 1,2-dichloroethane below 1.0 micrograms/liter.

It was estimated that all 195,595,000 persons using public drinking water supplies in the United States, including 3,129,520 bottle-fed infants, would be exposed to drinking water levels at or below 10 micrograms/liter. Because of limitations in the data, the number of persons exposed to more precise concentration intervals between 0-10 micrograms/liter could not be estimated. However, the majority of persons exposed to 1,2-dichloroethane in drinking water supplies are projected to receive concentrations less than 1.0 micrograms/liter.

1,1-Dichloroethylene

1,1-Dichloroethylene ($CH_2=CCl_2$), also known as vinylidene chloride, is a colorless liquid. It has a high vapor pressure (495 mm Hg at 20° C) and a low water solubility (0.035 g/100 g at 25° C).

Production of 1,1-dichloroethylene in 1978 was estimated to be 144,200 kkg. This figure included 1,1-dichloroethylene captively produced for the production of methyl chloroform. However, 1,1-dichloroethylene appears no longer to be used in methylchloroform production. In addition to its direct manu-

facture, 1,1-dichloroethylene may be produced indirectly during the production of other chlorinated chemicals.

The major uses of 1,1-dichloroethylene in 1978 were the production of methyl chloroform and the production of copolymers for use in resins, coating latex, and the manufacture of modacrylics. A minor application was its use in the production of chloroacetyl chloride, a component of mace and tear gas.

Total releases of 1,1-dichloroethylene from production and use processes in 1978 were estimated to be 2,000 kkg to air, 0 kkg to land, and 2 kkg to water. However, the production and use of methyl chloroform was estimated to account for 1,300 of the 2,000 kkg of 1,1-dichloroethylene released to air in 1978, and methyl chloroform no longer appears to be produced by this process. The production of copolymers is currently estimated to account for almost all environmental releases of 1,1-dichloroethylene. Quantities of 1,1-dichloroethylene entering water are expected to be volatilized and to undergo rapid degradation in the troposphere.

Two Federal surveys were used to estimate levels of 1,1-dichloroethylene in the nation's public drinking water supplies: the National Screening Program for Organics in Drinking Water (NSP) and the Groundwater Supply Survey (GWSS). Additional state data are also reported, but were not used in developing the national estimates. These data came from only

a few states and were not well-characterized with respect to water type and system size sampled.

Using the combined survey data and the multinomial approach for estimating national occurrence, it was calculated that about 98% of the groundwater systems (Table C-7), of all sizes contain either no 1,1-dichloroethylene or levels less than 0.2 micrograms/liter. It is not possible, however, to estimate how many of these systems contain 1,1-dichloroethylene at low levels and how many are free of 1,1-dichloroethylene contamination. Of the estimated 858 systems expected to have levels higher than 0.2 micrograms/liter, 81 (0.2% of total groundwater systems) are projected to have concentrations >5 micrograms/liter; none are expected to have levels >10 micrograms/liter. The state data, however, indicate that there may be some supplies with levels substantially higher than 10 micrograms/liter.

For surface water supplies (Table C-8), it is estimated that about 99.7% will have either no 1,1-dichloroethylene present or levels <0.2 micrograms/ liter. It is estimated that 35 surface water systems have levels $\geq$0.2 micrograms/liter (0.3% of total surface water systems).

It is important to note that some of the data used in computing the national estimates are from samples held for a prolonged period of time prior to analysis, with possible

biodegradation of 1,1-dichloroethylene. Therefore, these projections of national occurrence may underestimate actual contaminant levels.

Using combined data from surface water and groundwater supplies, it was estimated that 209,630,000 persons (97.8% of the population served by public drinking water systems) are receiving water with no 1,1-dichloroethylene or levels less than 0.2 micrograms/liter. Of the 4,789,000 persons (2.2%) receiving water containing 1,1-dichloroethylene levels $\geq$0.2 micrograms/liter, an estimated 52,000 (<0.1%) are exposed to levels >5 micrograms/liter. No individuals are estimated to be exposed to levels >10 micrograms/liter. Of the approximately 4.7 million people estimated to be exposed to levels ranging from 0.2 to 5 micrograms/liter, 52% obtain water from surface water supplies. All exposure to 1,1-dichloroethylene in drinking water at levels aboue 5 micrograms/liter is expected to be from groundwater sources.

Perchloroethylene

Perchloroethylene (CCl_2=CCl_2), also known as tetrachloroethylene, is a clear, non-flammable liquid with an ethereal odor. Commercial perchloroethylene has a high vapor pressure (20 mm Hg at 26.3° C) and a low water solubility (0.11 g/100 g at 25° C), two properties which make it a useful solvent.

Production of perchloroethylene in 1981 amounted to 313,000 kkg. In addition to its direct manufacture, perchloroethylene may be produced indirectly during the production or breakdown of other chemicals.

The major use of perchloroethylene is in dry cleaning and textile processing, which were estimated to account for 63% of its use in 1978. Other more minor applications include metal cleaning, fluorocarbon production, use as a specialty solvent, and use as an antihelminthic. Dry cleaning and metal cleaning were estimated to account for almost all environmental releases of perchloroethylene in 1978. Areas of releases from these processes are expected to parallel areas of greatest population density. Total releases of perchloroethylene from production and use processes in 1978 were estimated to be 153,000 kkg to air, 104,000 kkg to land, and 400 kkg to water. Some perchloroethylene may enter drinking water systems through the use of vinyl-lined A/C pipe. Quantities of perchloroethylene entering water are expected to be volatilized and to undergo degradation in the troposphere, while quantities in disposal sites (e.g., landfills) appear to be capable of moving through the soil to underground aquifiers.

Five Federal surveys were used to estimate levels of perchloroethylene in the nation's public drinking water supplies: the National Organic Monitoring Survey (NOMS), the National Screening Program for Organics in Drinking Water (NPS), the

1978 Community Water Supply Survey (CWSS), the Rural Water Survey (RWS), and the Groundwater Supply Survey (GWSS). Additional state data are also reported, but were not used in developing the national estimates for several reasons. These data came from only a few states and were not well-characterized with respect to water type and system size sampled. Since states do not monitor water supplies routinely, the reported data are believed to represent sites of recognized contamination and could, therefore, bias the national estimates.

There is some relationship between system size (based on population served) and the frequency of occurrence of perchloroethylene at levels $\geq$0.5 micrograms/liter in the combined survey data for both groundwater and surface water. (The level of 0.5 micrograms/liter was selected for these analyses because it was a relatively common minimum quantifiable concentration across all surveys.) The frequency of occurrence of systems with levels $\geq$0.5 micrograms/liter in groundwater for various system sizes was 3.6% (25-500 served), 2.2% (501-3,300 served), 8.5% (3,301-10,000 served), 5.6% (10,001-100,000 served), and 12.8% (>100,000 served). Statistical tests showed that the difference between the two smallest size categories was not significant, nor was the difference between the two largest size categories statistically significant. For surface water, the frequency of occurrence of systems with levels $\geq$0.5 micrograms/liter was 0% (25-500 served), 1.7% (501-3,300 served), 0% (3,301-10,000 served), 6.1% (10,001-100,000 served), and

5.3% (>100,000 served). In the case of surface water, the differences among the three smaller size groups was not statistically significant, nor was the difference between the two larger size categories.

Using the combined survey data and the multinomial approach for projecting national occurrence, it was estimated that about 97% of the groundwater systems (Table C-9), of all sizes contain either no perchloroethylene or levels less than 0.5 micrograms/liter. It is not possible, however, to estimate how many of these systems contain perchloroethylene at low levels and how many are free of perchloroethylene contamination. Of the estimated 1,552 systems estimated to have levels higher than 0.5 micrograms/liter, 322 (0.7% of total groundwater systems) are projected to have concentrations >0.5 micrograms/liter; only 3 systems have levels >30 micrograms/liter. It is estimated that no systems will have perchloroethylene concentrations >70 micrograms/liter. Although a greater percentage of the large systems are expected to be contaminated, in absolute numbers more small groundwater systems are expected to have contamination because of the very large number of small groundwater systems in the United States.

For surface water supplies (Table C-10), it is estimated that more than 98% will have either no perchloroethylene present or levels <0.5 micrograms/liter. It is estimated that 180 surface water systems have levels between 0.5 and 5 micro-

grams/liter (1.6% of total surface water systems); none are estimated to have perchloroethylene above 5 micrograms/liter.

The state and miscellaneous data generally support the national projections. A few instances of groundwater contamination at levels >70 micrograms/liter are reported for supplies in New Jersey; however, the type of water sampled in these cases is generally unknown. With respect to surface water, two supplies in Connecticut were reported to have levels >5 micrograms/liter (6 micrograms/liter; 31.2 micrograms/liter).

It is important to note that some of the data used in computing the national estimates are from samples held for a prolonged period of time prior to analysis, with possible biodegradation of perchloroethylene. Therefore, these projections of national occurrence may underestimate actual contaminant levels.

Using combined data from surface water and groundwater supplies, it was estimated that 202,989,000 persons (94.7% of the population served by public drinking water systems) are receiving water with no perchloroethylene or levels at less than 0.5 micrograms/liter. Of the 11,430,000 persons (5.3%) receiving water containing perchloroethylene levels $\geq$0.5 micrograms/liter, an estimated 874,000 (0.4%) are exposed to levels >5 micrograms/liter. About 440,000 (0.2%) are estimated to be

exposed to levels >10 micrograms/liter, and 105,000 (<0.1%) are estimated to be exposed to perchloroethylene in drinking water at levels in excess of 60 micrograms/liter. Of the approximately 10.5 million people estimated to be exposed to levels ranging from 0.5 to 5 micrograms/liter, 67% obtain water from surface water supplies. All exposure to perchloroethylene in drinking water at levels above 5 micrograms/liter is projected to be from groundwater sources.

Trichloroethylene

Trichloroethylene ($CHCl=CCl_2$) is a clear, non-flammable liquid with an ethereal odor. Commercial trichloroethylene has a high vapor pressure (58.7 mm Hg at 20° C) and a low water solubility (0.11 g/100 g at 25° C), two properties which make it a useful solvent.

Production of trichloroethylene in 1981 amounted to 117,100 kkg. In addition to its direct manufacture, trichloroethylene may be produced indirectly during the production or breakdown of other chemicals.

The major use of trichloroethylene is metal cleaning, which was estimated to account for 83% of its use in 1978. Other more minor applications include use as a polyvinyl chloride chain terminator and uses in adhesives, textiles, and paints. Metal cleaning was estimated to account for almost

all environmental releases of trichloroethylene (95%) in 1978. The majority of metal cleaning operations occur in the northeastern and midwestern United States. Total releases of trichloroethylene from production and use processes in 1978 were estimated to be 103,000 kkg to air, 7,500 kkg to land, and 190 kkg to water. Quantities of trichloroethylene entering water are expected to be volatilized and to undergo degradation in the troposphere, while quantities in disposal sites (e.g., landfills, lagoons) appear to be capable of moving through the soil to underground aquifiers.

Five Federal surveys were used to estimate levels of trichloroethylene in the nation's public drinking water supplies: the National Organics Monitoring Survey (NOMS), the National Screening Program for Organics in Drinking Water (NSP), the 1978 Community Water Supply Survey (CWSS), the Rural Water Survey (RWS), and the Groundwater Supply Survey (GWSS). Additional state and miscellaneous data are also reported, but were not used in developing the national estimates for several reasons. These data came from only a few states and were not well-characterized with respect to water type and system size sampled. Since states do not monitor water supplies routinely, the reported data are believed to represent sites of recognized contamination and could, therefore, bias the national estimates.

There was an apparent relationship between system size

(based on population served) and the frequency of occurrence of trichloroethylene at levels $\geq$0.5 micrograms/liter in the combined survey data for both groundwater and surface water. (The level of 0.5 micrograms/liter was selected for these analyses because it was a relatively common minimum quantifiable concentration across all surveys.) The frequency of occurrence of systems with levels $\geq$0.5 micrograms/liter in groundwater for various system sizes was 2.3% (25-500 served), 3.8% (501-3,300 served), 5.0% (3,301-10,000 served), 7.2% (10,001-100,000 served), and 25% (>100,000 served). The differences among the three smallest size categories were not statistically significant, however. For surface water, the frequency of occurrence of systems with levels $\geq$0.5 micrograms/liter was 0% (25-500 served), 1.7% (501-3,300 served), 5.1% (3,301-10,000 served), 10.6% (10,001-100,000 served), and 16.8% (>100,000 served). In the case of surface water, the differences among the three smaller size groups was not statistically significant, nor was the difference between the two larger size categories.

Using the combined survey data and the multinomial approach for projecting national occurrence, it was estimated that about 97% of the groundwater systems (Table C-12), of all sizes contain either no trichloroethylene or levels less than 0.5 micrograms/liter. It is not possible, however, to estimate how many of these systems contain trichloroethylene contamination. Of the estimated 1,632 systems expected to have

levels higher than 0.5 micrograms/liter, 421 (0.9% of total groundwater systems) are projected to have concentrations >5.0 micrograms/liter, and 133 systems (0.3%) have levels >50 micrograms/liter. It is estimated that only 64 systems (0.1%) nationally will have concentrations >100 micrograms/liter. Although a greater percentage of the large systems are expected to be contaminated, in absolute numbers more small groundwater systems are expected to have contamination because of the very large number of small groundwater systems in the United States.

For surface water supplies (Table C-13), it is estimated that about 96% will have either no trichloroethylene present or levels <0.5 micrograms/ liter. It is estimated that 498 surface water systems have levels $\geq$0.5 micrograms/liter (4.5% of total surface water systems); only 9 (<0.1%) are expected to have levels above 5 micrograms/liter; none are estimated to have trichloroethylene above 40 micrograms/liter.

The state and miscellaneous data tend to support the national projections, showing several instances of groundwater contamination at levels >100 micrograms/liter while these data show no surface water contamination above 20 micrograms/liter.

It is important to note that some of the data used in computing the national estimates are from samples held for a prolonged period of time prior to analysis, with possible biodegradation of trichloroethylene. Therefore, these projec-

tions of national occurrence may underestimate actual contaminant levels.

Using combined data from surface water and groundwater supplies, it was estimated that 189,288,000 persons (88.3% of the population served by public drinking water systems) are receiving water with no trichloroethylene or levels at less than 0.5 micrograms/liter. Of the 25,131,000 persons (11.7%) receiving water containing trichloroethylene levels $\geq$0.5 micrograms/liter, an estimated 1,844,000 (0.9%) are exposed to levels >5 micrograms/liter. About 212,000 (0.1%) are expected to be exposed to levels >50 micrograms/liter, and 42,000 (<0.1%) are estimated to be exposed to trichloroethylene in drinking water at levels in excess of 100 micrograms/liter. Of the approximately 23 million people exposed to levels ranging from 0.5 to 5 micrograms/liter, 76% obtain water from surface water supplies. However, of the 1.8 million people exposed to levels >5 micrograms/liter, 62% use groundwater sources. All exposure to trichloroethylene in drinking water at levels above 40 micrograms/liter is projected to be from groundwater sources.

Vinyl Chloride

Vinyl chloride (CH_2=CHCl) is a colorless, sweet-smelling gas. It has a high vapor pressure (2,580 mm HG at 20° C) and a low water solubility (0.11 g/100 g at 28° C).

Production of vinyl chloride in 1981 was 3,117,000 kkg. The major use of vinyl chloride is in the production of polyvinyl chloride, which is then manufactured into a wide range of products. The production of polyvinyl chloride polymers and copolymers is estimated to account for the majority of environmental releases of vinyl chloride (greater than 54,000 kkg in 1975). Unreacted vinyl chloride monomer present in polyvinyl chloride products accounts for small releases to water and land. Quantities of vinyl chloride entering water are expected to be volatilized and to undergo degradation in the troposphere.

Three Federal surveys were used to estimate levels of vinyl chloride in the nation's public drinking water supplies: the National Organics Monitoring Survey (NOMS), the National Screening Program for Organics in Drinking Water (NSP), and the Groundwater Supply Survey (GWSS). Additional state data are also reported, but were not used in developing the national estimates. These data came from only a few states and were not well-characterized with respect to water type and system size sampled.

Using the combined survey data and the multinomial approach for projecting national occurrence, it was estimated that about 99.9% of the groundwater systems (Table C-14) of all sizes contain either no vinyl chloride or levels less than 1.0 micrograms/liter. It is not possible, however, to estimate

how many of these systems contain vinyl chloride at low levels and how many are free of vinyl chloride contamination. Of the 31 systems estimated to have levels higher than 0.5 micrograms/liter, 17 (0.9% of total groundwater systems) are projected to have concentrations >5.0 micrograms/liter. Three of these are expected to have levels as high as 60-70 micrograms/liter, but none are estimated to have levels >70 micrograms/liter.

For surface water supplies (Table C-15), it is also estimated that about 99.9% will have either no vinyl chloride present or levels <1.0 micrograms/liter. It is estimated that 12 surface water systems have levels between 1.0 and 5.0 micrograms/liter (0.1% of total surface water systems); none are estimated to have vinyl chloride above 5 micrograms/liter.

It is important to note that some of the data used in computing the national estimates are from samples held for a prolonged period of time prior to analysis, with possible biodegradation of vinyl chloride. Therefore, these projections of national occurrence may underestimate actual contaminant levels.

Using combined data from surface water and groundwater supplies, it was estimated that 212,497,000 persons (99.1% of the population served by public drinking water systems) are receiving water with no vinyl chloride or levels at less than 1.0 micrograms/liter. Of the 1,922,000 persons (0.9%) receiv-

ing water containing vinyl chloride levels $\geq$1.0 micrograms/ liter, an estimated 591,000 (0.3%) are exposed to levels >5 micrograms/liter. About 118,000 (0.1%) are estimated to be exposed to levels >50 micrograms/liter. No individuals are estimated to be exposed to vinyl chloride in drinking water at levels in excess of 70 micrograms/liter. Of the approximately 1.3 million people estimated to be exposed to levels ranging from 1.0 to 5 micrograms/ liter, 65% obtain water from surface water supplies. All exposure to vinyl chloride in drinking water at levels above 5 micrograms/liter is projected to be from groundwater sources.

Population-Concentration

One of the ultimate goals of this analysis is to provide a profile of conditions in the United States answering the basic question: "How many people are being served drinking water containing each level of the pollutant from zero to the maximum quantity detected?" The national occurrence profile can then be converted into population-concentration estimates which summarize the national experience.

The population-concentration is calculated as the sum of the population exposed (P) times concentration (c) at various concentration levels, or

$$\text{population-concentration} = \underset{\text{n concentrations}}{\text{summation over}} \text{ of } (c_i \times P_i)$$

For example, if 10,000 persons were exposed to a chemical at 5 micrograms/liter and 4,000 additional persons were exposed to the same chemical at 10 micrograms/liter, the population-concentration calculated would be:

population-concentration = (10,000 persons x 5 micrograms/liter) + (4,000 persons x 10 micrograms/liter) = 90,000 micrograms/liter x persons.

These calculations would be straightforward if the number of individuals exposed to specific concentrations of the chemical were known. However, it is not possible to determine the specitic concentration to which a population is exposed.

Instead, projections have been performed which estimate the number of individuals exposed at various intervals of concentration (e.g., >5-10 micrograms/liter).

One way to express the uncertainty created by data grouped into intervals is to create upper and lower bounds if clusters occurred at the top and bottom of each interval. This becomes critical for the category of occurrence labeled "less than" where we do not know if the findings represent occurrence below detection or an artifact of the analytical methods. Mean best case and mean worst case total estimates were calculated using the lower and upper bounds of the first concentration interval (i.e., the values below the detection limit), and the mean value of the other concentration intervals. The lower and upper bounds of the first concentration interval were used rather than the mean value because of the higher uncertainty in knowing where the values fall in this interval in comparison to the other intervals. Example calculations for vinyl chloride follow:

Mean best case --

$$118 \times 10^3 \times 65 + 472 \times 10^3 \times 7.5 + 1331 \times 10^3 \times 2.75 = 14,870,250$$
$$= 1.5 \times 10^7 \text{ person x micrograms/liter}$$

Mean worst case --

$$118 \times 10^3 \times 65 + 472 \times 10^3 \times 7.5 + 1331 \times 10^3 \times 2.75 + 212,497 \times 10^3 \times 0.5 = 121,118,750$$
$$= 1.2 \times 10^8 \text{ person x micrograms/liter}$$

Calculation of the mean best and mean worst population concentration totals for each chemical are presented in Table C-18.

Another way of dealing with uncertainty is to assume a random distribution of data within the intervals and compute a standard error around the mid-point.

Using this method the population concentrations and standard errors can be calculated as follows: Let k be the number of systems sampled for a given contaminant. Suppose these k systems are broken down into m distinct subgroupings as follows: Let i denote the ith sub-group, i=1, . . ., n. Let sub-group i have: (1) c_i as the mid-point of the concentration (c_i = 7.5 if in the 5 to 10 range); (2) m_i as the number of persons served by systems in the ith range group (this must be available); and (3) f_i as the frequency of systems in the ith range group. Then, an estimate of the population concentration is PC = population-concentration = summation over i from 1 to n of $(c_i m_i f_i)$ and its standard error can be estimated by:

s.e. (PC) =

$$\left(\text{summation over } i \text{ from 1 to n}\right) c_i m_i^2 f_i \left(1 - \frac{f_i}{k}\right) - \frac{1}{k} \text{summation for i not = to j } c_i c_j m_i m_j f_i f_j)^{1/2}$$

We have k = summation over f_i. Also, the c_i for the first interval (0 to .5) would again be .25. The advantage to this

approach is that it also allows an estimate of the variability (the standard error) of the estimated PC. The results of using this relationship are shown in Table 5.

Factors in Human Exposure

Contaminants present in drinking water are transferred to people through the fluids they ingest, the enriched indoor air they inhale, and the contaminated water coming into contact with their skin. The intake of air and water are governed by metabolic rates reflecting age, sex, and physical condition and modified by personal activity. The assumptions made in this analysis about the magnitude of intake over a lifetime for the U. S. population provide a basis for a population risk analysis may be at wide variance for specific individuals.

Ingestion of Water

Risks projected for individual chemicals are based upon the chronic ingestion of 0.03 liters of drinking water and its contaminant, per kilogram of body weight and that 100% of the ingested chemical as absorbed. The factor for ingestion was developed using a standard of 2 liters intake per 70 kilograms adult.

