November 1984

RESEARCH AND DEMONSTRATION OF IMPROVED METHODS FOR CARRYING OUT BENEFIT-COST ANALYSES OF INDIVIDUAL REGULATIONS

by

David Harrison, Jr.
Harvard University
Principal Investigator

VOLUME IV

STRATEGIES FOR DEALING WITH UNCERTAINTY IN INDIVIDUAL REGULATIONS

Final Report
Submitted to U.S. Environmental Protection Agency
Cooperative Agreement 68-809-702-01-0

Project Officer

George Provenzano
United States Environmental Protection Agency
Washington, D.C. 20460

Energy and Environmental Policy Center
John F. Kennedy School of Government
Harvard University
Cambridge, Massachusetts 02138

DISCLAIMER

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

VOLUME IV

TABLE OF CONTENTS

LIST OF TABLES

PART 9: AN OVERVIEW OF SCIENTIFIC UNCERTAINTIES IN BENEFIT ESTIMATION

PART 10: THE VALUE OF IMPROVED EXPOSURE INFORMATION IN BENEFIT-COST ANALYSES OF TOXIC SUBSTANCES

PART 10 (continued)

PART 11: THE VALUE OF ACQUIRING INFORMATION UNDER SECTION 8(a) OF THE TOXIC SUBSTANCES CONTROL ACT: A DECISION-ANALYTIC APPROACH

LIST OF FIGURES

PART 9: AN OVERVIEW OF SCIENTIFIC UNCERTIAINTIES IN BENEFIT ESTIMATION

PART 10: THE VALUE OF IMPROVED EXPOSURE INFORMATION IN BENEFIT-COST ANALYSES OF TOXIC SUBSTANCES

PART 11: THE VALUE OF ACQUIRING INFORMATION UNCER SECTION 8(a) OF THE TOXIC SUBSTANCES CONTROL ACT: A DECISION-ANALYTIC APPROACH

PART 11 (continued)

PART 9

AN OVERVIEW OF SCIENTIFIC UNCERTAINTIES IN BENEFIT ESTIMATION

John S. Evans
Katherine Walker

I. INTRODUCTION

The basic components of quantitative risk assessment are exposure assessment and hazard assessment. Exposure assessment involves estimation of the concentrations of pollutants to which individuals or populations are exposed. Hazard assessment is concerned with estimation of the health risks associated with given patterns of exposure.

For an exposure to occur one or more persons must come into contact with the pollutant. Therefore, exposure assessment involves (at least implicitly) not only determination of the spatial and temporal pattern of pollutant concentrations but also analysis of human activity patterns. See, for example, Ott (1980) and Duan (1981). The spatial and temporal aspects of the field of pollutant concentrations are typically estiamted by measurement, modelling or some combination of the two.

Hazard assessment is the determination of the health risks posed by exposure to the pollutant concentrations obtained from the exposure analysis. The steps involved in hazard assessment depend largely on the type of toxicity being considered (carcinogenic or noncarcinogenic) but generally include

estimation of dose and the toxic potency of the pollutant. Determination of the human potency of pollutants itself may involve several steps -- e.g., determination of animal potency, extrapolation from animal to man, and extrapolation from high to low dose.

Each step in risk assessment involves some uncertainty, and the uncertainties in each phase of the analysis combine to produce a final risk estimate which is uncertain. Several methods are available for analysis of the propagation or cascading of uncertainty. To illustrate the propagation of error, we introduce a simple model which has been used by Crouch and Wilson (1981) to estimate the risk:

$$P = \alpha C \beta \tag{1}$$

where β is the human potency (cases per mg/kg per day), C is the concentration of the pollutant in the media of interest to which people are exposed (mg/m^3 in air or mg/1 in water), and α is the parameter which relates exposure to dose rate ((mg/kg per day) per mg/m^3 in air or (mg/kg per day) per mg/1 in water).

In a simple multiplicative model such as this, the propagation of error may be analyzed with Gauss' Law of Error Propagation:

$$\sigma_p^2 \approx (C\beta)^2 \sigma_\alpha^2 + (\alpha\beta)^2 \sigma_c^2 + (\alpha C)^2 \sigma_\beta^2 \tag{2}$$

which may be re-expressed as:

$$\left[\frac{\sigma_P}{P}\right]^2 \approx \left[\frac{\sigma_\alpha}{\alpha}\right]^2 + \left[\frac{\sigma^2}{C}\right]^2 + \left[\frac{\sigma_\beta}{\beta}\right]^2 \qquad (3)$$

Thus if $(\sigma_\alpha/\alpha) = 0.05$, $(\sigma_C/C) = 0.10$, and $(\sigma_\beta/\beta) = 0.30$:

$$\frac{\sigma_P}{P}\left[0.0025 + 0.0100 + 0.0900\right]^{\frac{1}{2}} = 0.32 \qquad (4)$$

In risk assessment, however, it is common for uncertainties to be large in comparison with central estimates of parameters. Although, as Seiler (1982) has demonstrated, Gauss' Law of Error Propagation can be extended to cover large errors, a more common form of error analysis involves the use of lognormal distributions to characterize key parameters. See, for example, Crouch and Wilson (1981). If the uncertainties in α, C and β are characterized as lognormal, then

$$\sigma^2_{\ln P} = \sigma^2_{\ln\beta} + \sigma^2_{\ln C} + \sigma^2_{\ln\beta} \qquad (5)$$