TABLE 5

POPULATION CONCENTRATIONS
(in units of persons x micrograms/liter)

			TOTAL
BENZENE			
	ground	$(22 \pm 0.8) \times 10^6$	
	surface	$(34 \pm 1.7) \times 10^6$	$(56 \pm 2) \times 10^6$
CARBON TETRACHLORIDE			
	ground	$(24 \pm 1.3) \times 10^6$	
	surface	$(86 \pm 7.1) \times 10^6$	$(120 \pm 7) \times 10^6$
1,1-DICHLOROETHYLENE			
	ground	$(13 \pm 1.0) \times 10^6$	
	surface	$(49 \pm 1.2) \times 10^6$	$(62 \pm 2) \times 10^6$
1,2-DICHLOROETHYLENE			
	ground	$(21 \pm 0.8) \times 10^6$	
	surface	$(49 \pm 2.8) \times 10^6$	$(70 \pm 3) \times 10^6$
TETRACHLOROETHYLENE			
	ground	$(39 \pm 3) \times 10^6$	
	surface	$(1.3 \pm 0.2) \times 10^6$	$(40 \pm 3) \times 10^6$
TRICHLOROETHYLENE			
	ground	$(70 \pm 2.8) \times 10^6$	
	surface	$(74 \pm 8.4) \times 10^6$	$(144 \pm 9) \times 10^6$
VINYL CHLORIDE			
	ground	$(48 \pm 2) \times 10^6$	
	surface	$(55 \pm 2) \times 10^6$	$(103 \pm 3) \times 10^6$

Individuals may experience many times this intake rate. Formula-fed infants and young children have average intake rates as high as 8 times greater than those of "average adults". Adults in tropical areas may consume twice as much liquid as the "average" as may the athletically inclined when engaged in strenuous physical activity. The diseased may consume many fold more water than is indicated. Social behavior ritualized around the drinking of tea or coffee may lead to increased water consumption.

A recent Canadian study shows that children under 5 years of age, 10% of the children 6-17 years old and 2% of the adult population consume more than 0.03 liters/kilograms/day of drinking water (Canadian 1 81).

Inhalation

Each of the chemicals mentioned in this analysis have been shown to transfer from water into air if the water is heated or aerated (Sorrell, 1982). Monitoring data indicate that this process is continuous within the home leading to an immediate enrichment of respirable air at the point of water use and a diffusion throughout the home.

Showers taken within an enclosed bathroom using approximately 20-30 gallons of water containing volatile organic compound result in the liberation of all or a portion of the

compound into the air. A person who showers inhales air containing a highly enriched level of the chemical while within the bathroom.

Exposure is due to showering and other uses of water in the home. These sources mix with general home air resulting with some higher level of the pollutant volatilized into the indoor air.

The level of exposure for household members is determined by their personal habits and the air exchange in the home. Open windows and fans would make a large difference. Showering may take thirty (30) minutes during which the person could be inhaling air containing a large amount of a volatilized water pollutant. Calculations indicate that the resultant dose may be equivalent (within a factor of two) to that obtained by ingestion.

Very little data is available on the enrichment of general indoor air with volatile organics. Data from New Jersey validates its occurrence at specific locations within the home (Bishop, 1982). No work has been done on overall enrichment of indoor air with the organics. However, research has been conducted on radon enrichment of indoor air in Texas (Gessel and Pritchard, 1980). These data indicate that an enrichment of indoor air (amount of radon per liter of air) is about 10^{-4} to the concentration found per liter of drinking water. This

model assumes an air change rate of 1 per hour. Also see United Nations (1977), Hess (1982) and Kahlos (1980) for more discussion of this transfer process.

The total inhaled exposure is made up of two components, one in the shower and the other a sum of all other inhalation exposures indoors. Preliminary calculations indicate that except for shower exposure respiratory exposure would be equivalent to 16-20% of the oral intake for individuals spending a large amount of time indoors. Time spent within the highly enriched environment of a shower stall and bathroom would contribute to far more (83-125%) of the oral exposure.

For this analysis it has been assumed that respiratory exposure from volatile organic chemicals present in drinking water results in exposure equivalent to that from ingestion.

To compare the relative exposures from inhalation and ingestion, assume that for both pathways 100% of that which enters the body reaches the bloodstream. Then if water with a concentration of X micrograms/liter is ingested, the daily intake to the bloodstream is:

X micrograms/liter x 2 liter/day = 2X micrograms/day

If water with a concentration of X micrograms/liter transfers it all to indoor air and it all goes to the bloodstream,

then the intake to the bloodstream is:

$$X \frac{\text{micrograms}}{\text{liter}} \times \frac{10^{-4}\ \text{l(water)}}{\text{l air}} \times \frac{20\ \text{m}^3\text{(air)}}{\text{day}} \times \frac{10^6\ \text{cm}^3\text{(air)}}{\text{m}^3\text{(air)}} \frac{\text{l(air)}}{1000\ \text{cm}^3}$$

$$= 2X \text{ micrograms/day.}$$

Thus the exposure under these conditions from ingestion and inhalation is equal.

Dermal Absorption

The skin provides an outer barrier preventing free passage of foreign substances into the body. Water moves slowly through the skin because of its inability to pass through oil and waxes covering the skin. Chemicals more easily dissolved in oils pass through the skin more rapidly.

There are no data on the absorption of volatile organics present in diluted solutions found in bath water. It is evident that whole body emersion in 30 gallons of water places the body in contact with a relatively large amount of pollutants present at 1 ppm concentration (Figure 15).

Lacking data upon which to make this judgment, it was assumed that the contaminants passed through the skin at the rate of flux for water (0.2-0.5 $mg/cm^2/hr$). Although dermal

absorption during whole body emersion would be 2-3 times greater per unit of body weight for infants rather than adults, dermal absorption would contribute less than 1.5% of the oral dose experienced by "average adults" and less than .05% of the oral dose experienced by formula-fed infants (see Table 2). This analysis assumes that dermal absorption from bathing is not a significant source of exposure and risk for these compounds. Experimental data are needed to confirm this assumption.

POPULATION RISK CALCULATIONS

The population risks were estimated by using the procedure described earlier in this paper. Risk reduction analysis is also included. As can be seen in Tables 5 through 13 (Table 13 is a summary), the population risks are determined by a product of the number of people affected and the individual risk. Table 5 (benzene), for example, shows the estimated number of people exposed to a given concentration of benzene in drinking water. In each case the concentration is the average of the range. The lowest concentration range (less than 0.5 micrograms per liter) represents the minimum reported concentration for the national surveys and does not necessarily represent the detection limit of the chemical in drinking water. While we have used the rough average of 0.25 micrograms/liter for this category, it is possible that all of the people in this category could be exposed to 0 micrograms/liter or all of them could be exposed to the maximum of 0.5 micrograms/liter. For each chemical the individual lifetime risks were determined from the model fits described earlier.

In each case the curve giving the lowest individual risks and the curve giving the highest individual risks were used to determine the range of population risks. See Uncertainty section for discussion of rationale for this approach to including the uncertainty in this estimate. The reasoning for including only the uncertainty in choice of model is that

it is by far the largest uncertainty. When calculations yielded less than one person, the number was listed that way rather than the fractional number ot people. Also included in these tables is the risk reduction analysis. The risk reduction analysis is a determination of the number of cases averted if the standard is set corresponding to individual risk rates of 10^{-4}, 10^{-5} or 10^{-6}. In these individual risk rates the corresponding range of concentration and the cases averted are listed in Tables 5 through 13. The concentration range was determined from the graph shown earlier.

As seen in Table 5, the population risk estimates are listed as ranges. The range is indicative of the uncertainty in this estimate. For benzene the lower estimate for each concentration range is so small as to be less than one. The sum of the upper population risk estimates is 3,000 when rounded off to one significant figure. The majority of the population risk is contributed by the lowest concentration category. This category represents calculations below the detection limit and may be an artifact of the mathematical analysis scheme used to estimate the population risk. The range is largely a reflection of the uncertainty in choice of the mathematical model selected to estimate individual risk. The contribution to uncertainty from other factors is negligible by comparison.

Drinking water concentrations corresponding to individual

risk rates of 10^{-6} and 10^{-5} can fall below the minimum detectable level of 0.5 micrograms per liter. In the case of benzene less than 1 to 70 cases are due to drinking water concentrations above the "minimal reportable level". Therefore, if a drinking water standard (MCL) was established at 7.5 micrograms per liter, nationwide compliance would result in the aversion of between less than 1 and 10 cases (Table 6).

Tables 7-11 provide similar information on the remaining chemicals. Population risk estimates vary widely as do the estimated number of cases averted. A comparison of these results is presented in Table 12.

Investigation into the actual mechanism or the effect of vinyl chloride on health has revealed that it is actually a metabolite that does the damage (Gehring and Blau, 1977 and Gehring, 1978). This mechanism has the consequence that the responses predicted in Figures 10 and 11 are two to three orders of magnitude too high since the basis of that analysis was that the vinyl chloride was the direct contributor to health effects (Vay Ryzin, 1984). Thus, it appears that the estimates for vinyl chloride are too high. The experimental points for trichloroethylene and 1,2-dichloroethane appear similar to those of vinyl chloride (see Figures 4 and 6). Since all three of these estimates (see Table 12) were the highest of the VOC's analyzed, it might be suspected that the same mechanism is causing them to be high and this may be an

artifact of the analysis.

TABLE 6

BENZENE-POPULATION RISK ESTIMATES
FOR CURRENT LEVELS OF DRINKING WATER EXPOSURE

Mean Drinking Water Concentration (Micrograms/ Liter)	Number of People Being Served	Total Litetime Individual Risk for the Mean Concentration*		Lifetime Population Risk
		Low (Probit)	High (Weibull)	
0.25	2.1×10^{8}	$<10^{-10}$	1.7×10^{-6}	<1 - 360
2.75	4.5×10^{6}	$<10^{-10}$	1.3×10^{-5}	<1 - 60
7.5	1.2×10^{5}	$<10^{-10}$	3.1×10^{-5}	<1 - 4
15	6.2×10^{4}	2×10^{-9}	5.1×10^{-5}	<1 - 3
25	1.6×10^{4}	1×10^{-8}	8.9×10^{-5}	<1 - 1
35	6×10^{3}	3×10^{-8}	1.3×10^{-4}	<1 - 1
45	3×10^{3}	4×10^{-8}	1.6×10^{-4}	<1 - 1
55	2×10^{3}	5×10^{-8}	1.7×10^{-4}	<1 - 1
65	1×10^{3}	7×10^{-8}	2.2×10^{-4}	<1 - 1
75	1×10^{3}	1×10^{-7}	2.4×10^{-4}	<1 - 1
100	1×10^{3}	5×10^{-7}	3.2×10^{-4}	<1 - 1
			Total†	<1 - 400

Benzene-Risk Reduction Analysis
For Limiting Drinking Water Concentration

Maximum Allowable Drinking Water Concentration (Micrograms/Liter)	Approximate Individual Risk Rate for Maximum Concentration	Cumulative Cases Averted°
75	1×10^{-8} - 2×10^{-4}	<1 - 2
35	7×10^{-9} - 1×10^{-4}	<1 - 6
7.5	$<1 \times 10^{-10}$ - 8×10^{-5}	<1 - 10
2.75	$<1 \times 10^{-10}$ - 7×10^{-5}	<1 - 70

† Rounded to one significant figure.

* The total individual risk was determined by assuming that the risk due to inhalation exposure is equal to that due to ingestion.

° Number of cases averted for concentrations shown in the first column.

TABLE 7

1,2-DICHLOROETHANE-POPULATION RISK ESTIMATES
FOR CURRENT LEVELS OF DRINKING WATER EXPOSURE

Mean Drinking Water Concentration (Micrograms/ Liter)	Number of People Being Served	Total Lifetime Individual Risk for the Mean Concentration*		Lifetime Population Risk
		Low (Probit and Multistage)	High (Weibull)	
0.25	2.0 x 10^8	3 x 10-8	3.8 x 10-4	6 - 76,000
2.75	1.3 x 10^7	2 x 10-6	1.3 x 10-3	26 - 17,000
7.5	0	7 x 10-6	1.8 x 10-3	0 - 0
15	1.4 x 105	2 x 10-5	2.6 x 10-3	3 - 360
			Total†	40 - 90,000

1,2-Dichloroethane-Risk Reduction Analysis
For Limiting Drinking Water Concentration

Maximum Allowable Drinking Water Concentration (Micrograms/Liter)	Approximate Individual Risk Rate for Maximum Concentration	Cumulative Cases Averted°
15	6 x 10-7 - 8 x 10-3	3 - 360
7.5	1 x 10-6 - 1 x 10-2	3 - 360
2.75	1 x 10-5 - 2 x 10-2	29 - 17,000

† Rounded to one significant figure.

* The total individual risk was determined by assuming that the risk due to inhalation is equal to that due to ingestion.

° Number of cases averted for concentrations shown in the first column.

TABLE 8

PERCHLOROETHYLENE-POPULATION RISK ESTIMATES
FOR CURRENT LEVELS OF DRINKING WATER EXPOSURE
(USE INHALATION CURVE)

Mean Drinking Water Concentration (Micrograms/ Liter)	Number of People Being Served	Total Lifetime Individual Risk for the Mean Concentration*		Litetime Population Risk
		Low (Probit)	High (Weibull)	
0.25	2.0 x 10^{8}	<10^{-10}	3.1 x 10^{-8}	<1 - 6
2.75	1.0 x 10^{7}	<10^{-10}	3.4 x 10^{-7}	<1 - 3
7.5	4.3 x 10^{5}	2 x 10^{-10}	9.4 x 10^{-7}	<1 - 1
15	2.5 x 10^{5}	5 x 10^{-9}	1.8 x 10^{-6}	<1 - 1
25	8.2 x 10^{4}	1 x 10^{-8}	3.0 x 10^{-6}	<1 - 1
65	1.1 x 10^{5}	6 x 10^{-8}	8.2 x 10^{-6}	<1 - 1
			Total°	<1 - 10

Perchloroethylene-Risk Reduction Analysis
for Limiting Drinking Water Concentration

Maximum Allowable Drinking Water Concentration (Micrograms/Liter)	Approximate Individual Risk Rate for Maximum Concentration	Cumulative Cases Averted°
65	1 x 10^{-8} - 1 x 10^{-4}	<1 - 1
15	1 x 10^{-10} - 3 x 10^{-5}	<1 - 2
7.5	<1 x 10^{-10} - 1 x 10^{-5}	<1 - 3
2.75	<1 x 10^{-10} - 8 x 10^{-6}	<1 - 6

† Rounded to one significant figure.

* The total individual risk was determined by assuming that the risk due to inhalation is equal to that due to ingestion.

° Number of cases averted for concentrations shown in the first column.

TABLE 9

TRICHLOROETHYLENE-POPULATION RISK ESTIMATES
FOR CURRENT LEVELS OF DRINKING WATER EXPOSURE

Mean Drinking Water Concentration (Micrograms/ Liter)	Number of People Being Served	Total Lifetime Individual Risk for the Mean Concentration*		Lifetime Population Risk
		Low (Probit)	High (Multistage)	
0.25	1.9 x 108	<10-10	2.4 x 10-4	<1 - 45,600
2.75	2.3 x 107	<10-10	7.3 x 10-4	<1 - 16,790
7.5	4.3 x 105	<10-10	1.3 x 10-3	<1 - 559
15	2.1 x 105	<10-10	1.7 x 10-3	<1 - 367
35	7.4 x 105	7 x 10-8	2.3 x 10-3	<1 - 1,702
45	2.6 x 105	3 x 10-7	2.6 x 10-3	<1 - 676
55	4.2 x 104	4 x 10-7	2.8 x 10-3	<1 - 117
75	1.3 x 105	6 x 10-7	3.2 x 10-3	<1 - 416
100	4.2 x 104	1.2 x 10-6	3.7 x 10-3	<1 - 155
			Total†	<1 - 70,000

Trichloroethylene-Risk Reduction Analysis
For Limiting Drinking Water Concentration

Maximum Allowable Drinking Water Concentration (Micrograms/Liter)	Approximate Individual Risk Rate for Maximum Concentration	Cumulative Cases Averted°
100	1 x 10-7 - 1 x 10-2	<1 - 155
45	1 x 10-9 - 1 x 10-2	<1 - 1,365
7.5	<1 x 10-10 - 6 x 10-3	<1 - 3,993
2.75	<1 x 10-10 - 4 x 10-3	<1 - 20,783

† Rounded to one significant figure.

* The total individual risk was determined by assuming that the risk due to inhalation is equal to that due to ingestion.

° Number of cases averted for concentrations shown in the first column.

TABLE 10

VINYL-CHLORIDE-POPULATION RISK ESTIMATES
FOR CURRENT LEVELS OF DRINKING WATER EXPOSURE
(MALTONI-OLD)

Mean Drinking Water Concentration (Micrograms/ Liter)	Number of People Being Served	Total Lifetime Individual Risk for the Mean Concentration*		Lifetime Population Risk
		Low (Probit)	High (Weibull)	
0.25	2.1×10^{8}	2×10^{-9}	2.1×10^{-4}	<1 - 50,000
2.75	1.3×10^{6}	3×10^{-7}	1.1×10^{-3}	<1 - 1,430
7.5	4.7×10^{5}	1×10^{-6}	2.2×10^{-3}	<1 - 1,030
65	1.2×10^{5}	9×10^{-6}	3.5×10^{-3}	<1 - 420
			Total†	<1 - 50,000

Vinyl Chloride (Maltoni)-Risk Reduction Analysis
For Limiting Drinking Water Concentration

Maximum Allowable Drinking Water Concentration (Micrograms/Liter)	Approximate Individual Risk Rate for Maximum Concentration	Cumulative Cases Averted°
65	$6 \times 10^{-6} - 6 \times 10^{-3}$	<1 - 420
7.5	$3 \times 10^{-7} - 2 \times 10^{-3}$	<1 - 1,450
2.75	$1 \times 10^{-7} - 1 \times 10^{-3}$	<1 - 2,880

† Rounded to one significant figure.

* The total individual risk was determined by assuming that the risk due to inhalation is equal to that due to ingestion.

° Number of cases averted for concentrations shown in the first column.

TABLE 11

VINYL CHLORIDE-POPULATION RISK ESTIMATES
FOR CURRENT LEVELS OF DRINKING WATER EXPOSURE
(FERON-NEW)

Mean Drinking Water Concentration (Micrograms/ Liter)	Number of People Being Served	Total Lifetime Individual Risk for the Mean Concentration*		Lifetime Population Risk
		Low (Probit)	High (Weibull)	
0.25	2.1×10^{8}	$<10^{-10}$	5.4×10^{-5}	<1 - 11,300
2.75	1.3×10^{6}	$<10^{-10}$	6.1×10^{-4}	<1 - 790
7.5	4.7×10^{5}	$<10^{-10}$	1.4×10^{-3}	<1 - 660
65	1.2×10^{5}	3×10^{-6}	1.1×10^{-2}	<1 - 1,320
			Total†	<1 - 10,000

Vinyl Chloride (Feron)-Risk Reduction Analysis
For Limiting Drinking Water Concentration

Maximum Allowable Drinking Water Concentration (Micrograms/Liter)	Approximate Individual Risk Rate for Maximum Concentration	Cumulative Cases Averted°
65	1×10^{-7} - 1×10^{-2}	<1 - 1,320
7.5	8×10^{-8} - 8×10^{-3}	<1 - 1,980
2.75	$<1 \times 10^{-10}$ - 1×10^{-3}	<1 - 2,770

† Rounded to one significant figure.

* The total individual risk was determined by assuming that the risk due to inhalation is equal to that due to ingestion.

° Number of cases averted for concentrations shown in the first column.

TABLE 12

CASES AVERTED IF A STANDARD IS SET AT THE CONCENTRATION SHOWN

Contaminant	Population Risk†	Micro-grams/ Liter	Cases Averted†	Micro-grams/ Liter	Cases Averted†	Micro-grams/ Liter	Cases Averted†
Benzene	<1 - 400	2.75	<1 - 70	7.5	<1 - 10	75	<1 - 2
Carbon Tetrachloride	DATA	INSUFFICIENT	FOR	ANALYSIS			
1,2-Dichloroethane*	40 - 90,000	2.75	30 - 20,000	7.5	3 - 400	15	3 - 400
1,1-Dichloroethylene	DATA	INSUFFICIENT	FOR	ANALYSIS			
Perchloroethylene	<1 - 10	2.75	<1 - 6	7.5	<1 - 3	65	<1 - 2
Trichloroethylene*	<1 - 70,000	2.75	<1 - 20,000	7.5	<1 - 4,000	100	<1 - 200
Vinyl Chloride*							
(Maltoni-Old)	<1 - 50,000	2.75	<1 - 3,000	7.5	<1 - 1,000	65	<1 - 400
(Feron-New)	<1 - 10,000	2.75	<1 - 3,000	7.5	<1 - 2,000	65	<1 - 1,000

† Rounded off to one significant figure.

* See text.

UNCERTAINTY

Since the uncertainty was so large in these estimates it was felt to be important to closely examine its composition. Thus it is in this chapter disected to determine what the individual contributions are. In a scientific or technical endeavor such as a risk estimate that produces numbers, the most important part of the number is the uncertainty involved. The actual number is meaningless without an indication of the error or uncertainty. For example, to give an estimate as 100 cases/lifetime, it has a different meaning if it is 100 $\pm$ 1 or 100 $\pm$ 80. In the first case, the tens and units places have significance while in the second case they do not.

There are many ways to determine the uncertainty in the number, however, we will consider three here: viz, a statistical, practical, and the socalled Delphi approach. Too often these and other approaches are not used. It has been recommended that the level of concern about uncertainty should be increased (Morgan, et al., 1983).

The statistical approach requires knowing the frequency distribution involved. To determine this, many measurements must be made. When the data are sufficient and the frequency distribution becomes defined the standard deviation can be calculated and used as a measure of the error or uncertainty. This approach is the ideal one as it would allow the develop-

ment of a risk profile or actual distribution of probable risks as a function of the magnitude of the risk. The luxury of such a plethora of information is unlikely for risk estimation such as those considered in this paper. It is more likely that the data is so sparse that the frequency distribution is unknown and the statistical approach cannot be used.

The practical approach to determine the uncertainty is well known to those who have taken the laboratory in Physics I. By a practical approach is meant the determination of the contribution to the uncertainty from observation and experience with the measurement tools. If we were measuring the length of a table the uncertainty can be determined by noting the smallest measure on the measuring device.

For exposure some of the contribution to uncertainty that can be observed are: known uncertainty for the analytical method used, variation with respect to time, geography, well depth, house size and ventilation rate. By simple observation of the uncertainty the measurement tools and an approach to combining these errors in an overall uncertainty can be estimated.

A third way of estimating the uncertainty is to ask the experts--sometimes called the Delphi approach. This can be used to develop an estimate of the nature and magnitude of uncertainty. As an example of this approach a recent study

sought the judgments of experts concerning the environmental effects of sulfur dioxide (Morgan, et al., 1983). In that study the experts were asked to estimate the plume flight time of sulfur dioxide from coal fired power plants and an estimate of the health impacts. There was a small range of disagreement (or uncertainty) among the experts concerning the flight time but considerable uncertainty with regard to the health aspects.

Potential difficulties with the Delphi approach are due to the different way people make judgments in the face or uncertainty (Whittmore, 1983). It has been found that anxiety about uncertainty leads to overconfidence, that the seriousness of the risk can lead to overestimates and the method of expressing the risk can influence judgment (for example people react differently if the risk is expressed as 1 in 1000 of dying or a chance or 99.9% of living).

This discussion will use a combination of the practical and Delphi approaches to determine the uncertainty. Further, the contributions to uncertainty that cannot be quantified will be listed and discussed.