And, if for example α is known to within a factor of 1.2, C is known to within a factor of 2.0, and β is known to within a factor of 5.0 (i.e., $\sigma_{\ln\alpha} = 0.18$, $\sigma_{\ln C} = 0.69$ and $\sigma_{\ln\beta} = 1.61$), then:

$$\sigma_{\ln P} \approx \left[0.0324 + 0.4761 + 2.5921\right]^{\frac{1}{2}} = 1.7609 \qquad (6)$$

Equation (6) indicates that the risk, P, could be estimated to within a factor of $e^{1.76}$ or 5.8 Both (4) and (6) illustrate an

important general feature of uncertainty analysis -- when one component of error is large relative to other components it may dominate the uncertainty in the final risk estimate.

Of course (1) is highly simplified. A somewhat more comprehensive model might be:

$$P = (\alpha_a C_a + \alpha_w C_w)\ EK_{ha}\ \beta_a \qquad (7)$$

Here, both air and water exposures are considered as well as the possibility that a potency factor may not be available from epidemiological analysis of human exposures in the range of concentrations of interest. In this case, a potency factor from animal bioassay, β_a, may have to be used in conjunction with an interspecies conversion factor, K_{ha}, and a low dose extrapolation factor, E. Recognizing that $\alpha_a C_a + \alpha_w C_w$ is simply the dose rate, d, (7) could be re-expressed as a purely multiplicative model in which error propagation could be analyzed using (3) or (5).

Other methods of error analysis are available for more complex models of risk. See, for example, Fiering et al. (1982). However, it is not our intent to review methods of error analysis, but rather to use these simple approaches to analyze and illustrate the propagation of uncertainty in risk assessments of toxic compounds in air and water.

The sections which follow discuss the components of typical risk assessments for toxic compounds: environmental transport and fate modelling, dose estimation, epidemiology and animal

bioassay. The focus in each section is on sources of uncertainty. In addition, the problems of interspecies conversion and low-dose extrapolation and the attendant uncertainties are briefly reviewed.

II. EXPOSURE ASSESSMENT

Exposure assessment often begins with estimation of the concentrations in air or water which are expected to result from specified patterns of emissions. The primary tools used in such estimation are air and water pollution models. The basic air and groundwater models are discussed below.

Air Pollution Dispersion Models

Air pollution dispersion models provide a link between pollutant emissions and exposures estimates necessary for risk assessment. In order to be most useful, the models should provide data in an appropriate form and for the same conditions likely to be encountered in situations requiring risk assessment;

- short and long term averaging periods
- receptors near and far from source
- urban and rural areas
- complex (mountains, valleys, near large bodies of water), and smooth terrain
- non-reactive and reactive pollutants.

The ability of air pollution dispersion models to provide accurate concentration estimates under all the conditions listed above is limited. Gaussian dispersion models in particular were developed with certain conditions in mind, and provide the best results when these conditions are met. By examining briefly the mathematical and scientific basis for dispersion models, this section attempts to promote an understanding of the conditions

under which dispersion models are most valid. The section concludes with estimates of model accuracy reported in the scientific literature.

Air pollution dispersion models may be divided into two broad categories -- those suitable for estimating exposures near to the source and those suited for predicting exposures at long distances from the source. The meteorological conditions and processes which govern pollutant transport and dispersion differ in the near and far field. The exact dividing point between short and long range is not well defined, but has been estimated by Turner (1979) to be roughly 50 km.

Gaussian dispersion models are most commonly used for modeling concentrations within short ranges of a source. The UNAMAP (User's Network for Applied Modeling of Air Pollution) series of Gaussian models has been widely used to help current or proposed sources meet regulatory requirements under the Clean Air Act (CAA). Advection-dispersion models based on numerical approximation solutions to advection dispersion equations have also been developed and are theoretically better able to simulate three-dimensional dispersion than their Gaussian counterparts. However, they generally require more complete data, are not as routinely available, and have not been verified in as many field situations as the Gaussian models (Mahoney, 1979, and Turner, 1979). They will therefore not be discussed further in this review.

Table 1. Domains of Validity of Air Pollution Transport Models

SPACE	
Long range (regional, mesoscale)	>50 km
Near-field/short-range	<50 km
TIME	
Short term	3 hr, 24 hr
Long term	month, year

The discussion of long range transport models will focus on the Lagrangian (moving coordinate system) rather than Eulerian (fixed coordinate system) numerical advection-dispersion models, again because of the relative limited availability of such models for routine use (Bass, 1981). Several Lagrangian models are currently operational and easily applied to long range transport problems (Bass, 1981).

Gaussian Dispersion Models

The Gaussian dispersion model has been called the basic "workhorse" of air pollution dispersion modeling (Hanna, et al., 1982). With the passage of the 1970 and 1977 Clean Air Act amendments, the requirements for dispersion modeling to maintain air quality and to support new source permits has increased. This increasing demand for modeling has had the beneficial effects of spurring the development of the UNAMAP series and encouraging better validation and field verification of the models. However, the modeling demands have also created the temptation to apply models to situations for which they are not particularly valid and to use their results without proper regard for their accuracy. The purpose of this discussion is to develop an understanding of the accuracy of the Gaussian dispersion models as applied to various field conditions.

The basic form of the Gaussian model (for a conservative pollutant with complete reflection at the ground surface) is:

$$C(x,y,z;H)=\frac{Q}{2\pi\sigma_y\sigma_z\bar{u}}\exp\left\{-\frac{y^2}{2\sigma_y^2}\right\}\left[\exp\left\{-\frac{(H-z)^2}{2\sigma_z^2}\right\}+\exp\left\{-\frac{(H+z)^2}{2\sigma_z^2}\right\}\right] \quad (8)$$

where

C is the time averaged concentration of pollutant (mass/length^3);

x, y, and z are the distances downwind, crosswind, and vertically upward, respectively (length);

H is the effective source height above ground level (H is equal to the sum of the physical stack height h_s and the plume rise ΔH) (length);

Q is the source strength (mass/time):

σ_y is the standard deviation of the time-averaged plume concentration distribution in the vertical direction (length);

$\bar{u}$ is the time-averaged wind speed at the level H (length/time, usually 10 minute or one hour averages) (Spengler, et al., 1982).

The equation states that the concentration observed at a given distance from a source is a function of the initial peak center line concentration, the height of the plume and the horizon and vertical spread of the plume. The model assumes complete reflection of the plume at the ground surface.

The amount of plume rise, ΔH, is determined by the plume's initial momentum and buoyancy and by the stability of the atmosphere. The formulae for calculating plume rise require input of stack physical dimensions, effluent exit velocity,

temperature and density, atmospheric stability data, and mean wind speed at the height of the stack.

The dispersion parameters σ_y and σ_z, describe the spread of the plume in the horizontal and vertical directions as a function of distance from the source and atmospheric turbulence, although dispersion is a function of both local and larger scale atmospheric turbulence. For modeling purposes, turbulence is more commonly characterized in terms of atmospheric stability; the greater the stability, the less vertical mixing of the atmosphere and therefore the less turbulence. Atmospheric stability has been divided into seven classes (A-G) ranging from extremely unstable (A) to extremely stable (G). Values of σ_y and σ_z as a function of stability class have been documented in the form of standard curves, the most common of which are those developed by Pasquill-Gifford and Turner.

The average wind speed, $\bar{u}$, at stack height is usually not measured directly but is estimated from wind speed at 7 to 10 meters above the ground using the power law formula:

$$u_z = u_{10} \left[\frac{z}{10}\right]^p \qquad (9)$$

where,

z = height to which wind speed extrapolated

p = parameter which varies as the function of stability class and site (urban/rural)

u_{10} = wind speed at 10 meters usually obtained from regional weather station.

The Gaussian model is based on the basic diffusion equations and a set of simplifying assumptions which are presented in Table 2. With a few exceptions, these assumptions are not easily classifiable as "protective" or "unprotective". The conservation of mass assumption, unless modified by the use of a decay term tends to result in the over-estimation of concentrations. The use of simple stability-based dispersion parameters in conditions for which dispersion is dominated by other turbulence factors (complex terrain, layer scale atmospheric motion) may also result in the over-estimate of ambient concentrations. The effect of the other assumptions is dependent upon the conditions under which the model is applied.

Most Gaussian dispersion models require similar meteorological and emissions data. The meteorological data required include wind speed, direction, stability class and mixing layer depth. Models used for predicting long-term average concentrations may require monthly or yearly windroses for speed and direction and stability classes. Precipitation records may also be necessary to assess the probability of pollutant wash out near the source. The emissions data required includes physical stack parameters (height, diameter), stack gas exit velocity and temperature, and source strength.

Standardized current and historical measures of surface level and upper air meteorological parameters are available from a number of national data bases. The National Climatic Center in Ashville, North Carolina maintains current and historical records of measured values of wind speed and direction, stability

Table 2. Basic Assumptions of Gaussian Model

	Assumption	Comment
1)	Plume centerline originates at a point, to equilibrium height and is parallel to ground	
2)	No mass is lost to ground or by conversion	Conservative
3)	Wind speed is uniform, parallel to ground and constant	
4)	No dispersion/diffusion in direction of wind flow	
5)	Concentrations decay from plume centerline concentration in Gaussian (bivariate normal) manner in the y and z directions	Conservative under certain conditions
6)	Steady-state emissions	
7)	Concentrations described by Gaussian form are time-averaged concentrations (usually over 10 min. or 1 hour)	

Sources: Spengler, et al. (1982) , Mahoney (1974)

classes, and upper air parameters. The Center also tabulates windroses (speed, direction, stability) over five and ten year periods for use in predicting long-term average concentrations. The National Weather Service (NWS) records wind speed and direction, temperature, pressure dewpoint, and precipitation on an hourly basis at stations throughout the country. Upper air observations of temperature, humidity, wind speed and direction, and height of pressure surface levels are obtained from approximately 130 radiosonde (radio transmitters carried aloft by balloons) and 70 Pilot balloon (PIBAL) stations. Since 1969, the Environmental Meteorological Support Units (MSU) set up by the NWS in cooperation with the EPA have provided upper air observations in 17 major urban areas. Although these sources of data may not be suitable for all site-specific applications of Gaussian models, they provide a first approximation of meteorological conditions for most sites.