HEALTH EFFECTS UNCERTAINTY

This section is an attempt to list, define, and quantitate the uncertainty that may be an integral part of toxicological experimental design. Biology by definition is not an exact science. When an experiment is designed, a carefully reasoned question is asked. If the answer is known, then one has an 80% chance of getting the same experimental results. Rarely, and with good reason, are experiments repeated that take years to perform. Therefore, the presentation below attempts to catalogue the potential and actual uncertainty that is inherent in the design, extrapolation and interpretation of such experiments. The results of this analysis are shown in Table 13.

We will attempt to focus on, but not limit the discussion to, long-term carcinogenicity or chronic toxicity experiments. This discussion will center on the VOC's because the authors have extensive experience, access to data, and have performed extensive analysis in these areas: 1) toxicology, 2) risk estimation and 3) occurrence/exposure. Numbers, details, estimates and information that are unreferenced originate from the author's experience.

The probability of choosing the appropriate end-point is about seventy percent using a chemical from a well characterized class such as the VOC's. The chance of being wrong is larger in an unknown or poorly characterized chemical class.

TABLE 13

HEALTH EFFECTS UNCERTAINTY

	Category	Contribution to Uncertainty*
1.	Choice of endpoint	Less than 10% of the time chose the wrong endpoint--some exception like chlorinated hydrocarbon where expect liver cancer (U)
2.	Personnel capabilities	All or nothing (E)
3.	Choice of species, strain, age and sex of animals	All or nothing (E)
4.	Test compound purity (contamination, decay and vehicle contribution)	For the VOC's is generally trivial but could be two orders of magnitude (O)
5.	Inappropriate statistical test methodology	All or nothing (E)
6.	Distribution of animals among doses an number used	One or two orders of magnitude (E)
7.	Selection of dose levels	Up to two orders of magnitude (E)
8.	Lack preliminary tumor change information such as hyperplasia	For VOC's is less than 10% but could for other contaminants be as large as a factor of three (U)
9.	Experimental surroundings	Could be orders of magnitude (O)
10.	Dietary considerations	Factor of two (E)
11.	GLP's (good laboratory procedures)	One to two orders of magnitude (U)

* U-leads to an underestimate of the risk.
O-leads to an overestimate of the risk.
E-could lead to an overestimate or an underestimate of the risk.

TABLE 13 HEALTH EFFECTS UNCERTAINTY (continued)

12.	Time-to-tumor	Could miss the effect or could be an underestimate by a factor of two
13.	Microscopic tissue examination	Could be a factor of two (E)
14.	Diseases in test animals	All of nothing (E)
15.	Statistical noise	Factor of 2 (E)
16.	Outcomes of cancer (if include non-fatal cancers)	Could be two orders of magnitude or more (O)
17.	Conventional choice of p level (e.g., 0.05)	All or nothing (E)
18.	Synergism/Antagonism	Many orders of magnitude (E)
19.	No corresponding tissue in humans	All or nothing (E)
20.	Most sensitive vs. average	Several orders of magnitude (E)
21.	Animal to man	Conservatively two orders of magnitude (E)
22.	Body weight vs. surface	One order of magnitude (E)
23.	Use of upper 95% confidence limit	Up to an order of magnitude (O)
24.	Choice of dose-response model	5-6 orders of magnitude when considering risk levels in the 10^{-4} to 10^{-6}/lifetime range (E)

* U-leads to an underestimate of the risk.
O-leads to an overestimate of the risk.
E-could lead to an overestimate or an underestimate of the ris

This is illustrated by the unexpected occurence of brain cancer that was produced in animals exposed to acrylonitrile (Quast, 1980 a and b).

The possibility of choosing appropriate dose levels to produce the desired effect is dependent upon several factors. These include the number of dose levels, the regimen of dosing and the route of exposure. One can give the animals one or more doses daily. The material can be given via ingestion (in food or other vehicle) or inhalation. The inhalation exposure could be for up to 24 hours a day for seven days a week. The most common method used in inhalation, however, is 6 hours a day for 5 days a week. This latter commonly used regimen introduces an uncertainty becau e all of the extrapolations are for lifetimes of continuous exposure. There are many examples of choosing the wrong dose. The NCI bioassay program adjusts their dose requirement if the animals begin to die or exhibit morbidity at an unacceptable rate (NCI, 1978b).

The facilities in which acceptable experiments are performed have many components. Most important are the quality of personnel and the adequacy of the physical plant. An example of unacceptable data due to poor plant facilities is from Gulf South Research Institute (Moore, 1983). In IBT (U. S. vs. Kepplinger, 1983) the personnel though qualified were dishonest. In the case of GSRI (Moore, 1983) the personnel were honest but incompetent.

The experimental model (choice of species, strain, age and sex) is an integral part of the design of the experiment. NCI recommends the use of specific species and strains of rodents (Weisberger, 1983). The NCI has determined that the use of the Fisher 344 rat and the B6C3F1 mouse produce the best combination of sensitivity and low backround for use in carcinogenicity studies. This is not true in every case. It is entirely likely that for a particular chemical these animal models may be insensitive. An example of this kind of phenomena is thalidomide. Only certain strains of rabbits and some monkeys were capable of demonstrating phocomelia (Casarett, 1980).

One should select the statistical procedure before performing the study. This is to eliminate the possibility of introducing the personal bias and desire of the experimenter. Without this objectivity, the experimenter may search for the statistics that confirm the significance of the already subjectively determined result.

The experimental design may play a major role in reducing uncertainty. One can choose an unusual design such as a truncated or pyramidal model in which the lower dose levels have significantly more animals. This increases the sensitivity approximately equal to the square root of the increased number (Land, 1980 and OTA, 1981). One must know ahead the shape of the dose-response curve for this design to be successful. Therefore as the numbers of animals are increased (by squaring)

one can reduce the dose level being investigated by one order of magnitude (Land, 1980). This procedure is limited to a total of 20,000 animals because only a small number of facilities can handle such large numbers of animals. The cost is approximately 7 million dollars for an experiment using 20,000 animals. Using this large number of animals the lowest measured dose level is reduced by only one order of magnitude compared to that achieved with the normal 1500 animal experiment.

Often similar experiments yield data that appear different. For example, the data for benzene shown in Table 14 (a-d) seem to be quite different. However as shown in Table 14e the concentration predicted for the 10^{-6}/lifetime level is within a factor of approximately two for these four cases.

The maximum tolerated dose (MTD) is derived from an experiment. It is the dose that produces a 10% depression in weight gain over 90 days. The MTD is a surrogate that attempts to predict the dose level at which no more than 20% of the animals exposed will die over the duration of a lifetime experiment. The next standard level chosen is usually some fraction of the MTD--say 1/2 or 1/4. This ignores the dose-response curve. It may be wiser to choose the 5% and 1% projected response for 90 days as the second and third dose levels. If the MTD is not a sufficiently high dose level, then the true carcinogenic response will be missed. If it is too high, we will lose the entire dose group, and not have enough dose level information

TABLE 14a Incidence Rates of Zymbal Gland Carcinomas in Male Rats

Experimental Dose (mg/kg/day)	Human Equivalent	Incidence Rates
0	0	2/48
50	6.1	6/50
100	12.2	10/50
200	24.4	17/50

TABLE 14b Incidence Rates of Zymbal Gland Carcinomas in Female Rats

Experimental Dose (mg/kg/day)	Human Equivalent	Incidence Rates
0	0	0/50
25	2.5	5/50
50	5.1	5/50
100	10.1	14/49

TABLE 14c Incidence Rates of Malignant Lymphomas and Zymbal Gland Carcinomas in Female Mice

Experimental Dose (mg/kg/day)	Human Equivalent	Malignant Lymphomas Incidence Rates
0	0	4/49
25	1.43	9/48
50	2.85	9/49
100	5.7	15/48

TABLE 14d Incidence Rates of Malignant Lymphomas and Zymbal Gland Carcinomas in Female Mice

Experimental Dose (mg/kg/day)	Human Equivalent	Malignant Lymphomas Incidence Rates
0	0	15/48
25	1.43	24/45
50	2.85	24/50
100	5.7	19/49

TABLE 14e

Table Number	Concentration in Drinking Water that Produces a 10^{-6}/Lifetime Risk (micrograms/liter)
15a	1.7
15b	0.8
15c	1.0
15d	2.3

to use for extrapolation. The NCI, therefore, has made interim changes in dose levels to compensate for the incorrect selection. The NCI then reports the dose level the animals were exposed to as a time weighted average. It is difficult to determine the actual dose level that will reproduce the effect using the time weighted average.

Prior knowledge of the historical background incidence of the species and strain of animals under test is critical. Initially one would like to design an experiment in which the target organ, if known, normally has a low tumor incidence. This is because a high backround incidence of tumors in the target organ system would make a low-level carcinogenic response impossible to detect (Watanabe, 1983).

Many of the early NCI reported experiments were called into question because animals who were exposed to different volatile organic carcinogens were housed in the same room as animals exposed to other volatile organic chemicals (NCI, 1976).

The NCI has kept track of the historical control background cancer levels. They have segregated the population on the basis of different laboratories who routinely performed their experiments. Over time and many experiments later, it became obvious that certain of these laboratories had a significantly higher incidence of background cancers (Haseman,

1984). One must assume, since the NCI supplied randomly selected animals of identical species and strain, that differences in the way the animals were handled or housed were the genesis of the laboratory difference.

These differences could be the result of:

1. Noise stress
2. Laminar flow vs. HEPA filtered air
3. Purity of drinking water
4. Cage cleanliness

Dietary considerations can be thought of in two categories. One and most often talked about is trace contamination by pesticides or other organic chemicals. The NCI distributed a carefully formulated and tested chow for mice and rats. This chow was not generally available until 1979. Therefore, earlier experiments are subject to questions when their results were marginally significant. The second often ignored phenomena is that diets deficient in one or another element can significantly enhance disease processes (Revis, 1983).

There are certain housekeeping aspects to experimental procedures that may have significant effects. The first and most ignored is the randomization of cage placement of animals in their racks. There is some evidence that the location in the racks had greater significance than the experimental treatment.

Good laboratory procedures (GLP's) have mandated that each animal be given a unique identification. This is as a result of the IBT experience in which animals died and were reborn during a chronic experiment (US vs. Kepplinger).

Another source of problems addressed by GLP's was the handling of recently dead and moribund animals. Significant amounts of needed tissue were lost to gross and microscopic examination due to inadequate procedures. GLP's now require frequent enough examination to prevent cannibalism and autolysis. Lack of these procedures produced major effects on the experimental outcomes.

GLP's also specify that experienced and qualified personnel must supervise and perform autopsies. It is likely that subtle effects would be missed if only a technician alone performed necropsies. (A small footnote to necropsy procedure is included. The 'standard' is to preserve tissue in a formalin solution. If one is looking for immune or autoimmune effects, they will be destroyed by formalin preservation.)

An attempt should be made to characterize time-to-tumor phenomena. While this information does not change the outcome of a cancer experiment, it provides a good measure for potency. Unpalatable feed or drinking water can produce erroneous results. This is due to reduced nutritional intake. One must decide how often to examine and collect data from the animals.

Weight gain data should initially be collected very often and then less often as the animals mature.

Microscopic characterization of grossly observed lumps and bumps is used to determine the carcinogenic character of the lesion. Today, panels of 'experts' are assembled, duplicate, double blind, coded slides sets distributed, and a consensus is used as the final result. This does not take into account what the 'true' answer really is (NCI, 1977). It is not that these talented experts differ in the characterization of the structural architecture that they are examining. It is that a clear definition of what a carcinogenic lesion in that tissue is lacking. Benign tumors are added to any malignant tumors and used in the calculation of risk (Surgeon General, 9?) and Plant 1975). Another example is to consider curable skin cancer due to arsenic exposure as life threatening (USEPA, 1983b).

Intercurrent animal diseases are difficult problems with which to deal. Murine pneumonia can sweep through an animal colony rapidly, negating two years of expensive and carefully conducted experiments. Unfortunately, results are sometimes reported in which the infection though noted, is ignored. The results are often quoted without murine pneumonia cited. The results of experiments in which the animals were or are sick is not acceptable.

We will define the concept or a statistical noise as the

level at which one cannot determine the difference between response and the natural occurrence in the general population. This level will vary depending upon the endpoint chosen to be measured. If the endpoint is an extremely rare occurrence, e.g., angiosarcoma, mesothelioma, brain tumors, it is easily shown to be significant. The second kind of endpoint that is difficult to measure statistically is a normalized standard measure--e.g., IQ. It is impossible to measure small changes in limited populations even though they may be real. The third example would be the yes or no parameter such as life or death. These kinds of yes/no answers lead to easily characterized statistical measures.

Another source of uncertainty is the lack of well defined decision criteria. For instance 1-5 Napthlenediamine is called a carcinogen while Clonitracid and 2-(chloromethyl) pyradine were not, even though all three had significant increases in uterine tumors (Salsburg, 1983).

There is a major error in the evaluation of carcinogenicity data, that is all tumors that are cancer are considered equal. This is shown in its most unreasonable form when one considers the cancers produced by arsenic (cancer of the skin) and those produced by benzene (leukemia). Skin cancer is curable 99.99% of the time. Benzene induced leukemias are incurably fatal. The risk model as applied predicts a much higher incidence of skin cancer vs. leukemia. The CAG, there-

fore, ranks arsenic as a much more dangerous carcinogen than benzene (USEPA, 1983b).

The choice of 0.05 as a level of significance is an arbitrary one (Fisher and Yates, 1957). There is no reason that a greater p would not also show significance. In those cases in which a very rare cancer occurs as cited above, statistical significance is not required since these are rare events.

Two tissues for which there is no human equivalent are the Harderian gland and Zymbal's gland. However experience has shown that these are predictive of cancer although in some other tissue (Maltoni, 1983). Animals and humans may not show the same effect. For example benzene causes cancer of the hematopoietic system in humans and not in animals, benzidine causes bladder cancer in humans and liver cancer in animals and cadmium causes lung and prostate cancer in humans and cancer of the testes in animals (Tomatis, 1979). Cigarettes have been shown to cause lung cancer in humans and not in animals and dioxin (2,3,7,8-tetrachlorodibenzo-p-dioxin) is one of the most lethal chemicals to animals but the effect observed to date in humans is only chloroacne and some possible soft tissue cancers. It would appear that to use animal bioassay data to predict the effects on humans will be wrong for some environmental contaminants.

There is a tendency to use the most sensitive animal model

to determine the potential carcinogenicity of chemicals. This may overestimate the potential of actual carcinogenicity in humans by several orders of magnitude.

One of the major considerations in the extrapolation of data from animal to man are the differences in metabolic patterns. The animal model may not produce a proximal carcinogen and one would miss characterizing a true carcinogen. An example is arsenic (USEPA, 1983b).

Animals may produce more of the active proximal carcinogen and less of the material to deactivate or bind the carcinogen than does man (USHEW, 1978) The distribution of the chemical may be different in man vs. animals. An example would be the binding of heavy metals in plasma and in red blood cells (USEPA, 1983a).

The choice of taking the cube root or using the direct ratio of the weights of humans to experimental animals can lead to an error of up to an order of magnitude.

The different models, as shown in Figures 4-14, predict responses that are 5-6 orders of magnitude different at the 10^{-4} to 10^{-6}/lifetime risk level.

UNCERTAINTIES IN OCCURRENCE ESTIMATES AND POPULATION CONCENTRATION ESTIMATES

The assessment of the frequency of finding VOC's in drinking water supplies across the United States and the concentration being served to various portions of the American population is a major input to the risk assessment process. This is called an occurrence estimate. The assessment process involves the assembly of monitoring data from national studies in which the criteria of system selection did not alter the probability of the system being representative of the pollution in other systems within its class, thereby allowing the formation of a distribution curve which can be used to project national occurrence estimates and populations being served. The uncertainties resulting from this process begin with conditions surrounding sampling design which ultimately determines sensitivity to low frequencies of occurrence and moves through all of the technical procedures for sampling, analysis, data verification and interpretation. Each of these procedures contribute some uncertainty to the final estimates. The overall list of uncertainties in estimating exposure and an estimate of their impact on the risk analysis are shown in Table 15.

Site Selection of Systems to be Sampled

Two sources of uncertainties must be dealt with during the procedure for selecting sampling locations. These uncertainties

TABLE 15

SOURCES OF UNCERTAINTY FOR OCCURRENCE, POPULATION CONCENTRATION AND EXPOSURE ESTIMATES USED IN THE ASSESSMENT OF VOLATILE ORGANIC CHEMICALS IN DRINKING WATER

Factors	Impact on Estimate of†: Occurrence	Population Concentration	Risk
Generation of Monitoring Data			
Proportion of population sampled	5% (U)		
Representativeness of systems selected			
Geographic distribution, system size and source of water	10% (E)		
Sampling methods		50%	Factor of 2
Site of sample collection	20% (E)		
Time of sample collection	20% (E)		
Method of sample collection	10% (U)		
Container type	10% (U)		
Stability during storage	100% (U)		
Sample analysis			
% recovery from sample	10% (U)		
Compound identification	10% (E)		
Accuracy of quantitative determination	40% (E)		
Assumptions During Data Analysis			
Lower limits of quantification		Factor of 3-4	Factor of 2(E)
Oral exposure rates			
Intake rate of water			10% (E)
Pollutant level in consumed water (hot vs. cold)			50% (O)
% Absorption for oral intake			10% (O)
Respiratory exposure rates			Factor of 3
Dermal exposure			Negligible

†U = leads to an underestimation of the risk.
O = leads to an overestimation of the risk.
E = could lead to an overestimation or an underestimation of the risk.

are:

- -- Was the selection of systems to be sampled conducted using criteria which provided an unbiased representation of the national universe of community water supplies?

- -- Was the study designed so as to detect levels of pollution which may occur infrequently?

The information used in this risk assessment to project national occurrence rates was formed as an amalgam from national survey data. In some studies the sampling locations were selected randomly. In others sampling locations were selected independently by individuals within different states and regions throughout the country with the only guidance being not to select systems known to be contaminated. Prior to forming an amalgam of these data sets, data were compared with data from "random" sites, and it was found there was no statistical difference between the two data sets. Thereafter, the data sets were merged to increase the number of samples available from which to assess national occurrence.

It would appear that there is less uncertainty surrounding the body of the national frequency distribution formed from these data than exists when one examines the high and low ends of that curve. The exact number of systems experiencing an extremely high level of contamination is unknown. Information exists from state agencies which demonstrates that such places exist yet were undetected in the national monitoring survey.

One possible source of this "false negative" could be caused by the lower limit of sensitivity created by randomly selecting 1,000 sampling locations from a population of 48,000 candidates.

The number of systems experiencing these high levels and the number of people they serve is probably quite small and is expected to have a minor impact on population concentration estimates. However if this distribution of systems having elevated concentration was affected by some factor causing a "clustering" of systems within a category not sampled during the survey the underestimate could be larger.

Uncertainty is also created because of our inability to detect true concentration, especially at the low end of the frequency distribution. For this analysis it has been assumed that places where the pollutant was not detected at a "minimum quantifiable level" (<0.5 micrograms/liter) contained 0.25 micrograms/liter (the median concentration between 0 and 0.5 micrograms/ liter).

We know that certain solvents are ubiquitous in societal use and appear at the micrograms/liter level in rain water. For these substances there are probably few if any, water supplies which are absolutely free of contamination. However, for other less ubiquitous chemicals, a number of water supplies are probably absolutely clean. By far the largest number of people fall into the "less than" category. The percent of

risk associated with this "less than" category representing the maximum overestimation, if all systems were truly zero, is presented below for each of the chemicals in this assessment:

	Percent in the "less than" category	Ration of total to fraction above the "less than" category
Benzene	78%	4.5
Carbon Tetrachloride	35%	1.5
1,2 Dichloroethane	60%	2.5
1,1 Dichloroethylene	90%	10
Tetrachloroethylene	51%	2.0
Trichloroethylene	29%	1.4
Vinyl Chloride	79%	4.7
	Rough average	3.5

Sampling Location Within the System

The data available from national monitoring studies were generated from a single sample taken at one point in time at one location within the distribution system of each water supply.

How representative is data from a grab sample of the drinking water quality being provided to users throughout the system?

There are several components to this question. How do

concentrations vary with time? Is the pollutant level created by stable long-term contamination of raw water? If so, then the data could be used to produce fairly representative estimates of exposure. However, if the pollutant level can effect a portion of the system or may change drastically over time, the estimates will be highly uncertain.

There is no information to help answer these questions. We know that some groundwater systems operate so that individual wells serve specific portions of a city. We also know that some chemicals may be added to specific portions of a system because of leaching from pipe linings (e.g., tetrachloroethylene). Under these conditions the exact sampling point within a water system will make a large difference in the finding. For others, such as carbon tetrachloride, it is felt that most of the problems arise from contamination of raw water or the use of contaminated chlorine disinfectant which would affect the water supply rather uniformly. Since this effect of this source of uncertainty cannot be well defined, it has been assumed that while results for specific systems may be overestimated, unbiased selection throughout the systems sampled has likewise produced an underestimate of conditions in other systems which when combined, produce a reasonably accurate national occurrence profile.

Maintaining Sample Integrity Prior To Analysis

The problem of partial or complete disappearance of the chemical pollutant contained in drinking water prior to analysis due to physical or biological processes may produce a large uncertainty in the analytical results. When partial disappearance is suspected, the concentrations represented in data may be taken as "minimum values" from an unknown but larger level present at the time of sampling. A problem arises in interpreting data reported as "zero or minimum detectable" found in such a data set for the possibility exist that higher levels were originally present which diminished in sample handling produce false negative readings at the time of analysis. Under these conditions, it would be impossible to identify low occurrences from false negatives. Preliminary data reported by Glick in 1979 demonstrated a rapid loss of benzene, ethybenzene and toluene from unutilized spiked tap water to background in about 10 days. Chlorobenzene, m-xylene and p-dichlorobenzene levels were found to diminish in 2-4 weeks of storage. The loss of benzene was especially rapid with 50% loss being found within 2 days. The situation with respect to the degradation of unsaturated aliphatic hydrocarbons is not well understood.

What level of uncertainty has been created for chemicals discussed in this report by loss during the storage of samples prior to analysis? To address this question one must realize that the only data used to project national occurrence were

those developed from water samples drawn after treatment in the distribution system. Most of these samples, especially those from systems using surface water, would have contained chlorine disinfectant. Data on the level of VOC's in surface water supplies was developed in studies predating the recognition of possible degradation problems in which holding times were long (see Table 16). Therefore, there exists a level of uncertainty surrounding surface water contamination with each of the VOC's, but especially benzene. Data from which groundwater predications have been made are also samples drawn from the distribution system. However, since a significant number of the smaller groundwater systems do not disinfect their water, it is indeed fortunate that data were derived from a more recent study (GWSS) which was designed to ensure minimal pollutant loss during sample storage.

The uncertainty for risk assessment on these VOC's is greater for benzene where occurrence predictions may underestimate national occurrence because of numerous false negatives in surface water supplies and lesser for each of the other VOC's.