On-site measurements can be taken to supplement these data if necessary. For instance, when sources are located in complex terrain -- valleys, mountains, or near large bodies of water -- regional meteorological data are more likely to be unrepresentative of conditions at the site. On-site data do not always assure greater reliability. Disadvantages of on-site data include potential lack of standardized instrument exposure, maintenance and calibration, and the lack of long historical records of climatological parameters.

Gaussian dispersion models are most valid when estimates of pollutant concentrations relatively near the source are needed. Mahoney (1974) suggested that the most appropriate application of

Gaussian models is within 250 meters to 20 km of a source (see Table 3). At distances close ($\leq$250 m) to the source, models are less capable of accounting for local turbulence effects created by the presence of the source (building wake and down wash). At increasing distances from the source, large scale motions in the atmosphere begin to play a more important role in plume transport and dispersion; these effects are generally not reflected in the dispersion estimates based on local classification of stability.

Gaussian dispersion models have been successfully applied to both single and multiple source problems. Turner (1979) categorizes the UNAMAP models into four groups; models for continuous elevated releases over relatively level countryside; models for elevated releases in urban areas over relatively level terrain; models developed for non-level terrain; and finally, models for transportation sources. Modeling pollutant dispersion in non-level or complex terrain is still one of the greatest weaknesses of Gaussian dispersion models, particularly when estimates of short term, peak exposures are necessary.

Both short- and long-term average pollutant concentrations have been estimated using Gaussian dispersion modeling. The Environmental Protection Agency's National Ambient Air Quality Standards require averaging periods as short as 1 hour for ozone and carbon monoxide concentration that are not to be exceeded more than once a year. Other standards, such as the nitrogen dioxide standard, require that concentrations be averaged over one year.

Table 3. Validity of Gaussian Model vs. Domain in Which it is Applied

Distance	Applicability	Comments/Reasons
<250 m	Questionable usefulness	Effect on turbulence of local urban influences (building wake effects, etc.).
250 m - 2 km	Some usefulness	Best reliability of σ_z, σ_y estimates.
2 - 10 km	Best usefulness	
10 - 20 km	Some useful	Projections for σ_z, σ_y less well founded.
> 20 km	Less useful	Little data available for verification of dilution rate estimates. Rate of vertical plume spread diminishes resulting in further transport of plume.
> 100 km	Not appropriate	Concentrations controlled by details of wind trajectories and vertical mixing.

Source: Mahoney (1974).

There are three major sources of error in model predictions. One is the representativeness of the model of real conditions -- the ability of the model to account for complexities in emission, transport, dispersion and chemical transformation or decay of the plume that differ from the relatively simple conditions under which the model was developed. The second major source of uncertainty in model predictions lies in the accuracy, resolution and representativeness of available meteorological and emission data. Finally, the world is stochastic and any individual prediction may be different from the actual state of nature simply by chance.

A primary limitation of the model is that the dispersion parameters as a function of stability class do not adequately account for dispersion that occurs as a result of processes in the upper atmosphere, changes in meteorological conditions at greater distances from the source, and in uneven terrain. The Gaussian model assumes continuous flow parallel to the ground and dispersion parameter estimates are based on empirical observations of non-buoyant emissions over short time periods, low elevations, and stability conditions close to the source (Smith, 1980). Consequently, Gaussian models do not generally represent flow and dispersion well for tall stacks (>150 m), in mountains or valleys, in urban areas, or near land water interfaces (Turner, 1979). There are several methods, other than Pasquill-Gifford, for deriving stability class and dispersion parameters. Although these are more accurate, they require data which are not often available.

Because Gaussian models were developed for steady-state emissions and continuous flow, they generally do not accurately describe exposures from short term releases, Estimates for 1 hour or 24 hour peak exposures are then likely to be less accurate than longer term average exposures from continuous releases. The continuous flow assumption also means that Gaussian models do not describe dispersion well under calm or near calm conditions, or when the wind direction is highly variable.

Gaussian models do not simulate well loss of a pollutant through processes of chemical and physical transformation, decay, rainout, adsorption and impaction. Some of the UNAMAP models have attempted to account for losses through all mechanisms by incorporating a single exponential decay term (Turner, 1979). This inability to account for pollutant losses is one of the most important sources of uncertainty for models dealing with large residence times (large regions or long averaging times); although dispersion dominates pollutant concentrations in the early stages after emission, chemical and physical transformations play an increasingly important role after a few hours or days (Mahoney, 1974, Smith, 1980). Consequently, unless models are calibrated, the assumption of no or limited loss for substances which are in reality reactive is likely to be quite conservative.

While meteorological data necessary to run Gaussian dispersion models is readily available from national and regional databases, the data are not necessarily representative of the site being modeled. The differences between vertical wind speed

profile, the prevailing winds, and stability classes may be quite marked even over distances of a few miles. Meteorologic data taken at stations located near large bodies of water or at rural locations can be particularly unrepresentative of conditions at inland or urban sites.

Wind direction is one of the most significant sources of error in model predictions, particularly of short term exposures or exposures at specific receptor sites. Longer term average exposures are determined on the basis of historical windrose tabulations and tend to be more accurate. Inaccuracies in predictions of wind direction may result from using regional rather than site-specific meteorological data or from use of surface wind data to predict wind direction at the height of the plume centerline. Wind direction changes as a function of height, so surface wind data may not accurately describe the direction of plume transport, especially in the case of highly elevated sources.

Errors in wind speed can also contribute to uncertainty in exposure predictions. The use of power law calculations to extrapolate from wind speeds measured at 7-10 meters to several 10's of meters can be inaccurate especially for short averaging periods (Turner, 1979).

Several authors have studied the accuracy of various Gaussian models (Table 4). The table is not exhaustive; it provides rough estimates of accuracy for best and worst conditions. The estimates of accuracy range from 10-20% for "research grade" predictions to over 100% for more typical applications.

Table 4. Error Estimates for Gaussian Models

Condition	Percentage Error	Source
1) Research grade observations from near surface sources, flat terrain	10-20%	Pasquill (1979)
2) Most "real world" applications to elevated plumes with few supporting meterologic data		
a) 1 hour average	100%	Pasquill (1974)
	>100%	Ruff (1980) Bowne (1981)
b) long term average	≤20%	Pasquill (1974)
c) urban regional average concentrations	±20%	Mahoney (1974)
3) Multiple source model (various conditions)	20-100%	Mahoney

Sources: American Meterological Society (1981)
Mahoney, J.R. (1974)

As might be expected, predictions become more accurate as averaging times and areas increase. These results were presented in most cases without specific reference to the pollutants being modeled, but are probably most representative of inert tracers. It is likely that errors would be greater for applications involving highly reactive pollutants.

Long-Range Transport Models

The need for long range transport models has been growing in recent years because of concerns about the role of long distance transport of sulfur and nitrogen oxides in acid rain and the contribution of distant sources to the degradation of visibility in our national parks (Bass, 1981, Hanna, et al., 1982). Long range transport models may also have an important role to play in risk assessment.

Long-range transport models are not as well developed as Gaussian dispersion models. One reviewer noted that there is "no definitive model for long-range transport and diffusion" (Hanna, et al., 1982). The models are in their "adolescent" stage, "promising, but not mature" and therefore have a somewhat limited history of application (Bass, 1981).

Of the long-range transport models that have been developed, the Lagrangian models appear to be most "mature". The advantages of Lagrangian models over other long-range transport models (e.g., numerical advection-dispersion models) are that they are operational, readily available and relatively easy for the non-specialist to understand and apply. Models for predicting short-term and long-term average concentrations are both available.

In Lagrangian models, pollutant concentrations are characterized with reference to Lagrangian coordinates. Changes in concentration over time and space are described as if from the perspective of the moving parcels of air rather than in relationship to fixed x, y and z axes.

> "The Lagrangian variable trajectory plume model represents a continuous plume emitted by a point source by the transport and dispersion of a succession of discrete plume elements (air parcels or massless trajectory points). These plume elements are advected and diffused by a spatially and temporally varying wind field. Each plume element carries an independent time history -- plume chemistry, dry deposition and scavenging. The time-average ground-level impact of the continuous plume at a given point is simulated by combining the contributions from all elements that independently traverse that point during the specified averaging time." (Bass, 1981).

Proponents claim that the models are able to simulate both small scale diffusion and the large scale meander believed to be largely responsible for long-range transport and dispersion (Hanna, et al., 1982). However, there appears to be little agreement on how these processes can or should be incorporated into the models (Bass, 1981).

A common form of Lagrangian dispersion models is the "Puff" model in which the continuous plume is represented by a series of puffs. The model assumes that: (1) puff diffusion is similar to plume diffusion; (2) at long ranges, plume dispersion may be described as some function of the Pasquill-Gifford dispersion parameters or as a function of time (e.g., σ_y = 0.5 t (sec), (Heffter, 1980); (3) pollutants are confined to the mixing layer.

The meteorological data required by the puff model are similar to those required by the Gaussian model, although they are used differently. Wind speed and direction data from surface and radiosonde stations are used to generate windfields. Windfields describe the spatial and temporal variations in wind speed and direction and are used to predict the transport and dispersion of the plume. When measured data are not available for regions within the windfield, data interpolated from adjacent regions must be used. Finally, the models require estimates of mixing layer depths over the model region.

As with Gaussian models, uncertainty in the Lagrangian model predictions are related to the ability of the model to mathematically simulate complex, real-world conditions and the availability of reliable data. An important source of uncertainty in the models is that the parameters used to describe diffusion and dispersion are still quite crude; they are not well founded theoretically or empirically. They cannot account well for the effect of windshear resulting from vertical variations in wind speed and direction. Dispersion parameters are consequently likely to underestimate dispersion. Model predictions are particularly sensitive to uncertainties in dispersion parameters for averaging periods less than 24 hours (Hanna, et al, 1982).

The inability of the models to describe vertical wind speed and direction shear may also result in significant error in prediction of plume trajectory. Trajectory error is generally greater under stable than in neutral or unstable conditions (Bass, 1981).

Upper air observations are critical for predicting plume transport, yet the twice daily radiosonde observations from stations typically 300-500 km apart effectively limit the temporal and spatial resolution of the windfield. Data extrapolated from surface station observations or predicted from numerical weather prediction models have been used to supplement upper air observations but the reliability of such data has not been fully tested. Accurate description of windfields is particularly troublesome in complex terrain and during "active weather conditions".