Detection and Quantification

The process of chemical analysis is filled with technical uncertainties. Procedures have been established to control the quality of the resultant data. Confirmation via alternate

TABLE 16

SAMPLE STORAGE TIME AND CONDITIONS

Study	Storage Conditions	Storage Time Before Analysis
NORS[a]	Iced	≤ 14 days
NOMS I[b]	Iced	7-14 days
NOMS II, III	Room temperature	21-42 days
NSP[c]		14-28 days
CWSS[d]	Some not refrigerated	150-800 days
RWSS[e]	Some not refrigerated	300-700 days
GWSS[f]	Acidified/bactericide added	≤ 28 days

[a]NORS - National Organics Reconnaissance Survey
[b]NOMS - National Organics Monitoring Survey
[c]NSP - National Screening Program
[d]CWSS - Community Water Supply Survey
[e]RWSS - Rural Water Supply Survey
[f]GWSS - Ground Water Supply Survey

analytical techniques (i.e., mass speculation) of a percent of the qualitative findings has been practiced on the data used on these surveys. Therefore, this source of uncertainty is relatively low. In addition, quantitative determinations may become a large surface of uncertainty. Once again, laboratory procedures were instituted for these studies to assure that the error was minimized (Taylor, 1982).

Uncertainties Exposure Assessment

The drinking water contribution to respiratory and oral exposure has been estimated using assumptions about the rate of water used during fluid intakes--bathing, the transfer rate to air of VOC's transported in the drink ng water, the turnover rate for air within the closed confines of a bathroom and house and the length of time with a corresponding volume of air inhaled within each environment.

Definitive data are not available to define the variance or average of any of these factors as they arfect the United States population. The assumptions made are based on "reasonable" extensions data confirmatory of a qualitative process but insufficiently precise to validate quantitative conclusions.

Oral Exposure

Assumption--Intake Rates

Procedures used in this risk assessment assumed that average intake of drinking water over a lifetime approximates 2 liters/ 70 kg rate established for "average adults" resulting in 0.028 micrograms/liter/kg absorbed dose per day. This consumption rate (0.03 liters/kg/day) has been selected as the population mean for this report. Comparison with theoretical average rates reveals that fluid intakes per unit body weight generally decrease with age, beginning at 7-9 times as high in infants and reaching a plateau by early teens and adulthood. A weighted average derived from a theoretical water consumption curve indicates that the lifetime average may be 20% higher than the current factor being used. Population studies on the variability or water consumption have not been conducted in the United States. Canadian data indicate that 13-16% of the adults may consume more than 2 liters of water per day.

Assumption - Concentration of Pollutant in Fluids Consumed Approximates the Concentration in the Tap Water Sampled.

This conservative assumption implies that individuals are receiving that static concentration in drinking water over their lifetime. It does not deal with population mobility nor potential changes in concentration with time. The vast

majority of the population does not remain in one city and is not served by a single water supply over their lifetime. However, the rate of movement from locations with higher pollutant concentrations and vice versa is unknown. The same lack of knowledge exists for the stability of drinking water concentrations over time, and therefore the uncertainty associated with these factors cannot be quantified.

If has been assumed that what is present in cold tap water is present in consumed fluids made with drinking water. We know from the Canadian study that the preponderance of drinking water consumed by children is in the form of cold drinks and the amount of hot drinks increases with increasing age. We also know that while the VOC's are transported in drinking water they are rapidly transferred into air with aeration, heating or upon standing. These factors would diminish oral exposure of adults significantly but tend to create point source liberations into indoor air which might increase respiratory intake. Since the percent of absorption from respiratory intake is assumed to be 50% of the oral absorption rate the overall exposure of adults might be reduced by 30-50%.

Respiratory Exposure

This risk analysis recognized the transference of drinking water borne pollutants into indoor atmosphere and the resultant exposure of household members to an enriched pollutant

atmosphere during time spent in the dwelling. It assumes that respiratory uptake is equivalent to oral uptake with the major contribution to exposure taking place during showering or bathing and the time spent within the enclosed bathroom immediately after these activities prior to dispersal of the pollutant concentration. This conclusion is based on several assumptions and their associated uncertainty:

-- Nearly 100% of the pollutant concentration of VOC's in drinking water is transferred to air upon heating, aeration or standing.

Laboratory and field testing data is available to support this assumption. Engineering documents produced for the USEPA by Environmental Sciences and Engineering of Gainesville, Florida, document a high efficiency of aeration processes for removing the VOC's from contaminated raw water. In addition, laboratory tests show that hot water has lost virtually 100% of its VOC contamination upon standing for 3 hours or being poured between containers 20 times. Cold water exhibits 20-80% loss under the same conditions. Since bathing and showers involve a large percent of heated water which is aerated during impact on body or tub, the volatility is enhanced. An assumption of 100% transference to air may be in excess by 5-10%.

Exposure During/After Showering and Bathing

Physical factors used to define the volume of air in a

bathroom into which the pollutant would be dispersed were chosen instinctively. The volume of air assumed to be present in the bathroom was 9.4 m^3 or 6' x 7' x 8' --the 30 gallons of water assumed to be used in a shower. The result of 100% transfer of the water pollutant was 113 micrograms of pollutant present at 1 microgram/liter in water for 9.4 m^3 of air or a resulting concentration of 1.2 micrograms/m^3. It was further assumed that the individual spending 30 minutes in the bathroom inhaled 0.42 m^3 of air containing 0.5 micrograms of pollutant of which 50% was absorbed. The net absorbed dose was estimated to be 0.0036 micrograms/kg/day.

The relationship between the volume of air contained in the bathroom, including turnover rate within a 30 minute period and "national averages" is unknown. Individuals in saunas, steambaths or hot tubs would most probably exceed these time estimates.

Volume of Air Inhaled Within the Confined Area of the Showering Room

The estimated air inhaled while in the confined area of a bathing room was .006 m^3/kg of body weight. These figures were developed as averages from an assumed respiratory rate of 20 m^3 per day per 70 kg individual. The uncertainty surrounding this factor is unknown.

Concentration in Confined Air

It was assumed that the pollutant equally dispersed throughout the air within the room remained in air throughout the 30 minute time period.

There is some data from field studies in New Jersey to indicate that pollutant concentration in a shower stall is greatest at the breathing zone. However, there are no data showing that the pollutant is removed from the air by binding to fabric, wood or plastics in the room. Since the scenario assumes 100% availability, it is likely to tend a worse case estimate. However, the assumed rate of absorption (50%) was an average from animal tests (30-80%) conducted at higher air levels.

Respiratory absorption at lower concentrations is probably increased substantially creating an underestimate by 10-40%.

General Household Respiratory Exposure

An average of 250 gallons of water pass through an American household each day At a pollutant concentration of 1 microgram/liter, 946 micrograms of a pollutant become available for dispersal within the household. Knowing that the VOC's are easily transferred from the water column into air, we also know that the bulk of VOC's in drinking water get

dispersed throughout the house. There are no measurements of the indoor air concentration resulting from this process. By analogy, in estimates made for radon it was assumed that the level in household air (in liters) would be equivalent to 10^{-4} of the level found in water (liters). At an initial concentration of 1 microgram/liter in water concentration in general household air would be .0001 microgram/liter (.001 microgram/liter equals .1 microgram/m^3). Individuals spending 12 hours per day within the home environment inhaling 10 m^3 of air would be exposed to 1.0 micrograms of pollutant of which 50% would be absorbed. The net absorbed dose is equivalent to .07 micrograms/ kg day.

There is no way to estimate the uncertainty associated with this estimate except to state that a comparison of the relative volatility of radon and the VOC's would indicate that the transfer of radon from water to air probably occurs more rapidly and therefore more completely in the home than would be expected for the VOC's.

Dermal Absorption

There is nothing known about the rate of dermal absorption for low levels of organics found in drinking water. It could be assumed that the VOC's pass the dermal layer in proportion to the flux of water (0.2-0.5 mg/cm^2/hr). Several factors might mitigate against this assumption:

-- Much of the VOC's in warm, hot or aerated water would be transferred to air.

-- The VOC's remaining in solution would pass through the skin at an enhanced rate due to their lipid solubility and enhanced blood flow to skin in warm water.

-- Soaps added to wash water would bind with the VOC's remaining in solution.

Since reality is unknown, dermal absorption was assumed to be proportional to water flux ranging from 2×10^{-6} to 9×10^{-7} micrograms/kilograms per day, per part, per billion of the pollutants present in drinking (micrograms/ liter).

Of all the uncertainties in the risk estimation process found here the largest is due to the choice of model. As shown in Table 15, this can be as large as 5 to 6 orders of magnitude. Since all of the other errors, combined even when in the maximum possible way, are not nearly this large, the overall error can be considered to be that due to the choice of model.

If a particular model can be chosen, the remaining uncertainty in the risk estimates is 2-3 orders of magnitude. The contribution to the overall uncertainty in either case from the exposure estimate is less than one order of magnitude.

DISCUSSION AND CONCLUSION

The number of people likely to be affected by a particular contaminant in a lifetime due to existing concentrations is called the population risk. To estimate the population risk we need to know how many people are exposed and at what concentration level and what the risk is at those concentrations. Both the exposure and toxicity information involve uncertainties. Some of these uncertainties can be estimated and some cannot.

Some, but not necessarily all, of the uncertainties involved are those due to: inherent measurement error, unrepresentativeness of estimating concentration in all supplies from a data for only a few, lack of data for concentrations below the limit of detection, estimating inhalation toxic effects from data derived by ingestion experiments and vice versa (also errors involved in estimating effects due to dermal exposure), using animal data to predict human effects and extrapolating the expected risk level several orders of magnitude below the actual dose levels used in the experiments.

The largest uncertainty, at the present, is in the extrapolation to low doses. Figure 16 shows such an extrapolation. The starred points are the actual data points for this fictitious contaminant (the figure is typical of those for most of the volatile organic chemicals found in drinking water). It

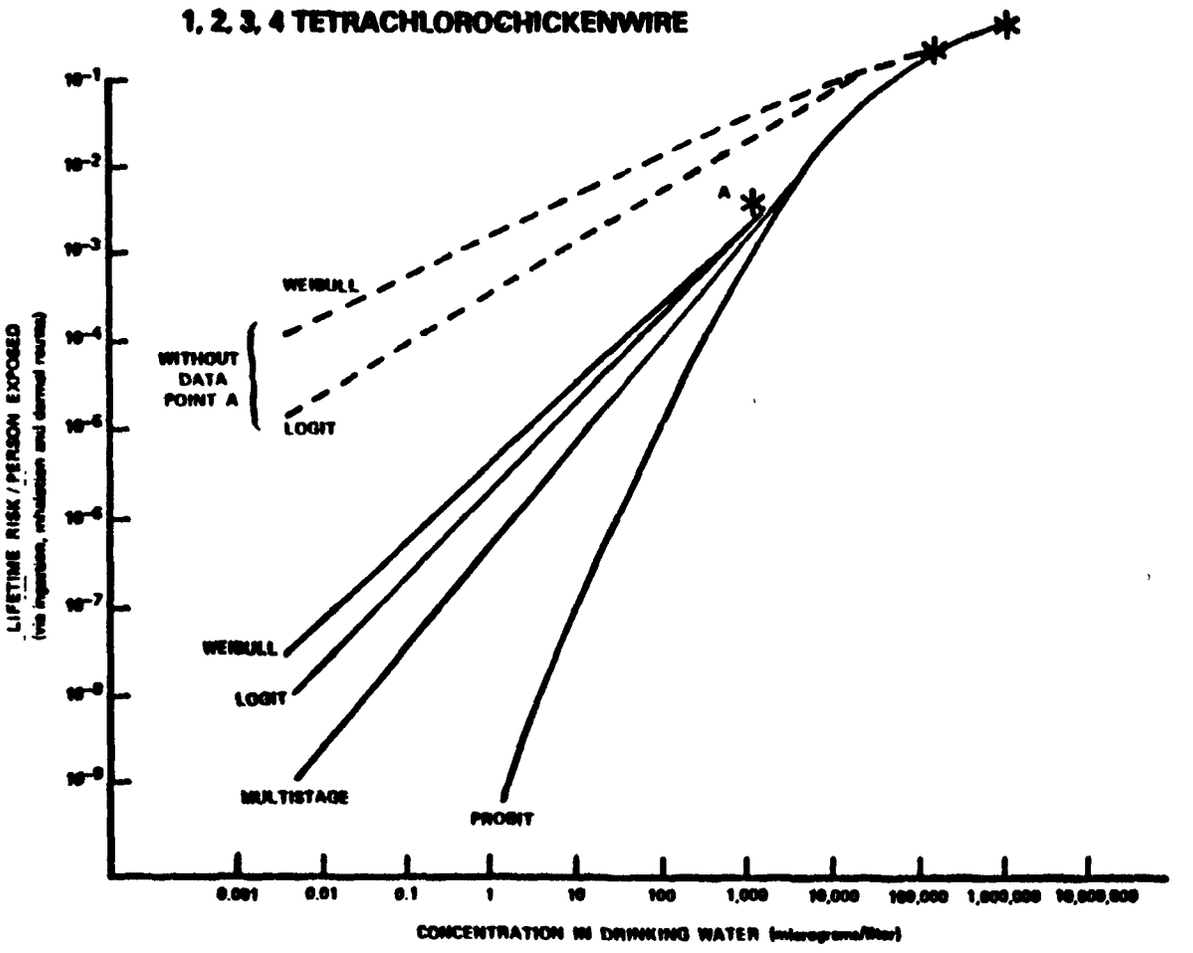

Figure 16. Hypothetical dose response curve demonstrating the effect on the extrapolations of an experimental point added at 'A'.

only the top two data points (along with a point at the origin which cannot be shown in the figure) are used, they can be fit to four of the currently popular models; viz, logit, multistage, probit and Weibull. If the point at 'A' in Figure 16 is excluded, the model fits are shown by the dashed lines for the Weibull and logit models. In that case the multistage and probit models predict curves close to the solid lines shown. If another data point is added (denoted 'A'), the multistage and probit estimates change very little but the Weibull and logit model predictions move down to the solid lines.

At a drinking water concentration of one microgram/liter in Figure 16, the range of risk prediction is about seven orders of magnitude (or a factor of 10,000,000) when point 'A' is excluded, while adding the point at 'A' reduces this range to a factor of about four orders of magnitude (or a factor of 10,000). It is possible to develop another model which could predict a risk value outside the range predicted by these four models. The probability of this cannot be estimated unless the frequency distribution of the data is known. In order to determine the frequency distribution, experimental points are needed at the low dose and low risk levels of interest and even at lower values. It is unlikely that this information will ever become available.

If we have no information at all about a contaminant, we

can predict with complete certainty that the population risk lies somewhere between less than one person and all the people (approximately 250 x 10^6 people in the U. S.). The object of risk estimation is to reduce this range. Suppose we consider Figure 16. Assume that about 5% of the U. S. was exposed to this hypothetical contaminant at a relatively low concentration of one microgram/liter. Then the range of population risk would be approximately:

10^{-3} cases/person exposed x 5% of 250 x 10^6 persons
= 10,000 cases

10^{-9} cases/person exposed x 5% of 250 x 10^6 persons
= 0.001 cases

when rounded off to one significant figure.

Does the above range have any meaning? Is it of any value? What does 0.001 cases mean? It would make more sense to express this as less than one case. Is this range of less than one to 10,000 different from another of say 100 to 10,000,000? Since the ranges overlap, the two contaminants could not be said to be necessarily different- larger or smaller.

There is little experience in extrapolating dose-response curves below the known data. In the area of health effects of ionizing radiation, effects have been seen down to the region

of one to ten rads while the region of natural background and that of regulatory interest is in the area of 0.1 rad (Land, 1980). The dose-response curve for alpha emitters appears to be linear down to the lowest known effect level and thus to get to the 0.1 rad level is an extrapolation of one or two orders of magnitude. The available data and model fits for health effects of ionizing radiation are found in the dose-response analysis of leukemia incidence from the atomic bomb survivors (Land, 1980). As shown in Figure 17, the data exist down to the 10^{-5}/lifetime. Since this is the region of interest for regulatory purposes, estimates of health effects are based on extrapolations of no more than an order of magnitude.

The megamouse experiment involved about 30,000 mice (exposed to 2acetylaminofluorine or 2-AAF) which is considerably more mice than is normally used in a bioassay study (Cairns, 1980). Although this large number was not enough to add a point as far down as that of 'A' in Figure 17, it did show that the dose-response curve looked linear for the late appearing liver neoplasm. Using this large number of rodents only extended knowledge one to two orders of magnitude into the unknown region. In addition a new and unexpected effect was revealed; viz, bladder neoplasm, which exhibited a minimal effect which could be interpreted as a threshold effect. Thus there can be surprises when the effect of low doses is investigated.

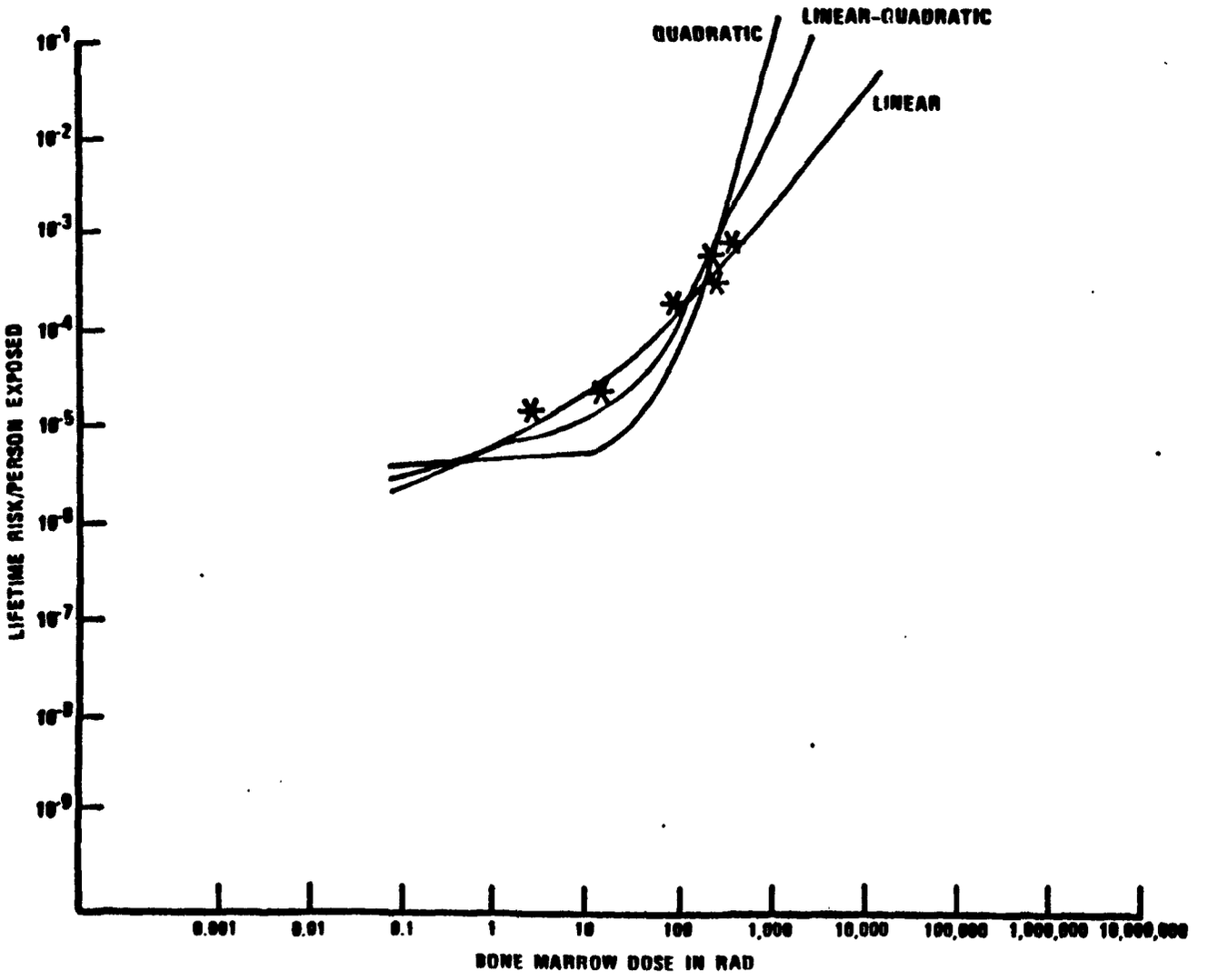

Figure 17 Dose response curves for leukemia from atomic

For the volatile organic chemicals as potential drinking water contaminants, the low dose region of interest is four to six orders of magnitude below the dose levels used in bioassay studies. There is no previous experience in extrapolating the dose-response curve so far into the unknown. Such extrapolations will involve considerable uncertainty as they go far beyond any previous experience in this field.

As can be seen from the extrapolation of the dose-response data in Figures 3-14, no one model is consistently the highest or lowest. This is in conflict with the commonly held idea that the multistage model predicts the highest risk or that the linear model predicts the highest risk (Moghissi, 1983).

A strong point of estimating risk as described here is that it does produce a number. But the weakness of the procedure is that it produces a large uncertainty. Thus, the number could be anywhere in the range and could be outside the range. If the range is 1 to 1,000, then 2 is the same number as 998 and we have the same confidence that either is correct. The magnitude of the uncertainty can be estimated by the range of model predictions shown in Figure 17, but the real value could lie outside that range.

The actual number can be anywhere in the indicated range. For the previous example of the range from 0.001 cases of 10,000,000 cases, the estimate of the lower end is equally as

possible as that of the upper end, and the probability of 0.002 cases is the same as the probability of 5,000 cases and so forth.

The magnitude of the range of uncertainty in this example can be compared to hitting a home run in a ball park. The idea of a ball park is that the ball can be anywhere in the park and the probability is the same for right field as left field and so forth. For hitting a home run in a major league park the uncertainty in its location is the order of several hundred feet. That can be compared to the uncertainty in the animal bioassay data. The corresponding uncertainty in the extrapolation is larger than that of a satellite position after launch if we are ignorant of the launch time and it could be anywhere in its orbit.

Although there are many interpretations that can be made of the calculations in this report, only a few of the major conclusions will be listed here. The report is an attempt to clearly show how the calculations of risk were made. No policy implications or discussion of the impact or connection to the overall regulatory process is included here. This is not meant to be a comment on the policy implications. The other contributions to overall regulatory decision making are to be found elsewhere.

The major observations that can be made from the calcula-

tions described in this report are:

-- The largest uncertainty in the calculations is due to the choice of model and is 5 to 6 orders of magnitude.

-- No single model consistently is the highest or the lowest on the dose-response plots.

-- If a single model were chosen, the overall uncertainty in risk estimates would be 2 to 3 orders of magnitude.

-- The exposure estimates contribute at most an order of magnitude to the uncertainty

-- It would appear that until a particular compound's mechanisms of cancer are better known, the uncertainty in the toxicity will not likely be improved.