Information on the accuracy of long-range transport predictions is extremely limited especially for short-term (3-24 hour averages. The difficulty and expense of data collection has led to a dearth of adequate tracer field monitoring studies, necessary for model verification. The ongoing Cross Appalachian Pollution Transport Experiment (CAPTEX) is designed to provide such data. Numerical analysis of model sensitivity has consequently been more common, but according to Bass (1981), little of the work in this area has been reported in the open literature.

The results of some field verification studies for long- and short-term long-range transport models have been reported by Bass (1981) and are shown in Table 5. In general, the performance of long-range transport models is worse than that of Gaussian models. Predictions are particularly poor for short-term averages. For instance, experiments designed to predict three-hour concentrations at ground-level in the Northern Great Plains -- an area that has been relatively well characterized in plume transport studies -- found that predicted concentrations frequently differed from observed concentrations by more than a

Table 5. Empirical Estimates of the Accuracy of Long-Range Transport Models

A. Short Term Models:

Model/Conditions	Results	Source
Weekly averages of Krypton-85 transport of 1000 km (H version of model)	50% within factor of 2	Heffter (1977)
	90% within factor of 10	

B. Long Term Models:*

Model	Correlation Coefficient	
Western European LTRAP data base, EURMAP-1 model	SO^2 0.7-0.8	Mancuso (1979)
	SO^4 0.6-0.7	
NADB and EPRI/SURB data bases for eastern U.S. Lagrangian Puff model with more complex vertical diffusion	SO^2 0.7+	Meyers, et al. (1979)
	SO^4 0.6-0.8	

*Monthly--Averaged--Ground-level Concentrations

Source: Bass (1981).

factor of two (Bass, 1981). Again, these results are derived from studies of primary pollutants or inert tracers and it is anticipated that errors would be greater for reactive pollutants.

GROUND-WATER MODELS

Ground-water supplies drinking water for roughly half the United States population (Konikow, 1981). Increasing concerns about the contamination of this supply with toxic chemicals has led to greater pressure on ground-water modeling to provide information for the prevention management, or cleanup of contamination problems. The long time frames involved and the expense of exploratory drilling and testing for physical characterization of sites make model simulation a theoretically attractive management tool for both existing and proposed sites.

The need for ground-water modeling for routine application to field studies of contaminant migration has generally outstripped the development of the necessary models and data to run them (Anderson, 1979 and Konikow, 1981). Ground-water modeling, like long-range transport modeling for air pollutants, is still in relatively early stages of development. There is no equivalent among ground-water models of the Gaussian models that are widely used in air pollution dispersion modeling. Instead, there is greater emphasis placed on more complex numerical advection dispersion models for which there are still substantial theoretical and practical, (e.g., data requirements) limitations. Ground-water modeling shares with air pollution dispersion modeling limited capability to provide accurate results under complex

real-world conditions -- long distances, multiple dimensions, complex geological conditions, reactive contaminants, and multiple sources. This limitation is particularly severe in the use of modeling to predict movement and concentration of plumes.

The EPA has begun in the last few years to evaluate existing ground-water modeling capabilities and to develop guidelines for the appropriate application of models to risk assessment for field situations (JRB, 1982). As yet, though, no coherent framework for evaluating the use and performance of ground-water modeling seems to exist.

The following sections seek to develop an understanding of the uncertainties in ground-water modeling and their effect on uncertainties in risk assessments. The analysis begins with a presentation of the theoretical bases and data requirements for ground-water models and is followed by a discussion of the present application of models, the major sources of error and estimates of the magnitude of the error in model outputs.

An understanding of the basic structure and use of ground-water models is helpful for clarifying the sources of error in modeling contaminant transport. The following introduction is a brief overview of a complex subject; readers interested in more complete reviews should refer to valuable works by Anderson (1979), Faust and Mercer (1980), Mercer and Faust (1980 a,b), Pope-Reid (1982), Freeze and Cherry (1979) or Konikow (1981).

Two basic processes govern the transport of contaminants in ground-water;

1) advection: movement of a solute attributable to the bulk motion of flowing ground-water the rate of which is controlled by the average linear velocity of the water.

2) hydrodynamic dispersion: process in which the solute is spread out in directions other than that dominated by ground-water flow and which results from mechanical mixing in the soil and to a lesser extent from molecular diffusion.

Most models developed to study contaminant transport in ground-water consist of solutions to partial differential equations describing both advection and dispersion (Anderson, M.P., 1979). In some models, advection equations alone may be used to describe solute transport.

The advection or flow equation for flow in a non-homogeneous, anisotropic[1] medium has the following general form:

$$\frac{\delta}{\delta x_i}\left(T_{ij}\frac{h}{\delta x_j}\right) = S_s\frac{\delta h}{\delta t} + W \qquad i,j = 1,2,3 \qquad (10)$$

where:

T = transmissivity (aquifer thickness multiplied by hydraulic conductivity)((Length)^2/Time)

h = head (Length)

S = specific storage (Length^{-1})

W = volume flux per unit area

x = longitudinal distance (L)

y = transverse distance (L)

t = time

Hydraulic conductivity is a poorly understood property which describes the ability of the geological medium to conduct fluid. It is a function of both the porous medium and the fluid transmitted. The head is a measure of the potential energy in the ground-water system and is estimated by the height to which water will rise in a open stand-pipe. The change in head over a distance of hydraulic gradient $(\frac{\delta h}{\delta x_i})$ is the driving force for ground-water flow in the flow equation.