-- If exposures greater than 7.5 micrograms/liter were eliminated, the number of lifetime cases averted would be less than 10 for benzene and perchloroethylene and less than 400 for 1,2-dichloroethane, less than 2,000 for vinyl chloride and less than 4,000 for trichloroethylene. Except for 1,2-dichloroethylene (which is 3), the lower limits are all less than one.

Some research effort could help to reduce the uncertainty in the risk estimates presented here. Clearly any information that allows only one model would be the biggest contribution to reducing the uncertainty Thus any research effort that sheds light on the mechanisms of cancer would be the most helpful. To determine what other research might be helpful, it is instructive to refer to the table of uncertainties, especially Table 15 for the uncertainties associated with toxicity. The next largest uncertainties involve the distribution of animals among doses, the doses selected and the use of averages rather than considering the most sensitive group. In this latter area some research regarding the importance of

species, age and sex would be useful. Also the question of the consequences of using homogenous, inbred strains could be better defined by some research efforts.

ACKNOWLEDGEMENTS

The authors wish to thank Mr. Charles V. Berry for help in using the computer programs to analyze the data, Mr. Robert W Bargh for help in developing a program to analyze uncertainty in the occurrence data and Dr. Steven P. Bayard (of the Office of Research and Development, USEPA) for many helpful discussions. Also thanks to Dr. Fred Abramson (George Washington University), Dr. John Van Ryzin (Columbia University) and Dr. Jacqueline Michel (Research Planning Institute, Columbia, SC) for their critical reading of the report and their many helpful suggestions for improving the presentation of this information.

REFERENCES

Anderson, E. L., 1982. "Quantitative Methods in Use in the United States to Assess Cancer Risk", Paper presented at the Work-Chemicals at Rome, Italy, July 12, 1982. For more information contact OHEA, USEPA, Washington, D.C 20460.

Bailey, J. R., Beniot, R. J., Dodson, J. L., Robb, J. M. and Wallman, H., 1969. A Study of Flow Reduction and Treatment of Waste Water from Households prepared by General Dynamics, for the Federal Water Quality Administration, Department of the Interior under Contract No. 14-12-428, Cincinnati, Ohio, Water Pollution Control Research Series, 1105UFKE.

Bayard, Steven, 1983. Private Communication, Cancer Assessment Group, USEPA, Washington, D.C.

Bishop, B., 1982. In home evaluation of the impact of benzene in drinking water and indoor air, New Jersey Department of Environmental Protection, Trenton, NJ, Personal Communication.

Boland, P. A. National Screening Program for Organics in Drinking Water, prepared by SRI International for Office of Drinking Water, USEPA, Washington, D.C., Contract No. 68-01-4666.

Brass, H. J., 1977. The National Organic Monitoring Survey: A Sampling and Analysis for Purgeable Organic Compounds in: Drinking Water Quality Enhancement Through Source Protection. (Pojasek, R.B., ed.) Ann Arbor, MI, Ann Arbor Science, pp. 393-416.

Cairns, T., 1980. The ED_{01} Study: Introduction, Objectives and Experimental Design, Journal of Environmental Pathology and Toxicology, 3:1-8, and succeeding articles.

Canadian Environmental Health Directorate, 1981. Tap Water Consumption in Canada, Ministry of National Health and Welfare, Publication number 82EHD80.

Casarett, L. J. and Doull, J., 1980. Toxicology, Macmillan Publishing Co., NY.

Coniglio, Wm., Miller, K., Mackeever, D. The Occurrence of Volatile Organics in Drinking Water--a briefing paper for USEPA. Deputy Assistant Administrator for Drinking Water, March 1980.

Crump, K., 1982. Science Research Systems, Rustin, Louisiana.

Dambrauskas and Cornish, 1970. Toxicol. Appl. Pharacol., 17:83-97.

Doull, J., 1984. Draft report, Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation, The University of Kansas, Department of Pharmacology Toxicology and Therapeutics, Kansas City, Kansas 66103.

Dow Chemical Company 1980. Burek, J. D., Nitschke, K. D., Bell, T. J., Wackerle, D. L., Childs, R. C., Beyer, J. E., Dittenber, D. A., Rampy, L. W., McKenna, M J., 1980. Methylene Chloride: A Two-Year Inhalation Toxicity Oncogenicity Study in Rats and Hampsters, Toxicology Research Laboratory Health and Environmental Sciences, Dow Chemical Company, Midland, MI, December 31, 1980. Unpublished.

Feron, V. J., Hendrikson, C. F. M., Speek, A. J., Til, H. P. and Spit, B. J., 1981. Lifespan Oral Toxicity Study of Vinyl Chloride in Rats. Fd. Cosmet. Toxicol. 19:317-331.

Fisher, R A. and Yates, F., 1957. Statistical Tables, 5th Edition, Tables VI, VI_1 and VI_2, Oliver and Boyd, Edinburgh.

FRDS, 1983. Federal Reporting Data System, Facilities and population served by primary water supply source (FRDS07) April 19, USEPA, Washington, D.C. 20460.

Gehring, P. J. and Blay, G. E., 1977. Mechanisms of Carcinogenesis: DoseResponse, J. of Environmental Pathology and Toxicology, 1:163-179.

Gehring, P. J., Watanabe, P. G. and Park, C N., 1978. Toxicol. Appl. Pharmacol., 44:581-591.

Gesell, T F. and Pritchard, H. M , 1980. The Contribution of Radon in Tap Water to Indoor Radon Concentration, Prepared for the Third International Symposium on the Natural Radiation Environment, Houston, Texas, April 23-28. Published in NTIS, as Natural Radiation Environment III.

Haseman, J., Huff, J. A. and Borman, G. A., 1984. Historical tumor incidence in the B6C3F1 mouse and F 344 rat, CRN Press, See, G. A. Borman, National Cancer Institute, Box 27709 Research Triangle Park, NC 27709.

Hess, C. T., Weiffenbach, C. V. and Norton, S. A., 1982. Variations of Airborne and Waterborne Rn-222 in Homes in Maine, Environ. Inter. 8:59.

JRB Associates, 1982. Evaluation of Drinking Water Data in EPA Studies, Draft report, Prepared for the Office of Drinking Water, USEPA, Washington, D.C. 20460, Contract #68-01-6185.

Kahlos H. and Asikainen, M., 1980. Internal Radiation Doses from Radioactivity of Drinking Water of Finland, Health Physics, 39:106.

Krewski, D., Koron, J. and Bickin, M , 1982. Optimal Experimental Designs for Low Dose Extrapolating to Appear in Topics in Applied Statistics. (T. W. Dwivedi, ed.), Marcell Dekker, New York.

Krewski, D., 1982a. Health Protection, Board of Health and Welfare, Canada, Ottawa, Ontario, Canada.

Krewski, D. and Van Ryzin, J., 1981. Dose-Response Models for Quantal Response Toxicity Data, In Statistics and Related Topics, Csorgo, M., Dowson, D. A., Rao, J. N. K. and Saleh, A K. Md. E., eds. North Holand Publishing Company, New York.

Kuzmack, A. M., 1983. Characterization of the Water Supply Industry (FY82) Memorandum of May 16, 1983, Office of Water, USEPA, Washington, D.C. 20460.

Land, Charles, 1980. Estimating Cancer Risks From Low Doses of Ionizing Radiation, Science, 209:1197.

Letkiewicz, F., Johnson, P., Macaluso, C., Elder, R., Yu, W., and Bason, C., 1983. Series of documents on occurrence of 1,1-dichloroethylene, 1,2-dichloroethane, dichloromethane, vinyl chloride, carbon tetrachloride, trichloroethylene, tetrachloroethylene and benzene, prepared by JRB Associates for the Office of Drinking Water, USEPA.

Lenaweaver, F. P., Jr. "Report on Phase One, Residential Water House Research Project", The Johns Hopkins University Department of Sanitary Engineering. October, 1963. Reported in "A Study of Flow Reduction and Treatment of Wastewater from Households", Bailey, J. R., General Dynamics Report for the Federal Water Quality Administration, 1969.

Mabey, W. R., 1981. Aquatic Fate Process Data for Organic Priority Pollutants. USEPA, OWRS Washington, D.C. EPA-440/4-81-014.

Morgan, M. G., Morris, S. C., Henrion, H., Amoral, D. A. L. and Rish, W. R., 1983. Technical Uncertainty in Quantitative Policy Analysis--a Sulfur Air Pollution Example, Manuscript submitted to Science from the Department of Engineering and Public Policy, Carnegie-Mellon University, Pittsburgh, PA 15213.

Munro, I. C. and Krewski, D. R. 1981. Risk Assessment and Regulatory Decision Making, Fd. Cosmet. Toxicol. 19:549-560.

Maltoni, C., Conti, B., Cotti, G., Mandrioli A., Scarmato, C., Valgimigli, L., 1982. Benzene: An experimental multipotential carcinogen, Institute of Oncology Bologna and Bentivogoio, Italy.

Maltoni, C., Lefimine, G., 1975. Carcinogenicity Bioassays of Vinyl Chloride: Current Results, Annals of the New York Academy of Sciences 246:195-218.

Maltoni, C., Conti, B. and Cotti, G., 1983. Benzene: A Multipotential Carcinogen, Results of Long-Term Bioassays Performed at the Bologna Institute of Oncology, Am. J. of Industrial Medicine, 4:589-6301.

Moore, J., 1983. Memorandum, Withdrawal of Gulf South Research Performed NCI Studies, USEPA, Office of Pesticides and Toxic Substances, TS-788, Washington, D.C. 20460.

National Academy of Sciences, 1982. Drinking Water and Health, Washington, D.C.

NCI (National Cancer Institute), 1976. Report on Carcinogenesis Bioassay of Chloroform, National Toxicology Program, Box 12233, Research Triangle Park, NC 27709.

NCI (National Cancer Institute), 1976a. Report on the Carcinogenic Assay of Chloroform, Carcinogenesis Program, Division of the Cancer Cause and Prevention, Technical Information Section, Carcinogenesis Testing Program, National Toxicology Program, Landow Building, Room A306, Bethesda, MD 20205. March 1, 1976.

NCI (National Cancer Institute), 1976b. Carcinogenesis Bioassay of Trichloroethylene, National Cancer Institute Carcinogenesis Technical Report Series No. 2. NCI-CG-TR-2, Technical Information Section, Carcinogenesis Testing Program, National Toxicology Program, Landow Building, Room A306, Bethesda MD 20205. February, 1976.

NCI (National Cancer Institute), 1977a. Bioassay of 1,1,1-Trichloroethane for Possible Carcinogenicity, National Cancer Institute Carcinogenesis Technical Report Series No. 3. NCICG-TR-3, January, 1977, Technical Information Section, Carcinogenesis Testing Program, National Toxicology Program, Landow Building, Room A306, Bethesda, MD 20205.

NCI (National Cancer Institute), 1977b. Bioassay of Tetrachloroethylene for Possible Carcinogenicity, DHEW Publication No. 77-813, Technical Information Section, Carcinogenesis Testing Program, National Toxicology Program, Landow Building, Room A306, Bethesda, MD 20205.

NCI (National Cancer Institute), 1977c. Bioassays of Nitriloacetic Acid, Nitriloacetic Acid (NTA) and Nitrilotriacetic Acid, Trisodium Salt Monohydrate ($Na_3NTA\ H_2O$) for Possible Carcinogenicity, Technical Report Series, Number 6, U. S. Department of Health, Education and Welfare, Bethesda, MD 20014.

NCI (National Cancer Institute), 1978a Bioassay of Ethylenedichloride for Possible Carcinogenicity, Technical Information Section, Carcinogenesis Testing Program, National Toxicology Program, Landow Building, Room A306, Bethesda, MD 20205.

NCI (National Cancer Institute), 1978b. Bioassay of Picloram, Technical Report Series No. 23, U. S. Department of Health, Education and Welfare, National Institutes of Health, Bethesda, MD 20014. See also, reports numbered: 2, 3, 5, 42, 43, and 126.

NCI (National Cancer Institute), 1983. Study of Bioassay of 1,1,1-Trichloroethane for Possible Carcinogenicity, Technical Information Section, Carcinogenesis Testing Program, National Toxicology Program, Landow Building, Room A306, Bethesda, MD 20205.

Nutrition Foundation, 1983. Report of the Ad Hoc Working Group on Oil/Gavage in Toxicology, 888 Seventeenth Street, NW, Washington, D.C. 20006.

Office of Technology Assessment (OTA), 1981. Assessment of Technologies for Determining Cancer Risks From the Environment, U. S. Government Printing Office, Washington, D.C. 20402.

Plant, P. A., 1975. A Dangerous Definition, Chem., Engineering News, Oct. 20, 1975.

Quast, J. F., Wade, C. E., Humiston, C. G., Carreon, R. M., Herman, E. A., Park, C. N. and Schwetz, B. A., 1980. A Two-Year Toxicity and Oncology Study with Acrylonitrile Incorporated in the Drinking Water of Rats, Toxicology Research Laboratory, Dow Chemical USA, Midland, MI 48640. Jan. 22, 1980.

Quast, J. F., Schuetz D. J., Balmer, M. F., Gushow T. S., Park, C. N. and McKenna, M. J., 1980. A Two-Year Toxicity and Oncology Study with Acrylonitrile Following Inhalation Exposure of Rats, Toxicology Research Laboratory, Dow Chemical USA, Midland, MI 48640. Dec. 9, 1980.

Rai, K. and Van Ryzin, J., 1979. Risk Assessment of Toxic Environmental Substances Using a Generalized Multi-Hit Dose-Respose Model, in Energy and Health, Breslow, N. E. and Whittemore, A. S. eds. SIAM Press, Philadelphia, PA

Rai, K. and Van Ryzin, J., 1981. A Generalized Multihit Dose-Response Model for Low-Dose Extrapolation, Biometrics, 37:341-352.

Rampy, L. W., Quast, J. F., Balmer, M. F., Leong, B. K. J., Gehring, P. J., 1978. Results of a Long-Term Inhalation Toxicity Study on Rats of a Perchloroethylene (Tetrachloroethylene) Formulation, Toxicology Research Laboratories, Dow Chemical Company Midland, MI.

Revis, N. W., Douglas, B. H., McCauley, P. T., Witschi, H. P. and Bull, R., 1983. The Relationship of Drinking Water Chlorine to Coronary Atheroschlerosis, Pharmacologist, 25:Abstract 732.

Salsburg, D., 1977. Use of Statistics When Examining Lifetime Studies in Rodents to Detect Carcinogenicity, J. of Tox. and Env. Health, 3:617-628.

Salsburg, D., 1983. The Lifetime Feeding Study in Mice and Rats--An Examination of its Validity as a Bioassay for Human Carcinogen, Fundamental and Applied Toxicology, 3:63-67.

Schneiderman, M. A., Decoufle, P. and Brown, C. C., 1979. Thresholds for Environmental Cancer: Biologic and Statistical Considerations. Annals of the New York Academy of Services, 329:92.

Sorrell, K., 1982. Preliminary Experimental Results on the effect of aeration on the VOC content of Drinking Water, USEPA, Municipal Research Laboratory, Cincinnati, OH. Personal Communication.

Sperling, F. and Marcus W. L., 1965. Thalidomide in Blood, Pharmacologist, 1:147.

Surgeon General, 1970. Report of the Ad Hoc Committee on the Evaluation of Low Levels of Environmental Chemical Carcinogens, Evaluation of Environmental Carcinogens Re-

Symons, J. M., et al., National Organics Reconnaissance Survey for Halogenated Organics. J. Amer. Water Works Assoc. 667 (11):634-647. 1975.

Tomatis, L., 1979. The Predictive Value of Rodent Carcinogenicity Test in the Evaluation of Human Risks, Annual Reviews of Pharmacology and Toxicology, 19:511-530.

United Nations Scientific Committee on the Effects of Atomic Radiation, 1977. Report to the General Assembly. Sources and Effects of Ionizing Radiation (United Nations, New York, NY).

USEPA: The National Public Water System Program. FY 82 Status Report. Office of Drinking Water, USEPA, 1983.

USEPA: National Interim Primary Drinking Water Regulations, USEPA. 570/9-76-003, Office of Drinking Water, 1976.

USEPA: Community Water Supply Survey Computer Data File on Inorganic and Organic Chemicals. Office of Drinking Water, USEPA, Washington, D.C., 1976.

USEPA: Rural Water Survey Computer Data provided by Department of Sociology, Cornell University for the Office of Drinking Water, 1982.

USEPA: Community Water Supply Survey, Bureau of Water Hygiene, Public Health Service, Department of Health, Education and Welfare, Washington, D.C., 1970.

USEPA, 1983. Health Assessment Document for Carbon Tetrachloride Cancer Assessment Group, USEPA/ECAO, EPA 600/8 82-001, Revised Draft, March, 1983.

USEPA, 1983a. Air Quality Criteria for Lead, EPA-600/8-83-028A, Washington, D.C. 20460.

USEPA, 1983b. Health Assessment Document for Inorganic Arsenic, EPA-600/ 8-83-021A, Washington, D.C.

U. S Department of Health Education and Welfare 1978. Special Occupational Hazard Review of Trichloroethylene, Center for Disease Control, NIOSH, Rockville, MD 20857.

U. S. vs. Kepplinger, et al., 1983. Northern District of Illinois, #81CR235.

Van Ryzin, J., 1980a. Quantitative Risk Assessment, J. Occupational Medicines, 22:321-326.

Van Ryzin, J. and Rai, K., 1980b. The Use of Quartel Response Data to Make Prediction, Witsch, H., ed., The Scientific Basis of Toxicity Assessment, North-Holland Biomedial Press, New York

Van Ryzin, J., 1984. Columbia University, Biostatistics Department, Private communication.

Waktanabe, P. G., 1983. Personal Communication, Concerning Sprague-Dawley Rats, Spartan Substran, Dow Chemical Co., 1803 Building, Midland, MI 48640.

Watson, K. S. Farrell, R. P. and Anderson, J. S., "Contribution from Individual Homes to the Sewer System", JWPCF 30:12. 2039-2054. December, 1967.

Weisburger, J. H and Williams, G. M., 1983. Chemical Carcinogens, Chapter 2, Bioassay of Carcinogens: In Vitro and In Vivo Tests, Washington, D.C.

Whittmore, A. S , 1983. Facts and Values in Risk Analysis for Environmental Toxicants, Risk Analysis, 3: 23-33.

Westrick, J. J., Mello, J. W. and Thomas R. F., The Ground Water Supply Summary of Volatile Organic Contaminant Occurrence Data, EPA Technical Support Division, Office of Drinking Water, Cincinnati, OH. 1983.

APPENDIX A

EXPLANATION OF HOW ANIMAL DATA WAS EXTRAPOLATED TO A CONTINUOUS HUMAN EXPOSURE EQUIVALENT

Benzene

Reference - Maltoni, et al., 1982.

The health endpoint considered for benzene was the zymbal gland carcinoma in female rats. The following is the bioassay data for this contaminant:

Animal, mg/kg/day, 4 to 5 times/week for 52 weeks	Human Equivalent, mg/kg/day	Animals Affected/Total
0	0	0/30 = 0
50	2.61	2/30 = 0.067
250	13.1	8/32 = 0.25

To convert the given dose to mg/kg/day continuous human equivalent, multiply by 52/104, since the exposure was for one year and the lifetime for rats is two years, and multiply by the cube root of the ratio of the rat and human weights, or $(0.3/70)^{1/3} = 0.162$. To convert the four to five times per week feeding to continuous exposure, multiply by 4.5/7. Thus, for the 50 mg/kg for the 4 to 5 times/week, the conversion is:

$$50 \text{ mg/kg} \times (4.5/7)\ (52/104)\ (0.3/70)^{1/3}$$

$$= 2.61 \text{ mg/kg/day human equivalent.}$$

Carbon Tetrachloride

Reference - NCI, 1976a, 1976b, USEPA, 1983.

The bioassay data from the NCI rat study is shown as follows:

Animal mg/day	Animal mg/kg/day	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	0	0/37
11	24	4.5	2/45
18	40	7.4	4/46
21	47	8.7	2/47
36	80	14.9	1/30

To convert from mg/day to mg/kg/day, divide by the average body weight of the rats, or 0.45 kg, and to get the human equivalent, multiply by the cube root of the ratio of the rat and human body weights, or $(0.45/70)^{1/3} = 0.186$. The bioassay data from the NCI mouse study is shown as follows:

Animal mg/day	Animal mg/kg/day	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	0	6/157
21	750	56	89/89
42	1500	112	90/93

To convert from mg/day to mg/kg/day, divide by the average body weight, or 0.028 kg, and to get the human equivalent, multiply by the cube root of the ratio of the mouse and human body weights, or $(0.028/70)^{1/3} = 0.074$.

1,2-Dichloroethane (Ethylene Dichloride)

Reference - Draft CAG (USEPA) document quoting 1978 NCI bioassay report.

The following is the bioassay data for hemangiosarcomas in the circulatory system of male rats:

Animal mg/kg/day	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	0/40
47	4.85	9/48
95	9.80	7/27

To convert from animal to human equivalent dose, multiply by 5/7 since the animals were dosed 5 days/week multiply by 78/104, since the experiment lasted 78 weeks and the average lifetime is 104 weeks, and multiply by the cube root of the ratio of the rat and human body weights, or $(0.5/70)^{1/3} = 0.193$.

1,1-Dichloroethylene (Vinylidene Chloride)

Reference - Maltoni, et al., 1980.

The bioassay for the resulting kidney adenocarcinomas of mice is shown below:

ppm/Daily for 12 months	Lifetime Continuance Equivalent, ppm	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	0	0/136
10	0.54	0.023	0/125
25	1.34	0.060	28/119

To convert from daily dose to lifetime continuous dose, multiply by 4/24 since exposure was 4 hours/day, multiply by 4.5/7 since the exposure was 4 to 5 times per week and multiply by 1/2 since the experiment was continued only for 1/2 the lifetime. Thus, $4/24 \times 4.5/7 \times 1/2 = 0.0536$.

To convert from ppm to micrograms/m^3, use the relationship 1 microgram/m^3 = $\frac{10\text{-}3 \text{ ppm}}{1.2 \times [MW (VDC)/MW (Air)]}$. Since the molecular weight of (Mn) vinylidene chloride is 97 and that of air is 28.8, then 1 ug/m^3 = 2.5×10^{-4} ppm. To convert from micrograms/m^3 to mg/kg/day, multiply micrograms/m^3 by ($20 m^3$/day x mg/1,000 micrograms x 1/70 kg) = 2.86×10^{-4}; multiply by the cube 70 kg root of the body weight ratios to convert to human equivalent $(0.03/70)^{1/3} = 0.075$ and divide by 2 on the assumption that 1/2 of that air is absorbed. Thus:

1,1-Dichloroethylene (Vinylidene Chloride) - continued

$$\frac{2.86 \times 10\text{-}4\ (0.075)}{2.5 \times 10\text{-}4 \times 2} = 0.042.$$

Perchloroethylene (Tetrachloroethylene) (Inhalation)

Reference - Rampy, et al., 1978.

The following are the bioassay results from this inhalation study:

Animal ppm	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	2/189
300	4.5	2/94
600	9.1	3/94

To change from ppm to mg/kg/day (human equivalent) multiply by 1.2×10^3 (165.8/28/8) = 6900, where 165.8 is the molecular weight of perchloroethylene and 28.8 is the molecular weight of air. Also multiply by the cube root of the ratio of the body weights $(0.350/70)^{1/3} = 0.171$ and 1/2 to adjust for 50% of that inhaled being absorbed. Then multiply by ($20 m^3$/day x 1 mg/1,000 micrograms x 1/70 kg) = 2.86×10^{-4}. Multiply by 6/24 since the exposure was for 6 hours/day and by 5/7 since the experiment was for 5 days/week and by 52/104 since the experiment lasted only 1/2 the lifetime. Thus to change from ppm to mg/kg/day human equivalent, multiply by:

$2.86 \times 10^{-4} \times 6900 \times 0.171 \times (6/24) \times (5/7) \times (52/104)$
$= 0.0151$.

Perchloroethylene (Tetrachloroethylene) (Ingestion)

Reference - NCI, 197?.

The bioassay data from resulting hepatocellular carcinomas of female mice are shown below:

Animal mg/kg/day	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	0/20
386	18.02	18/48
772	36.04	19/45

To convert from animal to human equivalent, multiply by 5/7 since the experiment was for 5 days/week and by 78/90 since the experiment was 78 weeks long and the average lifetime was 90 weeks, and by the cube root of the ratio of body weights $(0.03/70)^{1/3}$ equals 0.0754 or multiply altogether by:

$5/7 \times 78/90 \times 0.054 = 4.67 \times 10^{-2}$.

Trichloroethylene

Reference - NCI, 1976b.

The following are the bioassay data from hepatocellular carcinomas in male mice:

Animal mg/kg/day	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	1/20
1530	56.3	26/50
2700	112.6	31/48

To convert from animal to human equivalent, multiply by 5/7 since the experiment was done for 5 days/week, by 1.5/2.0 since the experiment lasted 1 1/2 years and the lifetime was 2 years and by the cube root of the ratio the body weights $(0.033/70)^{1/3}$ or:

$5/7 \times 1.5/2.0 \times (0.033/70)^{1/3} = 0.0417.$

Vinyl Chloride (Feron)

Reference - Feron, 1981.

The following are the bioassay data from female rats exposed to vinyl chloride:

Animal mg/kg/day	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	2/57
1.7	0.241	26/58
5.0	0.71	42/59
14.1	2.00	56/57

To convert from animal to human equivalent, multiply by the cube root of the ratio of the body weights $(0.20/70)^{1/3} = 0.142$.

Vinyl Chloride (Maltoni) (Inhalation)

Reference - Maltoni, 1975.

Animal ppm	Human Equivalent mg/kg/day	Animals Affected/Total
0	0	6/58
50	2.3	10/59
250	11	16/59
500	23	22/59
2500	110	32/59

To convert the animal doses to human equivalent chronic doses in micrograms/m^3, multiply by 1.2×10^3 [62.5/28/8] where 62.5 is the molecular weight of vinyl chloride and 28.8 is the molecular weight of air. Then multiply by 5/7 since the exposure was 5 days/week and by 52/104 since the exposure was for 52 weeks and the lifetime was 104 weeks. Then multiply by the cube root of the ratio of body weights or $(0.35/70)^{1/3} = 0.171$.

Then to convert to mg/kg/day multiply by 2.86×10^{-4}. Thus to go from animal dose in ppm to human equivalent chronic dose in mg/kg/day multiply by:

$1.2 \times 10^3 (62.5/28.8)(5/7)(52/104)(0.35/70)^{1/2} (2.06 \times 10^{-4}) = 0.0455$.

APPENDIX B

Point Estimate and Upper 95% Confidence Limit Values For the of Four Models (Logit, Multistage, Probit and Weibull) to the Bioassay Data

CONTAMINANT BENZENE
EXPOSURE ROUTE: INGESTION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Litetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0
2.61									0.067
13.1									0.25
Extrapolation									
1.0	2.2×10^{-2}	3.6×10^{-2}	2.9×10^{-2}	8.3×10^{-2}	2.8×10^{-2}	8.1×10^{-2}	2.3×10^{-2}	7.5×10^{-2}	
10^{-1}	2.2×10^{-3}	3.7×10^{-3}	3.3×10^{-3}	1.8×10^{-2}	3.2×10^{-3}	1.5×10^{-2}	7.6×10^{-4}	5.8×10^{-3}	
10^{-2}	2.2×10^{-4}	3.7×10^{-4}	5.0×10^{-4}	3.3×10^{-3}	3.5×10^{-4}	2.4×10^{-3}	6.7×10^{-6}	9.6×10^{-5}	
10^{-3}	2.2×10^{-5}	3.7×10^{-5}	6.5×10^{-5}	5.5×10^{-4}	3.9×10^{-5}	3.4×10^{-4}	1.6×10^{-8}	3.7×10^{-7}	
10^{-4}	2.2×10^{-6}	3.7×10^{-6}	8.5×10^{-6}	8.7×10^{-5}	4.3×10^{-6}	4.7×10^{-5}	9.6×10^{-12}	3.3×10^{-10}	
10^{-5}	2.2×10^{-7}	3.7×10^{-7}	1.1×10^{-6}	1.3×10^{-5}	4.8×10^{-7}	6.1×10^{-6}	1.6×10^{-15}	7.2×10^{-14}	
10^{-6}	2.2×10^{-8}	3.7×10^{-8}	1.4×10^{-7}	2.0×10^{-6}	5.3×10^{-8}	7.8×10^{-7}		3.7×10^{-18}	

CONTAMINANT: CARBON TETRACHLORIDE (NCI Rat)
EXPOSURE ROUTE: INGESTION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0
4.5									0.044
7.4									0.086
8.7									0.043
14.9									0.033
Extrapolation									
1.0	6.6×10^{-3}	1.1×10^{-2}	8.2×10^{-2}	3.0×10^{-1}	8.2×10^{-2}	3.0×10^{-1}	8.2×10^{-2}	2.9×10^{-1}	
10^{-1}	6.6×10^{-4}	1.1×10^{-3}	1.3×10^{-1}	8.6×10^{-1}	1.3×10^{-1}	8.3×10^{-1}	1.3×10^{-1}	7.3×10^{-1}	
10^{-2}	6.6×10^{-5}	1.1×10^{-4}	2.1×10^{-1}	1.0	2.0×10^{-1}	1.0	1.9×10^{-1}	1.0	
10^{-3}	6.6×10^{-6}	1.1×10^{-5}	3.2×10^{-1}	1.0	3.0×10^{-1}	1.0	2.6×10^{-1}	1.0	
10^{-4}	6.6×10^{-7}	1.1×10^{-6}	4.7×10^{-1}	1.0	4.2×10^{-1}	1.0	3.5×10^{-1}	1.0	
10^{-5}	6.6×10^{-8}	1.1×10^{-7}	6.5×10^{-1}	1.0	5.5×10^{-1}	1.0	4.4×10^{-1}	1.0	
10^{-6}	6.6×10^{-9}	1.1×10^{-8}	8.2×10^{-1}	1.0	6.7×10^{-1}	1.0	5.5×10^{-1}	1.0	
10^{-7}	6.6×10^{-10}	1.1×10^{-9}	9.4×10^{-1}	1.0	7.7×10^{-1}	1.0	6.4×10^{-1}	1.0	

CONTAMINANT CARBON TETRACHLORIDE (NCI Mouse)
EXPOSURE ROUTE: INGESTION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0.038
56									1.0
112									0.97
Extrapolation									
1.0	4.9×10^{-2}	6.1×10^{-2}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	
10^{-1}	5.0×10^{-3}	6.3×10^{-3}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	
10^{-2}	5.0×10^{-4}	6.3×10^{-4}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	
10^{-3}	5.0×10^{-5}	6.3×10^{-5}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	
10^{-4}	5.0×10^{-6}	6.3×10^{-6}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	
10^{-5}	5.0×10^{-7}	6.3×10^{-7}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	
10^{-6}	5.0×10^{-8}	6.3×10^{-8}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	9.6×10^{-1}	9.9×10^{-1}	

CONTAMINANT: 1,2-DICHLOROETHANE
EXPOSURE ROUTE: INGESTION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0
4.85									0.19
9.80									0.26
Extrapolation									
1.0	3.6×10^{-2}	5.2×10^{-2}	8.7×10^{-2}	2.7×10^{-1}	8.3×10^{-2}	2.8×10^{-1}	7.6×10^{-2}	2.9×10^{-1}	
10^{-1}	3.6×10^{-3}	5.4×10^{-3}	2.7×10^{-2}	1.6×10^{-1}	2.2×10^{-2}	1.5×10^{-1}	1.3×10^{-2}	1.2×10^{-1}	
10^{-2}	3.6×10^{-4}	5.4×10^{-4}	8.1×10^{-3}	7.0×10^{-2}	5.9×10^{-3}	5.7×10^{-2}	1.3×10^{-3}	2.3×10^{-2}	
10^{-3}	3.6×10^{-5}	5.4×10^{-5}	2.4×10^{-3}	2.8×10^{-2}	1.5×10^{-3}	1.9×10^{-2}	7.2×10^{-5}	2.1×10^{-3}	
10^{-4}	3.6×10^{-6}	5.4×10^{-6}	7.3×10^{-4}	1.0×10^{-2}	3.9×10^{-4}	6.2×10^{-3}	2.2×10^{-6}	9.3×10^{-5}	
10^{-5}	3.6×10^{-7}	5.4×10^{-7}	2.2×10^{-4}	3.7×10^{-3}	9.9×10^{-5}	1.9×10^{-3}	3.7×10^{-8}	2.2×10^{-6}	
10^{-6}	3.6×10^{-8}	5.4×10^{-8}	6.6×10^{-5}	1.3×10^{-3}	2.5×10^{-5}	5.6×10^{-4}	3.3×10^{-10}	2.6×10^{-8}	
10^{-7}	3.6×10^{-9}	5.4×10^{-9}	2.0×10^{-5}	4.4×10^{-4}	6.5×10^{-6}	1.6×10^{-4}	1.7×10^{-12}	1.6×10^{-10}	

CONTAMINANT: 1,1-DICHLOROETHYLENE
EXPOSURE ROUTE: INHALATION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Litetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0
0.023									0
0.060									0.24
Extrapolation									
1.0	1.0	1.0							
10^{-1}	5.1×10^{-1}	6.7×10^{-1}							
10^{-2}	7.2×10^{-3}	4.4×10^{-2}							
10^{-3}	7.2×10^{-5}	3.9×10^{-3}							
10^{-4}	7.2×10^{-7}	3.9×10^{-4}							
10^{-5}	7.2×10^{-9}	3.8×10^{-5}							
10^{-6}	7.2×10^{-11}	3.8×10^{-6}							

CONTAMINANT PERCHLOROETHYLENE
EXPOSURE ROUTE: INHALATION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0.01
4.5									0.021
9.1									0.031
Extrapolation									
1.0	2.4×10^{-3}	6.3×10^{-3}	2.3×10^{-3}	2.1×10^{-2}	2.3×10^{-3}	2.2×10^{-2}	1.9×10^{-3}	2.1×10^{-2}	
10^{-1}	2.4×10^{-4}	6.3×10^{-4}	2.3×10^{-4}	4.2×10^{-3}	2.2×10^{-4}	4.1×10^{-3}	6.7×10^{-5}	1.9×10^{-3}	
10^{-2}	2.4×10^{-5}	6.3×10^{-5}	2.3×10^{-5}	6.2×10^{-4}	2.1×10^{-5}	6.0×10^{-4}	1.1×10^{-6}	5.6×10^{-5}	
10^{-3}	2.4×10^{-6}	6.3×10^{-6}	2.2×10^{-6}	8.0×10^{-5}	2.0×10^{-6}	7.6×10^{-5}	7.9×10^{-9}	6.4×10^{-7}	
10^{-4}	2.4×10^{-7}	6.3×10^{-7}	2.1×10^{-7}	9.8×10^{-6}	2.0×10^{-7}	9.1×10^{-6}	2.5×10^{-11}	2.9×10^{-9}	
10^{-5}	2.4×10^{-8}	6.3×10^{-8}	2.1×10^{-8}	1.1×10^{-6}	1.9×10^{-8}	1.0×10^{-6}	3.5×10^{-13}	5.6×10^{-12}	
10^{-6}	2.4×10^{-9}	6.3×10^{-9}	2.0×10^{-9}	1.3×10^{-7}	1.8×10^{-9}	1.2×10^{-7}	1.1×10^{-16}	4.5×10^{-18}	
10^{-7}	2.4×10^{-10}	6.3×10^{-10}	2.0×10^{-10}	1.5×10^{-8}	1.7×10^{-10}	1.2×10^{-8}		1.5×10^{-18}	

CONTAMINANT: PERCHLOROETHYLENE
EXPOSURE ROUTE: INGESTION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0
18.02									0.40
36.04									0.42
Extrapolation									
1.0	2.0×10^{-2}	2.5×10^{-2}	2.7×10^{-1}	8.9×10^{-1}	2.8×10^{-1}	9.3×10^{-1}	2.9×10^{-1}	9.9×10^{-1}	
10^{-1}	2.0×10^{-3}	2.5×10^{-3}	2.1×10^{-1}	1.0	2.0×10^{-1}	1.0	2.2×10^{-1}	1.0	
10^{-2}	2.0×10^{-4}	2.6×10^{-4}	1.6×10^{-1}	1.0	1.5×10^{-1}	1.0	1.6×10^{-1}	1.0	
10^{-3}	2.0×10^{-5}	2.6×10^{-5}	1.2×10^{-1}	1.0	1.1×10^{-1}	1.0	1.1×10^{-1}	1.0	
10^{-4}	2.0×10^{-6}	2.6×10^{-6}	9.3×10^{-2}	1.0	7.5×10^{-2}	9.4×10^{-1}	7.3×10^{-2}	1.0	
10^{-5}	2.0×10^{-7}	2.6×10^{-7}	7.1×10^{-2}	8.9×10^{-1}	5.2×10^{-2}	7.9×10^{-1}	4.7×10^{-2}	9.5×10^{-1}	
10^{-6}	2.0×10^{-8}	2.6×10^{-8}	5.2×10^{-2}	7.7×10^{-1}	3.6×10^{-2}	6.3×10^{-1}	2.9×10^{-2}	7.2×10^{-1}	

CONTAMINANT: TRICHLOROETHYLENE
EXPOSURE ROUTE: INGESTION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0.05
56.3									0.52
112.6									0.65
Extrapolation									
1.0	9.8x10-3	1.3x10-2	7.3x10-2	2.9x10-1	3.8x10-2	2.0x10-1	2.2x10-2	1.7x10-1	
10-1	9.8x10-4	1.3x10-3	2.2x10-2	1.2x10-1	6.6x10-3	5.1x10-2	9.3x10-4	1.4x10-2	
10-2	9.9x10-5	1.3x10-4	6.5x10-3	4.8x10-2	1.1x10-3	1.1x10-2	1.2x10-5	3.1x10-4	
10-3	9.9x10-6	1.3x10-5	1.9x10-3	1.7x10-2	1.8x10-4	2.2x10-3	4.4x10-8	1.8x10-6	
10-4	9.9x10-7	1.3x10-6	5.7x10-4	6.0x10-3	3.0x10-5	4.5x10-4	4.9x10-11	2.8x10-9	
10-5	9.9x10-8	1.3x10-7	1.7x10-4	2.0x10-3	5.0x10-6	8.6x10-5	1.6x10-14	1.3x10-11	
10-6	9.9x10-9	1.3x10-8	4.9x10-5	6.8x10-4	8.3x10-7	1.6x10-5			
10-7	9.9x10-10	1.3x10-9	1.4x10-5	2.2x10-4	1.4x10-7	3.0x10-6			

CONTAMINANT: VINYL CHLORIDE (FERON)
EXPOSURE ROUTE: INGESTION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0.035
0.241									0.45
0.71									0.71
2.00									0.98
Extrapolation									
1.0	8.5×10^{-1}	9.0×10^{-1}	8.1×10^{-1}	8.9×10^{-1}	8.2×10^{-1}	8.9×10^{-1}	8.2×10^{-1}	8.9×10^{-1}	
10^{-1}	1.7×10^{-1}	2.1×10^{-1}	1.9×10^{-1}	2.9×10^{-1}	1.4×10^{-1}	2.3×10^{-1}	1.3×10^{-1}	2.3×10^{-1}	
10^{-2}	1.9×10^{-2}	2.3×10^{-2}	2.6×10^{-2}	5.3×10^{-2}	4.5×10^{-3}	1.2×10^{-2}	5.7×10^{-4}	2.6×10^{-3}	
10^{-3}	1.9×10^{-3}	2.3×10^{-3}	3.1×10^{-3}	8.2×10^{-3}	1.3×10^{-4}	4.8×10^{-4}	3.5×10^{-8}	3.4×10^{-7}	
10^{-4}	1.9×10^{-4}	2.3×10^{-4}	3.8×10^{-4}	1.2×10^{-3}	3.6×10^{-6}	1.7×10^{-5}	2.5×10^{-14}	4.3×10^{-13}	
10^{-5}	1.9×10^{-5}	2.3×10^{-5}	4.5×10^{-5}	1.7×10^{-4}	1.0×10^{-7}	6.0×10^{-7}		5.0×10^{-21}	
10^{-6}	1.9×10^{-6}	2.3×10^{-6}	5.5×10^{-6}	2.4×10^{-5}	2.9×10^{-9}	2.0×10^{-8}		6.0×10^{-31}	

CONTAMINANT. VINYL CHLORIDE (MALTONI)
EXPOSURE ROUTE: INHALATION

Human Equivalent Dose (mg/kg/day)	Risk (Cases/Lifetime/Person)								
	Multistage		Weibull		Logit		Probit		
	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Point Estimate	Upper 95% Confidence Limit	Bio-assay Data
Bioassay Data									
0									0.10
2.3									0.17
11									0.27
23									0.37
110									0.54
Extrapolation									
1.0	6.5×10^{-3}	9.2×10^{-3}	5.7×10^{-2}	1.2×10^{-1}	4.8×10^{-2}	1.0×10^{-1}	3.8×10^{-2}	9.5×10^{-2}	
10^{-1}	6.5×10^{-4}	9.2×10^{-4}	1.8×10^{-2}	4.7×10^{-2}	1.2×10^{-2}	3.5×10^{-2}	4.7×10^{-3}	1.9×10^{-2}	
10^{-2}	6.5×10^{-5}	9.2×10^{-5}	5.7×10^{-3}	1.8×10^{-2}	3.0×10^{-3}	1.1×10^{-2}	3.0×10^{-4}	1.9×10^{-3}	
10^{-3}	6.5×10^{-6}	9.2×10^{-6}	1.8×10^{-3}	6.6×10^{-3}	7.4×10^{-4}	3.1×10^{-3}	9.9×10^{-6}	9.1×10^{-5}	
10^{-4}	6.5×10^{-7}	9.2×10^{-7}	5.5×10^{-4}	2.3×10^{-3}	1.8×10^{-4}	8.8×10^{-4}	1.7×10^{-7}	2.1×10^{-6}	
10^{-5}	6.5×10^{-8}	9.2×10^{-8}	1.7×10^{-4}	8.4×10^{-4}	4.4×10^{-5}	2.5×10^{-4}	1.4×10^{-9}	2.4×10^{-8}	
10^{-6}	6.5×10^{-9}	9.2×10^{-9}	5.4×10^{-5}	2.4×10^{-4}	1.1×10^{-5}	6.7×10^{-5}	6.0×10^{-12}	1.3×10^{-10}	
10^{-7}	6.5×10^{-10}	9.2×10^{-10}	1.7×10^{-5}	8.2×10^{-5}	2.6×10^{-6}	1.8×10^{-5}	1.3×10^{-14}	3.5×10^{-13}	

APPENDIX C

OCCURENCE DATA FOR SELECTED VOC'S IN DRINKING WATER,
POPULATION CONCENTRATION AND STANDARD ERRORS
(FROM REFERENCE USEPA, JRB REPORTS)

This Appendix is taken directly from the Occurrence/ Exposure documents being developed for potential drinking water regulation. This section describes what methodology is used to develop estimates of national occurrence. The U. S. Environmental Protection Agency (EPA) is currently considering the proposal of national revised primary drinking water regulations under the Sare Drinking Water Act, as well as non-federal regulatory approaches, to limit human exposure to high levels of certain volatile synthetic organic chemicals (VOC's) that have been detected in drinking water (Advanced Notice of Proposed Rulemaking, 47 FR 9350, March 4, 1982). The specific VOC's of immediate interest to EPA are:

- Trichloroethylene
- Tetrachloroethylene
- Carbon Tetrachloride
- 1,1,1-Trichloroethane
- 1,2-Dichloroethane
- Vinyl Chloride
- Dichloromethane
- Benzene
- Chlorobenzene
- Dichlorobenzene(s)
- 1,1-Dichloroethylene
- 1,2-Dichloroethylene (cis and trans)

The objective of the analysis of the occurrence of the VOC's listed above is to support EPA's consideration of the need and alternatives for controlling VOC's in public water supplies in two principal areas:

- As input to the health risk assessments for the VOC's, the analysis provides estimates of the number of individuals in the United States exposed to various levels of VOC's in drinking water from public water supplies.
- As input to the assessment of the economic impact of the regulatory and treatment alternatives being considered, the analysis provides estimates of the number of public water supplies of various source (i.e., groundwater and surface water) and size (based on population served) categories having VOC's present, and the distribution of VOC levels in those water supplies.

The methodology for preparing these estimates involved the creation of a data base drawing from the results of several Federal surveys on the measured occurrence of the VOC's in public water supplies as a function of water source and supply size. Statistical models were then used to extrapolate from the observed frequency of occurrence of the VOC's in the supplies sampled in the Federal surveys to the universe of public water supplies having similar source and size characteristics. A separate report has been prepared for each of the VOC's listed. Section A is included in each report to provide detail on the sources of data and the methodology used to obtain these estimates. Some specifics presented in Section A may not be applicable to the VOC that is the subject of this report. Section B provides additional detail on the selection of an

appropriate model for estimating the national occurrence of the VOC that is addressed in this report.

A.1 SOURCES OF DATA

A.1.1 Number of Public Drinking Water Supplies in the United States and Size of Populations Served

It is currently estimated that there are approximately 60,000 public drinking water supplies in the United States serving approximately 214 million people. Table A-1 summarizes the estimated number of surface water and groundwater systems of various sizes and the associated populations served by them. These data, which correspond to the "FY 82 Characterization of the Water Supply Industry" presented by Kuzmack (1983), as updated by Schnare*, were derived from the Federal Reporting Data Systems (FRDS) for FY 1982 (FRDS, 1983).