The dispersion equation, also known as the mass transport equation, estimates contamination concentration changes in time and space:

$$\frac{\delta C}{\delta t} = \frac{\delta}{\delta x_i}\left[D_{ij}\ \frac{\delta C}{\delta x_j}\right] - \frac{\delta}{\delta x_i}\ (Cv_i) + R \qquad i=1,2,3 \qquad (11)$$

where:

C = concentration of chemical in the ground water (M/L^3)

D = coefficient of hydrodynamic dispersion (L^2/T)

V = ground-water velocity (L/T)

R= rate of generation or removal of solute from ground-water (M/L^3T)

The coefficient of hydrodynamic dispersion describes the tendency of the medium to spread out a contaminant in directions other than those produced by the principle direction of ground-water flow and is a function of the fluid velocity and the dispersivity of the medium. Dispersivity is a property of the medium relating hydrodynamic dispersion to the velocity of ground-water flow and like hydraulic conductivity is physically poorly understood.

Ground-water velocity is often determined using Darcy's Law:

$$V_i = \frac{K}{\eta} \frac{\delta h}{\delta x_i} A \qquad (12)$$

where:

V = ground-water velocity (L/T)

K = hydraulic conductivity (L/T)

η = porosity (percent void space in total volume)

A = cross sectional area of the aquifer (L2)

In sum, the data required to study contaminant transport using the advection-dispension equations generally include (Pope-Reid, 1982):

Boundary conditions: information on the geometry of the ground-water system, initial head distributions, location and type of flow at the boundaries (impervious, constant flux, etc.).

Physical characteristics of the system: hydraulic conductivity, porosity, compressibility.

Flow variables: Darcy (average) velocity in any coordinate direction, coefficient of hydrodynamic dispersion (advection-dispersion models only).

The difficulty of obtaining the necessary data poses the greatest obstacle to the use of ground-water models for the study of contaminant transport (Anderson, M.P., 1979). No national data bases comparable to those available for air pollution dispersion models exist for ground-water contaminant transport

models. While there are published data available for some soil properties, (e.g., porosity and hydraulic conductives), data must in general be obtained from each site.

The flow and dispersion equations may be solved analytically or numerically. In analytical models, the equations are solved exactly, but usually after they have been simplified by assuming idealized conditions, e.g., steady state conditions for groundwater velocity and dispersion, and an aquifer of infinite extent.

Analytical solutions of the equations are also used to verify the results of the approximation techniques used to solve the equations in numerical models.

Numerical models employ any of several techniques for approximating the partial differential equations. The finite difference method, method of characteristics, and finite element are the most common. In the finite difference method the continuous function is approximated by a series of linear difference equations (Pope-Reid, 1982). The method of characteristics is similar but involves the additional step of expressing the partial differential equations as their "characteristic" set of ordinary differential equations which are then solved by finite difference methods. In the finite element method the partial differential equations are first transformed into integral form in order to be solved. A fourth method for solving the dispersion equation is Monte Carlo simulation (e.g., discrete parcel random walk) in which dispersion is treated as a random rather than deterministic process (Pope-Reid, 1982).

Function does not necessarily follow form in the case of ground-water contaminant transport models. Unlike air pollution dispersion modeling, in which the Gaussian model is particularly suited to certain distance ranges and Lagrangian models to others, it is difficult to judge which type or combination of ground-water models might be best suited for a specific application. The model chosen depends on the model application, the availability and quality of input data, the skill and experience of the modeler, and the nature and accuracy of the model solution desired. The apparent reason for this lack of generalization about ground-water models lies in the fact that the models have largely been developed on a case by case basis for specific applications. There have been few attempts to evaluate the validity of the models for general applications.

Anderson (1979) makes a broad distinction between the uses of advection dispersion models and advection models. Advection-dispersion models are most appropriate for small or local scale problems (a few meters) and should be used when detailed investigation of the spread of the plume is desired. Numerical techniques are usually necessary to solve the dispersion equation.

Advection models using analytical or numerical solutions may be used alone or in conjunction with water quality (surface water) models for larger or regional scale problems (100 m -- a few kilometers) to provide a first approximation of average change in water quality or solute travel times. They are most appropriate when the effects of dispersion may safely be ignored (over large distances and rapid ground-water flow).

Although this distinction is not rigid, there appears to be a consensus that advection-dispersion models are more appropriate than advection models for tracking toxic contaminant plumes (Anderson, 1979; Konikow, 1981: JRB, 1982; Mercer and Faust, 1980c). The preference for the more complex models appears to be based on their more "realistic" underpinnings. However, there has been relatively little evidence presented to show that the predictions of advection dispersion models are necessarily always better than those of the simpler models, particularly under complex field conditions.

Gorelick (1983) has recently reviewed several innovative approaches by various authors to the use of modeling for managing ground-water quality. In these instances, ground-water and/or surface-water models have been developed in conjunction with various optimization techniques (e.g., linear and quadratic programming). Thus, theoretically, allowing important policy and financial constraints to play a role in water resource and quality management. The basic problem addressed by these hybrid models consisted of managing the joint use of an aquifer for waste disposal and drinking water while maintaining acceptable water quality at supply wells. While these models are very promising tools for water quality management, they are still in relatively early stages of development. While this discussion illuminates some of the possible applications for ground-water quality management, it would be misleading to suggest that the ready application of modeling techniques will provide a practical and realistic solution for managing contamination problems. There are several serious constraints on the application of

ground-water modeling to the water quality management which can generate considerable uncertainty in model results. The constraints include the availability and quality of input data, in the estimation of model parameters in the ability of the model to mathematically represent complex, non-idealized conditions, in the basic stochastic nature of ground-water/solute transport, and finally in the accuracy of the numerical approximations techniques.