It should be noted that FRDS (1983) does not provide a breakdown of the number of systems nor of the population served for the 10,001-25,000 and 25,001-50,000 size categories, but rather for 10,001-50,000 as a single category. The estimated number of systems in the 10,001-25,000 and 25,001-50,000 size categories as presented in Kuzmack (1983), were estimated by Dr. David Schnare of the Office of Program Development and Evaluation, EPA Office of Drinking Water from additional FRDS data. (Data for these additional categories are needed for the economic impact analysis.) The population served by systems in

---

* Personal communication between David Schnare, Office of Drinking Water, U. S. Environmental Protection Agency, and Frank Letkiewicz, JRB Associates, May 25, 1983.

these size categories were estimated by JRB as shown in Table A-2.

TABLE A-1

NUMBER OF SYSTEMS AND POPULATION SERVED
BY PRIMARY WATER SUPPLY SOURCE
(BY POPULATION CATEGORY)

System Size	(population served)	Surface			Groundwater		
		No. of Systems	Population (thousands)	Average Population Served	No. of Systems	Population (thousands)	Average Population Served
Very Small	25-100	1,525	86	56	19,125	1,031	54
	101-500	2,412	690	286	15,674	3,814	243
	501-1,000	1,377	1,051	763	4,877	3,590	736
Small	1,001-2,500	1,945	3,295	1,700	4,400	7,047	1,600
	2,501-3,300	495	1,445	2,900	891	2,583	2,900
	3,301-5,000	749	3,096	4,100	1,065	4,370	4,100
Medium	5,001-10,000	930	6,763	7,300	1,168	8,404	7,200
	10,001-25,000	915[a]	15,595	17,000	835[a]	12,276[b]	15,000
	25,001-50,000	400[a]	13,945[b]	35,000	290[a]	10,977[b]	38,000
Large	50,001-75,000	155	9,483	61,000	64	3,911	61,000
	75,001-100,000	82	7,131	87,000	14	1,184	85,000
Very Large	>100,000	217	78,366	360,000	55	14,286	260,000
TOTALS[c]:		11,202	140,948		48,458	73,475	

[a] Kuzmack, 1983, as updated by David Schnare, Office of Drinking Water, U. S. Environmental Protection Agency, in a personal communication with Frank Letkiewicz, JRB Associates, May 25, 1983.
[b] Estimated by JRB Associates (See Table A-2).
[c] Populations do not add to total due to rounding.

TABLE A-2

ANALYSIS FOR ESTIMATING THE POPULATION SERVED BY SURFACE WATER AND GROUNDWATER SUPPLIES IN THE 10,001-25,000 and 25,001-50,000 POPULATION SIZE CATEGORIES

System Size (population served)	Source		Total
	Surface Water	Groundwater	
10,001-25,000	w	x	27,870[a]
25,001-50,000	y	z	24,920[a]
TOTAL	29,540[b]	23,250[b]	52,790

[a] Estimated total population for surface water and groundwater provided by Kuzmack, (1983).

[b] Population served by surface water and groundwater supplies in the 10,001-50,000 category from FDRS, (1983).

$$w = \frac{29,540 \times 27,870}{52,790} = 15,595.$$

$$x = \frac{23,250 \times 27,870}{52,790} = 12,276.$$

$$y = \frac{29,540 \times 24,920}{52,790} = 13,945.$$

$$z = \frac{23,250 \times 24,920}{52,790} = 10,977.$$

A.1.2 Data on Measured Occurrence of VOC's in Public Water Supplies

There are three primary sources of information available on the occurrence of VOC's in drinking water supplies that were considered for use in preparing the national projections:

- Federal surveys
- State data
- Miscellaneous data

Only the Federal survey data were ultimately used for the national projections. While a substantial amount of state and miscellaneous published data were available for some of the VOC's examined, these data usually did not provide adequate details on the water source, size of the population served, or type of sample taken (raw, finished, or distribution), which precluded incorporating them into the analysis. Furthermore, there was usually no detailed information available on the sampling and analysis methodology used to obtain the data, which precluded their being subjected to a quality assurance review (performed as a separate task) (JRB Associates, 1982). In addition, much of the state and miscellaneous data appeared to have been obtained in response to spills, citizen complaints, or other evidence of contamination and, therefore, were not considered to be representative data for preparing the national projections. While the state and miscellaneous published data were not used for deriving the national projections,

these data are presented and discussed in the individual VOC occurrence reports.

The Federal survey data generally provided the information on water source, population served, and sample type studied that was necessary to perform the analyses in this report. These surveys also provided sufficient information on the sampling and analysis methods to be subjected to the quality assurance review. The following six Federal surveys were used for the national projections:

° National Organics Reconnaissance Survey (NORS)

The National Organics Reconnaissance Survey (NORS) was conducted early in 1975 for the purpose of determining levels of four trihalomethanes (chloroform, bromodichloromethane, dibromochloromethane, and bromoform), carbon tetrachloride, and 1,2-dichloroethane in finished water supplies from 80 cities across the country (Symons et al., 1975). A population base of 36 million individuals was covered during the study. Analysis of samples was performed by the Water Supply Research Laboratory of EPA in Cincinnati using purge and trap gas chromatography with an electrolytic conductivity detector.

° National Organic Monitoring Survey (NOMS)

The National Organic Monitoring Survey (NOMS) was conducted to determine the frequency of occurrence of specific organic chemicals in finished water supplies or 113 cities across the country (Brass et al., 1977). Among the chemicals surveyed were trihalomethanes, 1,2-dichloroethane, carbon tetrachloride, trichloroethylene, benzene, vinyl chloride, dichlorobenzene, and trichlorobenzene. Data from three phases (referred to as NOMS I, NOMS II, and NOMS III) of the study were collected over an eleven month period (March 1976 to January

1977) to reflect any long-term or seasonal variations. The analytical treatment of the samples was similar to that for the NORS samples. (Gas chromatography/mass spectrometry analyses were done for benzene.)

° National Screening Program for Organics in Drinking Water (NSP)

SRI International conducted a study from June 1977 to March 1981, entitled "National Screening Program for Organics in Drinking Water" (NSP), in which raw and finished drinking water samples were collected from 160 water facilities located in 35 states (Boland, 1981). The compounds sampled were 23 halocarbons, 6 aromatics, and 22 pesticides, phenols, and acids. The methods used for anaylsis included gas chromatography with electron capture detection for purgeable halocarbons and the base/neutral extraction fraction, and gas chromatography with flame ionization detection for purgeable aromatics.

° Community Water Supply Survey (CWSS)

The Community Water Supply Survey (CWSS) examined 106 surface water supplies, 330 groundwater supplies, and 16 supplies with mixed water or purchased sources in 1978. Trihalomethanes and other volatile organic chemicals, including carbon tetrachloride, chlorobenzene, 1,2-dichloroethane, cis- and trans-1,2-dichloroethylene, tetrachloroethylene, 1,1,1-trichloroethane, trichloroethylene, benzene, toluene, and xylenes were measured. One to five samples were collected from each system, including raw, finished, and/or distribution water. Gas chromatography with an electrolytic conductivity director was used for halocarbons and a flame ionization detector for aromatic analyses.

° Rural Water Survey (RWS)

The Rural Water Survey, conducted in 1978, was carried out in response to Section 3 of the Safe Drinking Water Act, which mandated that EPA "conduct a survey of the quantity, quality, and availability of rural drinking water supplies". Samples collected from 855 households in rural areas from across the

United States were analyzed for trihalomethanes and for carbon tetrachloride, 1,2-dichloroethane, cis- and trans-1,2-dichloroethylene, tetrachloroethylene, 1,1,1-trichloroethane, and trichloroethylene using gas chromatography with an electrolytic conductivity detector. The majority of the 855 samples were from households using private wells or small supplies serving fewer than 25 people. Using information provided by Dr. Bruce Brower at Cornell University, Department of Rural Sociology, it was determined that the RWS file had data for up to 207 groundwater and up to 45 surface water supplies serving more than 25 people.

- Groundwater Supply Survey (GWSS)

  The Groundwater Supply Survey (GWSS), conducted between December 1980 and December 1981, involved the national sampling of 945 public water supply systems using groundwater sources for 5 trihalomethanes and 29 other organic chemicals. Analyses were done using purge and trap gas chromatography with an electrolytic conductivity detector for halocarbons and a non-destructive photoionization detector for aromatics. There were 466 randomly selected supplies and 479 selected with state and EPA regional input based on the likelihood of finding some VOC contamination.

Table A-3 indicates which of the VOC's listed earlier were examined in each of the six Federal surveys used for the national projections.

TABLE A-3

VOC'S OF INTEREST EXAMINED IN THE SIX FEDERAL SURVEYS USED FOR NATIONAL PROJECTIONS

	NORS	NOMS	NSP	CWSS	RWS	GWSS
Trichloroethylene		X	X	X	X	X
Tetrachloroethylene		X	X	X	X	X
Carbon Tetrachloride	X	X	X	X	X	X
1,1,1-Trichloroethane		X	X	X	X	X
1,2-Dichloroethane	X	X	X	X	X	X
Vinyl Chloride		X	X			X
Dichloromethane		X	X			X[a]
Benzene		X	X	X		X
Chlorobenzene			X	X		X
Dichlorobenzene(s)		X	X			X
Trichlorobenzene(s)		X	X			
1,1-Dichloroethylene			X			X
1,2-Dichloroethylene (cis and trans)			X	X		X

a Dichloromethane data for the GWSS were not used due to a sample contamination problem.

A.2 GENERAL METHODOLOGY

A.2.1 Development of Survey Files

To facilitate the handling of data for preparing the national projections, suitable machine-readable files were developed for each of the six Federal surveys. JRB was provided access to existing computer files for the CWSS, RWS, and GWSS through the EPA Office of Drinking Water's Technical Support Division (OWD/TSD) in Cincinnati, Ohio. JRB used the published results of the NORS (Symons et al., 1975) and the NSP (Boland, 1981), and printed results for the NOMS provided by ODW/TSD to create machine readable files of those surveys. Three separate files were created for NOMS, one for each of the three phases. (In effect, NOMS I, II, AND III were treated as three separate surveys.) The final files for all chemicals were in SAS format. All computer efforts for this project utilized EPA's NCC-IBM (IBM 370) computer at Research Triangle Park North Carolina.

It was necessary to prepare working versions of each survey file containing the following minimum information for each of the sampled water supplies:

- Location of the supply (state and city)
- Population served by the supply
- Water source (groundwater, surface water, mixed, purchased, etc.)

- A single concentration value for each VOC sampled

With the exception of the RWS and NSP, the existing files and printed sources provided adequate information on the location of the supply sampled. The RWS design involved the collection of drinking water samples from households in rural areas of the United States. With the assistance of Dr. Bruce Brower at Cornell University's Department of Rural Sociology (responsible for the preparation of a detailed analysis of RWS results on inorganics, pesticides, and other parameters), it was possible to determine which of the 855 households for which VOC analyses were done obtained water from public water supplies. However, because of confidentiality restrictions on the RWS data, it was only possible to determine the location of the household and the public water supplies sampled at the state and county level, but not at the city or town level.

For NSP, the locations were not reported in Boland (1981); however, the Office of Drinking Water, Science and Technology Branch (ODW/STB) was able to provide copies of data sheets on the supplies sampled in NSP which provided information on location.

The existing files of the CWSS and GWSS each provided data on the size of the population served by the supplies sampled. For NORS, it was necessary to estimate the population served using information presented in Symons et al. (1975) on

the location of the supply and recent population data for those areas from other sources. It should be noted that most of the supplies sampled in the NORS fall into the large and very large size categories. Consequently, errors in the precise number used for persons served by those systems would not alter their classification or their impact on the national projections.

For NOMS data on the populations served by supplies were provided by ODW/TSD. In the case of NSP, the population served by each supply was not reported in Boland (1981). Again, those data were obtained from the data sheets provided by ODW/STB (There were three NSP locations for which the population was not specified.)

For RWS, data on the size of the population served by supplies were not collected. However, data were obtained on the number of service connections for each supply. With the assistance of Dr. Bruce Brower at Cornell University, it was possible to estimate the population served by each supply from the service connection data and data on the average number of individuals per household observed in the survey (3.034).

The identification of water source as groundwater, surface water, mixed, etc., was clearly designated in the CWSS, RWS, and GWSS files. For NORS, the source was determined from the descriptive information in Symons et al. (1975). For NOMS, source designations were provided by ODW/TSD. For NSP, source

information was given on the data sheets provided by ODW/STB.

Some public water supplies use a mixture of groundwater and surface water sources. Although these "mixed" supplies are counted as groundwater or surface water among the 60,000 supplies in FRDS based on the predominant source used, such supplies were excluded from the survey data for developing the national projections because the predominant source was rarely indicated in the survey file. Similarly, water supplies identified as purchasing water from another, usually unspecified, supply were excluded from the survey data for the national projections.

A 2.2 Computing Average Values for VOC's in Each Survey

In order to prepare the national projections, it was necessary that a single value be obtained for each VOC in each supply sampled. This requirement presented certain difficulties for several of the files where multiple sample results were reported for the supplies. NORS and NSP provided data on raw (i.e., untreated water sampled at the supply) and finished (i.e., treated water sampled at the supply) samples; CWSS provided data on raw, finished, and distribution (i.e., water sampled at a user's faucet) samples; NOMS and GWSS used finished water only; and RWS used distribution water only In order that the national projections be derived from data on drinking water samples most representative of what people

actually consume, all data on raw water were excluded from this analysis. Initially, consideration was given to excluding finished water sample data in the CWSS for supplies also having distribution samples. However, ODW/TSD staff indicated that inconsistencies in coding data for the CWSS resulted in some errors in designating water as finished or distribution. Therefore, all CWSS finished and distribution samples were included in the analysis as being equally representative of the water to which consumers are exposed.

The NORS, NOMS and GWSS provided a single analytical result for each VOC in finished water. For NSP and RWS, there was generally only a single value reported for each VOC, although a few systems had multiple samples. In CWSS, most supplies had multiple finished and/or distribution samples for each VOC. Where multiple samples occurred, a single "supply value" was computed for each VOC using the following rules:

- If positive values were reported for all samples, the supply value was computed as the arithmetic mean.
- If both positive values and values below the minimum quantifiable concentration were reported, the supply value was computed as the mean of the positive values and the minimum quantifiable concentration values.
- If values below the minimum quantifiable concentration were reported for all samples, the supply value was computed as the mean of the minimum quantifiable limits reported for each sample. (These means were recorded as "negative" values to indicate that the VOC was not observed.)

Computing a single value for supplies where the value for one or more of the samples was reported to be below the minimum quantifiable concentration was problematic. These "negative" values imply that the analyte in question may or may not be present and, if present, is so at a concentration below that measurable by the analytical method. In other words, the actual value is greater than or equal to zero and less than the minimum quantifiable concentration.

Where a supply was reported to have samples with both positive and negative values, two major alternatives were considered for treating the negative values. The first was to ignore or eliminate the negative values from the computation of the mean. This was considered unacceptable because it implies that the negative data are less valid than the positive data, which is not the case. This alternative would also necessarily result in a higher average value for that system than would be the case if the actual value for the negative data were known.

The second alternative was to assign the negative data a specific value for computing the supply average. Three possibilities were considered 0, the minimum quantifiable concentration, and the midpoint between 0 and the minimum quantifiable concentration. Assigning the negative samples a value equal to the minimum quantifiable concentration was selected since this gives the most conservative estimate of the

supply value. That is, if the analyte was in fact present, the maximum possible concentration it could have in the sample would be approximately (actually slightly less than) the minimum quantifiable concentration. For example, if a supply was reported to have one sample with a VOC present at 0.3 micrograms/liter and another sample in which the VOC was not observed at a minimum quantifiable concentration of 0.1 micrograms/liter, a supply value of 0.2 micrograms/liter was recorded in the working file. Using the minimum quantifiable concentration with other actual positive values to compute the mean results is the most conservative estimate of the supply value utilizing all sample data.

The treatment of supplies having only negative values reported derives from the treatment of those with negatives and positives described above. If, for example a VOC was not observed in two samples from a given system at minimum quantifiable concentrations of 0.1 micrograms/liter and 0.3 micrograms/liter, respectively, the system value retained in the working file was a "negative" 0.2 micrograms/liter. That is, the VOC was not observed, but if it had been present in both samples, the maximum possible average concentration for that supply would have been 0.2 micrograms/liter.

A.2.3 Combining Data from the Federal Surveys

Once the multiple samples for each VOC were averaged to

obtain a single value for each supply sampled, tables on the frequency of occurrence of each chemical were prepared for each survey as presented in Sections 2.2.1 and 2.3.1. In addition, the mean, median, range, and other statistics were computed for the positive values in each survey.

The next step in developing the national projections was to combine the results of all or the surveys together. In doing this, it was necessary to identify those supplies that had been sampled in more than one survey and compute an average supply value for each VOC. (The rules for averaging samples within a survey described in Section A.2.2 applied to computing averages across surveys.) It should be noted that supplies sampled in the RWS could not be matched against the other surveys since the RWS locations could only be determined at the state and county level, as previously described.

Table A-4 presents a list of those systems which were duplicated across the Federal surveys. When a system was sampled in two or more surveys, the population used for that system in the combined survey file was the one reported in the most recent survey, represented by the following chronological order (most recent first): GWSS NSP, CWSS, NOMS, and NORS.

TABLE A-4

SYSTEMS SAMPLED IN MORE THAN ONE SURVEY

Location	Source	Survey in Which the System was Sampled
Tuscon, AZ	G	GWSS, NORS
Fresno, CA	G	GWSS, NOMS
Jacksonville, FL	G	NOMS, NORS
Idaho Falls, ID	G	GWSS, NORS
Rockford, IL	G	GWSS, NOMS
Campbellsburg, IN	G	GWSS, CWSS
South Bend, IN	G	GWSS, NSP
Baton Rouge, LA	G	NSP, NOMS
Hammond, LA	G	GWSS, CWSS
Lafayette, LA	G	CWSS, CWSS
Senath MO	G	GWSS, CWSS
Webb City, MO	G	GWSS, CWSS
Greenville, MS	G	NOMS, NORS
Kearney, NE	G	GWSS, CWSS
Lincoln, NE	G	NOMS, NORS
Albuquerque, NM	G	NOMS, NORS
Baldwinsville, NY	G	GWSS, CWSS
Dayton, OH	G	NSP, NOMS, NORS
Aliquippa, PA	G	GWSS, CWSS
Sioux Falls, SD	G	GWSS, NOMS
Memphis, TN	G	NOMS, NORS
San Antonio, TX	G	GWSS, NOMS, NORS

TABLE A-4 SYSTEMS SAMPLED IN MORE THAN ONE SURVEY (continued)

Location	Source	Survey in Which the System was Sampled
Provo, UT	G	CWSS, NOMS
Marysville, WA	G	GWSS, CWSS
Spokane, WA	G	GWSS, NSP, NOMS
Madison, WI	G	NSP, NOMS
Powell, WY	G	GWSS, CWSS
Birmingham, AL	S	NSP, NOMS
Camden, AR	S	NOMS, NORS
Little Rock, AR	S	NSP, NOMS
Phoenix, AZ	S	NSP, NORS
Concord, CA	S	NOMS, NORS
Los Angeles, CA	S	NOMS, NORS
Oakland, CA	S	NSP, NOMS
Sacramento, CA	S	NSP, NOMS
San Diego, CA	S	NOMS, NORS
San Francisco, CA	S	NOMS, NORS
Denver, CO	S	NSP, NOMS, NORS
Pueblo, CO	S	NOMS, NORS
New Haven, CT	S	NSP, NOMS
Waterbury, CT	S	NOMS, NORS
Washington, DC	S	NSP, NOMS, NORS
Atlanta, GA	S	NSP, NOMS, NORS
Davenport, IA	S	NSP, NOMS, NORS
Chicago, IL	S	NSP, NOMS, NORS

TABLE A-4 SYSTEMS SAMPLED IN MORE THAN ONE SURVEY (continued)

Location	Source	Survey in Which the System was Sampled
Fort Wayne, IN	S	NSP, NOMS
Indianapolis, IN	S	NSP, NOMS, NORS
Whiting, IN	S	NOMS, NORS
Topeka, KS	S	NSP, NOMS, NORS
Louisville, KY	S	NSP, NOMS
Boston, MA	S	NSP, NOMS, NORS
Lawrence, MA	S	NSP, NORS
Baltimore, MD	S	NSP, NOMS, NORS
Portland, ME	S	NSP, NOMS
Detroit, MI	S	NSP, NOMS, NORS
Grand Rapids, MI	S	NSP, NOMS
Mount Clemens, MI	S	NOMS, NORS
St. Paul, MN	S	NOMS, NORS
Cape Girardeau, MO	S	NOMS, NORS
Kansas City MO	S	NSP, NOMS, NORS
St. Louis, MO	S	NSP, NORS
Jackson, MS	S	NSP, NOMS
Charlotte, NC	S	NSP, NOMS
Bismark, ND	S	CWSS, NOMS
Omaha, NE	S	NSP, NOMS
Manchester, NH	S	NSP, NOMS
Elizabeth, NJ	S	NSP, NOMS
Passaic, NJ	S	NSP, NOMS, NORS

TABLE A-4 SYSTEMS SAMPLED IN MORE THAN ONE SURVEY (continued)

Location	Source	Survey in Which the System was Sampled
Buffalo, NY	S	NOMS, NORS
Poughkeepsie, NY	S	CWSS, NOMS
Cincinnati, OH	S	NSP, NORS
Cleveland, OH	S	NSP, NOMS, NORS
Columbus, OH	S	NSP, NOMS NORS
Toledo, OH	S	NSP, NOMS
Oklahoma City, OK	S	NSP, NOMS, NORS
Tulsa, OK	S	NSP, NOMS
Corvallis, OR	S	NSP, CWSS, NOMS, NORS
Eugene, OR	S	NSP, CWSS, NOMS
Portland, OR	S	NSP, NOMS
Harrisburg, PA	S	NSP, CWSS
Philadelphia, PA	S	NSP, NORS
Pittsburgh, PA	S	NSP, NORS
Newport, RI	S	NOMS, NORS
Providence, RI	S	NSP, NOMS
Charleston, SC	S	NSP, NOMS, NORS
Huron, SD	S	NOMS, NORS
Chattanooga, TN	S	NSP, NOMS, NORS
Nashville, TN	S	NSP, NOMS, NORS
Brownsville TX	S	NOMS, NORS
Dallas, TX	S	NOMS, NORS
Fort Worth, TX	S	NSP, NOMS

TABLE A-4 SYSTEMS SAMPLED IN MORE THAN ONE SURVEY (continued)

Location	Source	Survey in Which the System was Sampled
Houston, TX	S	NSP, NOMS
Salt Lake City, UT	S	NSP, NOMS, NORS
Annandale, VA	S	NOMS, NORS
Richmond, VA	S	NSP, NOMS
Illwaco, WA	S	NOMS, NORS
Seattle, WA	S	NSP, NORS
Milwaukee, WI	S	NSP, NOMS, NORS
Huntington, WV	S	NSP, NOMS, NORS
Wheeling, WV	S	NOMS, NORS

G = Groundwater.
S = Surface water.

A.2.4 Testing for Biases in the Data

There was concern that the selection of sample sites in some of the surveys was biased toward VOC contamination, which, when combined with other random survey data, would bias the national projections toward a higher estimated frequency of occurrence and/or mean concentration. Purposeful selection of sites having a higher than average probability of contamination with VOC's was in fact the design of the nonrandom portion of the GWSS. It has been suggested that the NORS, NOMS, and NSP surveys also may have involved a bias towards systems known or suspected to be contaminated; however, this could not be confirmed.

Initially, consideration was given to excluding the GWSS nonrandom data for all VOC's when combining the survey data because the analysis presented by Westrick et al. (1983) showed that 21.2% of the random supplies had at least one VOC above its quantitation limit, whereas 27.3% of the nonrandom supplies showed contamination. The nonrandom portion of the GWSS was reported to have a higher frequency of occurrence of VOC's at all concentration levels for both large and small systems. At higher concentrations, there was a two to four times higher frequency of occurrence of VOC's in the nonrandom sample. It is important to note that these comparisons are based on the combined results for all VOC's. There was some concern, however, that sites selected because of suspected

contamination with one or anotner specific VOC would not necessarily be biased for other VOC's. Westrick et al. (1983) did not present a comparison of the random and nonrandom data on a chemical-by-chemical basis, so it was not clear whether it was appropriate to exclude the GWSS nonrandom data for each VOC. It was decided, therefore, that an analysis would be performed on each VOC to evaluate whether there was a statistically significant difference in the frequency of occurrence and in the mean of the positive values observed in the random and nonrandom portions of the GWSS.

To test the difference in frequency of occurrence, the results for each VOC at each site were classified as positive or negative and summarized in a two-way table as shown below:

Portion of GWSS	Results		Total
	Positive	Negative	
Random	n_{11}	n_{12}	$n_{1.}$
Nonrandom	n_{21}	n_{22}	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	$n_{..}$

Comparisons of relative frequencies of positive results in the random and nonrandom segments of the GWSS were based on the X^2 statistic:

$$2\ X^2 = \frac{(n_{11}n_{22} - n_{12}n_{21})\ n_{..}}{n_{1.}\ n_{2.}\ n_{.1}\ n_{.2}}$$

with n's defined in the table above. Under the null hypothesis

that the relative frequencies for the random and nonrandom segments are equal, the test statistic has an approximate chi-square distribution with 1 degree of freedom; this distribution was used to compute the P values for each VOC of interest that was examined in the GWSS.

The mean values of the positive samples in the random and nonrandom portions of the GWSS were compared using the "t" test, which tests the null hypothesis that the two means are the same (i.e., $u_1 = u_2$). The statistic for testing the equality of two population means u_1 and u_2 using independent samples from each are as follows:

$$t = (\bar{x} - x_2) \; / \; s^2 \; (1/n_1 + 1/n_2)$$

for n_1 and n_2 observations with a pooled cariance (s^2) of:

$$s^2 = [(n_1 - 1) \; s_1^2 = (n_2 - 1) \; s_2^2] \; / \; (n_1 + n_2 - 2)$$

where $\bar{x}_1$ and $\bar{x}_2$ are the observed sample means of the two groups and s_1^2 and s_2^2 are the corresponding variances. An underlying assumption for use of the "t" test is that the variables are normally and independently distributed in each group. The normal distribution is not considered an acceptable model for the values of the positive samples because of the positive skewness and large coefficient of variation of their distribution. Therefore, the tests were based on natural logarithms of the concentrations to make the normality assumption more reasonable.

Table A-5 presents the results of the comparison of the frequency of the positives and of the means of the positives for each VOC in the random and nonrandom portion of the GWSS. The P_t value indicates the significance level for evaluating the null hypothesis that the means of the random and nonrandom sample are equal. This value represents the probability that the null hypothesis (i.e., the population means are equal) has been rejected on the basis of the sample means when it is actually true. For example, the results for benzene shown in Table A-5 indicate that there is a 69% probability of being in error if one were to reject the hypothesis that the population means for randomly and nonrandomly selected sites are the same based on the sampling means of 4.1 micrograms/liter for the nonrandom and 6.2 micrograms/liter for the random sample observed. Similarly, the P_X2 value represents the level of significance for evaluating the null hypothesis that the frequency of positives are the same. Again, referring to benzene, the P_X2 value indicates that there is a 15% probability of being in error by rejecting the hypothesis that the frequency of occurrence of benzene in samples selected randomly and nonrandomly in the GWSS are the same based on the observed frequencies of 1.7% for the nonrandom and 0.7% for the random samples.

The critical P value selected for this analysis was 0.01, implying the acceptance of no more than a 1% probability of being in error by rejecting the null hypothesis of equal

TABLE A-5

RESULTS OF COMPARATIVE TESTS FOR RANDOM AND NONRANDOM GWSS DATA

Chemical	NR/R[a]	Number of Positive Samples[b] (n)	Percent Positive[b]	Mean (x) (micrograms/liter)	Standard Deviation (s) (micrograms/liter)	Range (micrograms/liter)	P_t[c]
Benzene	NR	8	1.7	4.1	4.9	0.50-12	0.69
	R	3	0.7	6.2	7.7	0.61-15	
Carbon tetrachloride	NR	15	3.2	2.2	4.2	0.20-15	0.54
	R	15	3.3	1.7	4.0	0.20-16	
Chlorobenzene	NR	1	0.2	2.7	-	-	-
	R	0	0	-	-	-	
o-Dichlorobenzene	NR	2	0.4	2.5	0.35	2.2-2.7	-
	R	0	0	-	-		
m-Dichlorobenzene	NR	0	0	-	-	-	-
	R	0	0	-	-	-	
p-Dichlorobenzene	NR	4	0.9	0.77	0.09	0.70-0.90	0.71
	R	5	1.1	0.75	0.31	0.52-1.3	
1,2-Dichloroethane	NR	7	1.5	3.4	2.9	1.1-9.8	0.01
	R	3	0.7	0.68	0.23	0.53-0.95	
1,1-Dichloroethylene	NR	15	3.2	0.59	0.71	0.22-3.0	0.28
	R	9	2.0	1.4	2.0	0.22-6.3	
cis,trans-1,2-Dichloroethylene	NR	38	8.0	8.5	21	0.21-120	0.06
	R	16	3.5	1.0	0.56	0.21-2.0	
Tetrachloroethylene	NR	43	9.1	4.7	11	0.22-69	0.04
	R	34	7.5	1.5	4.0	0.21-23	
1,1,1-Trichloroethane	NR	50	10.6	2.1	3.3	0.20-21	0.17
	R	27	5.9	1.7	3.6	0.20-18	
Trichloroethylene	NR	61	12.9	9.0	22	0.20-130	0.90
	R	30	6.6	8.1	17	0.24-78	
Vinyl chloride	NR	6	1.3	4.4	2.9	1.4-8.4	0.22
	R	1	0.2	1.1	-	-	

a NR = nonrandom; R = random.
b Based on 456 random and 473 nonrandom samples from supplies serving 25 more people.
c Based on the natural logarithms of the positive concentrations (see tex
d Data may be unreliable due to the small number of positive samples (i.e., <5 for random and/or nonrandom data).

means or equal frequencies. Using a critical value of 0.01, it can be assumed that, for those P values which are less than 0.01, the null hypothesis should be rejected (i.e., for a nul hypothesis $u_1 = u_2$, the alternative hypothesis, $u_1 \neq u_2$, should be chosen). On the other hand, for values that are greater than 0.01, it cannot be directly assumed that the null hypothesis is true. It is only known that the two valued are not significantly different; it is not known if they are statistically the same. However, assuming the null hypothesis is testing for equality between two values, the P values obtained from the test can be used as a general guide to determine how similar the two values are. A higher significance level (e.g., P = 0.70) would denote a greater similarity; a lower level (e.g., P = 0.05) would denote a greater difference.

In evaluating the null hypothesis based on the P value, the size of the sample on which the test of significance is made is also important. For small samples, the null hypothesis is likely to only be rejected if it is very wrong. Consequently, a hypothesis of equality may be accepted when it is wrong because the test for being significantly different at the 0.01 level is more stringent. In contrast, with a large sample small departures from the null hypothesis can be determined to be statistically significant, even though they are quite unimportant in practice.

As can be seen from the results shown in Table A-5,

there is an apparent general trend of higher frequencies or occurrence and higher means in the nonrandom samples. However, there are a few cases where the differences are statistically significant at the 0.01 level. With respect to means of positives, only 1,2-dichloroethane is found to have a statistically significant difference (actually, borderline) in the means (the nonrandom being higher, as would be expected). For the comparison of frequency of occurrence, statistically significant differences were found for only three of the VOC's: cis/trans-1,2-dichloroethylene, 1,1,1-trichloroethane, and trichloroethylene. Again, the frequencies were higher in the nonrandom samples as would be expected. (Note that the small number of positive samples for several of the VOC's makes the analysis of the frequencies of questionable validity.)

Based on these comparisons it was decided that in preparing the national projections, the GWSS nonrandom data would be included except for the four VOC's found to have statistically significant differences in the mean of the positives or frequency of positives.

A.2.5 Establishing a Common Minimum Quantifiable Concentration for the Combined Survey Data

As noted in the discussion on computing averages for VOC's in supplies the minimum quantifiable concentration for an analytical technique defines the level below which it cannot

be determined whether the VOC is present and, if so, at what concentration. If a common minimum quantifiable concentration was used across all surveys, then the national projections would provide estimates of the number of systems at various ranges of contamination levels above the minimum quantifiable concentration. Below that concentration, only a total number of systems would be estimated for the range of 0 micrograms/liter to the maximum possible concentration, i.e., just below the minimum quantifiable concentration. The number of systems having no contamination and the distribution of systems at various levels below the minimum quantifiable concentration could not be determined.

In many cases different minimum quantifiable concentrations were used in the different surveys and for different analyses in the same survey. Consequently some supplies in the combined data set were reported as having no measured VOC (i.e., negative systems) with a maximum possible concentration greater than quantified levels reported for other supplies analyzed with more sensitive methods.

In some cases, it was reasonable to use the highest "negative" value as the common minimum quantifiable concentration above which national estimates of the distribution of systems containing various VOC levels are made and below which only a total number of systems is given. Any positive values below this level were used as evidence that some supplies in

this latter group are contaminated, i.e., all negative values are not necessarily 0 micrograms/liter. These positive values were not included with positive values above the established common minimum quantifiable concentration used for selecting the appropriate model, etc. for estimating national occurrence.

In several instances, it happened that the selection of the highest negative value resulted in many, and sometimes most, of the measured positive values falling below that level, leaving too few positives above that level for completing the national projections. Three major alternatives were considered for handling high negative values in these cases. The first alternative was to establish the common minimum quantifiable concentration at some lower level that did not exceed substantial numbers of positive data points, and treat the high negative values as though they were positives at their maximum possible concentration. This would be generally consistent with the rules for averaging samples described earlier. It was observed, however that in a large number of cases this would result in a substantial number of the resulting total positive values being contributed from data where the VOC was not actually observed; in some cases this contribution would exceed the number of actual positive values. This was determined to be an improper use of the data.

The second alternative was to simply discard the high

negative data. This was also considered inappropriate since it meant essentially eliminating valid data because it could not be made to fit the projection methodology. Furthermore, its elimination would artificially raise the computed frequency of positives by lowering the total number of systems sampled.

The third alternative, and the one selected for the national projections, was to retain the data, but treat them as though they were negative data below the lower common minmum quantifiable concentration selected for the combined survey data. While it is true that VOC's could be present in those supplies at a higher concentration, the other data suggested that the probability is greater that they are not. This alternative also avoids the elimination of valid data from the analysis. (It should be noted that, while this third alternative is used for the reporting national projections projections were also calculated using the first alternative to give the more conservative estimate; these results are noted in the text.)

A.2.6 Model Selection

In developing methods for estimating the numbers of drinking water systems and people affected by different pollutant levels, it was necessary to select a statistical model for the distribution of positive concentrations for each data set. The following considerations guided the selection of a model:

- The same model should be applicable to data from different system size groups for a given pollutant (to facilitate comparisons and evaluation of estimation error).
- A continuous distributional model should be used where appropriate (to smooth out random variations in relative frequency among concentration intervals).
- The appropriateness of any continuous model should be checked through goodness-of-fit tests.
- Estimates from continuous models should provide upper bounds on the upper tail of the observed distribution (to avoid underestimating the number of systems with high-level contamination).

Three types of continuous models were investigated:

- Statistical distributions: e.g., the lognormal (Aitchison and Brown, 1957, and gamma (Johnson and Kotz, 1970) distributions
- Transformations: the Johnson (1949) system of transformation to normality
- Empirical models: fitting the cumulative frequency with a polynomial function of contration

The adequacy of different distributional models was tested by three goodness-of-fit tests: Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling (Stephens 1974). A model was considered unsatisfactory if it failed more than one of these tests at the 0.05 significance level. None of the three types of models worked consistently for different pollutant size-group data sets.

Of the models investigated, the delta distribution based on the lognormal distribution (Aitchison and Brown, 1957) seemed the most appropriate a priori. This model has been used successfully in a wide variety of water contamination problems. It allows for the positive probability of zero (not detected) results and the skewed distribution of positive results that generally characterize the drinking water data. However, in some cases this model failed the goodness-of-fit tests for data from at least one size group. In other cases there were insufficient positive results to test whether the model was appropriate.

Based on the evaluations described above, it finally was decided that no continuous model could be identified that would be useful for the drinking water data. Therefore, a discrete model had to be employed. The model adopted was the multinomial distribution (Johnson and Kotz, 1969), in which the proportion of the distribution in a specified concentration interval is estimated by the observed relative frequency for that interval. The intervals used were the ones of interest in the evaluation (MQC*, MQC-5 micrograms/liter, 5-10 micrograms/liter, 10 micrograms/liter intervals from 10 to 100 micrograms/liter, and >100 micrograms/liter).

---

* MQC is the Minimum Quantifiable Concentration set for the combined survey data (see Section A.2.5, above).

Estimating System Size Groupings

Consideration was given to developing estimates from data grouped by system size because it was thought that contamination might be more likely in larger systems located in more populous and probably more industrialized areas. The five system size groups shown below formed the starting point for grouping data by system size:

Group	Size Range (number of people served)
1 (very small)	1-500
2 (small)	501-3,300
3 (medium)	3,301-10,000
4 (large)	10,001-100,000
5 (very large)	>100,000

Frequency tables showing the number of systems sampled and the numbers of positive and negative results for each of these five size groups were produced for groundwater and surface water systems for each pollutant. It was considered desirable to consolidate these groups as much as possible (consistent with the data) because relatively few systems were sampled in some of the groups

The extent of further consolidation possible was evaluated by comparing the relative frequencies of positives in different groups through a statistical test procedure. Group relative frequencies were compared in the following order, and the groups were combined when no significant difference in

relative frequency of positives was found:

1. Groups 1 and 2
2. Groups 4 and 5
3. Group 3 with groups 2 and 2 combined
4. Groups 1, 2, and 3 combined with groups 4 and 5 combined

Step 3 was done only if groups 1 and 2 could be combined as a result of step 1; step 4 was done only if combinations in previous steps were possible. The order of the comparisons was chosen based on the possibility of a relationship (trend) between system size and the percentage of positive systems.

In performing the statistical test for equal percentages of positives in two size groups, the first step was to form a 2 x 2 summary table as illustrated below for groups 1 and 2:

	Test Results		
Size Group	Negative	Positive	Total
1	a	b	a + b
2	c	d	c + d
Total	a + c	b + d	n

where

a = the number of negative results for group 1.
b = the number of positive results for group 1.
a + b = the number of systems sampled in group 1.
n = a + b + c + d, the number of systems sampled in both groups, etc.

The statistic was:

$$x^2 = \frac{(|ad - bc| - 1/2n)^2 n}{(a + b)(a + c)(b + d)(c + d)}$$

When this statistic exceeded the critical value 3.84 (the 95th percentile of the chi-squared distribution with 1 degree of freedom), the hypothesis of equal relative frequencies of positives in the two groups was rejected, and the groups were not combined. The chi-squared test for homogeneity of percentages in two populations is discussed by Snedecor and Cochran (1967).

In some cases, the expected number of positive systems was too small for the x^2 test to be used (e.g., $(a + b)(b + d)/n < 5$ for group 1 under the hypothesis of equal proportions positive). In such cases, the usual alternative procedure, Fisher's Exact Test, was used to test equality of proportions positive. The application of Fisher's test is described in Ostle (1963).

° Projections of National Occurrence

After the final system size groups were selected for a pollutant, the national projections were computed. First the proportion of systems in each concentration interval of interest was estimated for each size group by the observed relative frequency for the sampled systems. The proportions of systems above the different specified concentrations also were estimated for each size group (again by the observed relative frequency for sampled systems). For example, if 100 systems were sampled in a given size group and three were found in the range 40 50 micrograms/liter, the estimated percentage of systems in

that range was 3%; if ten out of the 100 systems sampled were above 50 micrograms/liter, the estimated percentage of systems above 40 micrograms/liter was 10%.

The number of systems above each concentration limit for a given size group was estimated by multiplying the observed relative frequency (p_i) by the total number or systems (N_i) in that group. Then the total number of systems of all sizes above a given concentration was estimated by the sum of estimates for the k individual system size groups:

$$m = \sum_{i=1}^{k} N_i \, P_i .$$

It can be shown based on the multinomial model that m is an unbiased estimator of the total number of systems above a specified concentration (M), and that m has variance

$$Var(m) = \sum_{i=1}^{k} N_i^2 \, \frac{N^i - n_i}{N_i - 1} \quad \frac{P_i \, (1-P_i)}{n_i}$$

where n_i is the number of systems sampled out of the N_i systems in group i and P_i is the true percentage of systems above the specified concentration in group i. Following Cochran (1963), the term $(P_i \, (1-P_i)/n_i$ in Var(m) was replaced with its unbiased estimator, $P_i(1 - P_i)/(n_i - 1)$, to estimate Var (m). Then an approximate 95% confidence interval on M was calculated from

$$m \pm 1.96 \, [Var \, (m)]^{1/2}.$$

The accuracy of this interval improves with increasing sample size (n_i) and is bettern when P_i values are not close to zero.

When confidence limits obtained from the approximation were outside feasible M values (i.e., less than zero or greater than $N = N_i$), the limits were reset to the nearest feasible value.

Projections for numbers of people exposed to concentrations above specified limits were computed in the manner described above, letting N_i represent the number of people served by systems in the ith system size group.

B.1 SECTION B

B.1.1 System Size Grouping and Selection of Model for Estimating National Occurrence of 1,1-Dichloroethylene in Drinking Water Supplies

As indicated in Section A (A 2.6), an initial step in the estimation process was to group the data based on system size. Based on results of statistical tests comparing percentages positive in five initial size groups, groundwater data was collapsed into two groups:

- ≤10,000 persons served
- >10,000 persons served

(Summary statistics for initial and final groundwater groups are given in Table B-1. Results of the test are shown in Table B-3.) For surface water, data was collapsed into two groups:

- ≤10,000 persons served
- >10,000 persons served

(Summary statistics and test results are given in Tables B-2 and B-3, respectively.) Only one surface water system was sampled in groups 1-3 therefore, these groups were not combined with 4 and 5 because of insufficient data to test for differences in proportions positive.

For both groundwater and surface water data, the next step was to fit a delta distribution to each system size

grouping with sufficient positive data. The delta distribution has cumulative distribution function:

$$P(X \leq x) = \begin{cases} 0 & , x < 0 \\ & , x = 0 \\ + (1-) \quad F(x) & , x > 0 \end{cases}$$

with

$$F(x) = \int_{-\infty}^{(\log_e x -)/} f(z)dz,$$

where $f(z)$ is the standard normal probability density function. The mean and standard deviation of $\log_e$-transformed data were used to estimate the parameters and ; was estimated by the observed proportion of negative values. The goodness-of-fit of the delta distribution was tested, and the model was adopted if no more than one of the three tests failed for each size group. The results in Table B-4 show that the model failed for groundwater, and there was insufficient data to test the model for surface water. Therefore, the multinomial model was used as the basis for estimation.

TABLE B-1

FREQUENCY OF POSITIVE SYSTEMS BY SIZE GROUPS
FOR GROUNDWATER SYSTEMS

System Size		Number Negative	Number Positive	Number Sampled	Percent Positive
Initial Groups					
(1)	1-500	228	2	230	0.87
(2)	501-3,300	199	4	203	1.97
(3)	3,301-10,000	150	4	154	2.60
(4)	10,001-100,000	298	14	312	4.49
(5)	>100,000	38	1	39	2.56
	All	913	25	938	2.67
Final Groups					
(1,2,3)	≤10 000	577	10	587	1.70
(4,5)	>10,000	336	15	351	4.27
	All	913	25	938	2.67

TABLE B-2

FREQUENCY OF POSITIVE SYSTEMS BY SIZE GROUPS
FOR SURFACE WATER SYSTEMS

System Size		Number Negative	Number Positive	Number Sampled	Percent Positive
Initial Groups					
(1)	1-500	--	-	0	--
(2)	501-3,300	--	--	0	--
(3)	3,301-10,000	1	0	1	0.00
(4)	10,001-100,000	19	0	19	0.00
(5)	>100,000	81	2	83	2.41
	All	101	2	103	1.94
Final Groups					
(1,2,3)	≤10,000	1	0	1	0.00
(4,5)	>10,000	100	2	102	1.96
	All	101	2	103	1.94

TABLE B-3

CHI-SQUARED TESTS COMPARING
SYSTEM SIZE GROUP PROPORTIONS POSITIVE

System Size Groups Compared	Groundwater			Surface Water	
	No. of Systems	Test Stat.	Decision[a]	No. of Systems	Test Stat. Decisions[a]
1 vs. 2	433	b	Pass	0	Not done
4 vs. 5	351	b	Pass	102	b
1,2 vs. 3	587	b	Pass	0	Not done
1,2,3 vs 4,5	938	4.65	Fail	103	Not done

[a] Critical value for x^2 test ($a = 0.05$) is 3.84.

[b] Expected number of positives too small for X^2 test. Used Fisher's Exact Test (Ostle, 1963).

TABLE B-4

GOODNESS-OF-FIT TESTS FOR DELTA DISTRIBUTION
1,1-DICHLOROETHYLENE

System Type	System Size (population served)	Parameter Estimates[a] N			Anderson-Darling[b] Test Stat.	Decision	Cramer-von Mises[c] Test Stat.	Decision	Kolmogorov-Smirnov[d] Test Stat.	Decision
Groundwater	≤10,000	587	0.983	-0.42	4.04	(Insufficient data for test)				
	>10,000	351	0.957	-0.91	4.27	1.010	Fail	0.152	Fail	0.965	Fail
Surface Water	≤10,000	1	1.000	--	--	(Insufficient data for test)				
	>10,000	102	0.960	-1.14	1.74	(Insufficient data for test)				

[a] N = Number of systems sampled; and are estimates of the delta distribution parameters.

[b] Critical value for 0.05 level of significance = 0.787.

[c] Critical value for 0.05 level of significance = 0.126.

[d] Critical value for 0.05 level of significance = 0.895.