At the outset, many models which require numerical solutions are subject to errors that arise from the approximation techniques used to solve the advection dispersion equations. These numerical/mass balance errors are typically on the order of 10-15% and generally consist of two types: (1) numerical dispersion in which the contamination front predicted by the model is more smeared or dispersed than predicted by exact analytical solutions to the equations and (2) numerical oscillation in which the numerical solution overshoots and undershoots the values obtained analytically (Anderson, 1979; Pope-Reid, 1982). The finite difference method is particularly prone to these errors, while the finite element is less susceptible and the discrete parcel random walk method effectively eliminates them.

The data required to run the models present a far more serious problem.

> "We need, as a minimum, the permeability and porosity of the media and the hydraulic head gradients all in three dimensions. In addition, we need to know the sorptive characteristics of the media along all paths, and we need to estimate the variable rates at which the solidified wastes will enter the transporting fluids. Needed, in particular, is information on the distribution and extent of major heterogeneities. The need for such data severely taxes both

> the available data base and the technology for generating it. Most of the requisite available data have such large error limits that their usefulness in predictive models is limited." (Bredehoft, et al. in Anderson, 1979).

The velocity distribution and dispersivities necessary to solve the advection and dispersion equations are particularly difficult to obtain. According to Anderson (1979), "[t]o date, there are no well-tested, standard techniques for acquiring these data." Velocity may be measured directly using tracers but is more commonly determined indirectly using measurements of heads, hydraulic conductivity and Darcy's Law. The hydraulic conductivities used are either spatially averaged values obtained from analysis of field tests or are fitted parameters determined from the trial and error adjustment (calibration) of hydraulic conductivities until a flow model simulates the head distribution observed in the field.

The uncertainties in the velocity distribution and consequently in the rate of contaminant transport derive directly from the problems in describing the spatial variation in hydraulic conductivity. Inhomogeneities in the porous medium, such as "stringers" of more permeable materials or pockets of less permeable material, play a critical role in contaminant transport and dispersion. The difficulties in accurately representing these inhomogeneities can therefore effectively limit the predictive capabilities of the models.

Porosity, which is also required to compute ground-water velocities using Darcy's Law, further contributes to uncertainty in contaminant transport modeling. It can vary by several orders of magnitude over small distances, thus reducing the validity of average values measured in the lab or field.

Measurement of dispersivity, which is used to estimate the magnitude of contaminant transport attributable to hydrodynamic dispersion, is also very difficult. The primary problem is that the magnitude of the dispersivity measured depends highly on the scale on which the measurements are made (Anderson, 1979); Pope-Reid, 1982). Dispersivity measured in small scale laboratory and in larger scale field tests can yield values for the same medium that differ by several orders of magnitude. For instance, values obtained from laboratory tests typically range from 10^{-2} to 1.0 centimeters whereas field tests yield values in the range of 10-100 meters (Anderson, M.P., 1979). A major reason for the disparity in the results appears to be that laboratory tests are unable to account for the effects of large-scale inhomogeneities in the aquifer. Dispersivities may also be determined in the process of calibrating a model but again, as fitted parameters, they may not accurately describe the actual geological conditions.

There are also practical limits to the amount of data that can be collected to characterize a site. Not only is drilling expensive, but the number of boreholes that would theoretically

be necessary to characterize the spatial inhomogeneities that are so often crucial for ground-water flow and contaminant transport could change the geological properties of aquifer (Fiering, Personal communication, 1983).

A successful application and calibration of model for a given site depends heavily on the expertise and experience of the modeler at estimating the missing parameters of the model -- the gaps left by field measurements. The importance of the modeler to "intelligent" model use makes what Anderson (1983) calls the "institutionalized black boxing" of models so dangerous. The temptation for anyone to run complex models without the necessary training or understanding of the limitations of the results is great and indeed encouraged by current regulatory interest in the use of ground-water models.

As the discussion of data requirements shows ground-water flow contaminant transport models are by necessity simplified representations of complex hydrogeological conditions and processes. Faithful mathematical simulation of all processes affecting contaminant transport would not necessarily assure better results. However, additional assumptions that are commonly made to simplify more complex existing conditions that either poorly characterized or poorly understood contribute to the errors in and unreliability of model outputs. However, because of burdensome data requirements, one or two dimensional models are commonly used to represent three dimensional systems. When three dimensional transport is important, such models are likely, on average, to over-estimate concentrations at any given node. Another major source of error in the theoretical basis for

contaminant transport models lies in their inability to account accurately for the complex reactions that a chemical undergoes while in the soil. As in air pollution-dispersion models a single reaction term may be incorporated to represent chemical reactions, precipitation, absorption, desorption, ion exchange, volatization, etc. The incorporation of these processes into contaminant transport models has been effectively stymied by the paucity of experimental laboratory or field studies on the nature and rates of these reactions (Anderson, 1979).

The common assumption that the chemical species of interest in a modeling situation is non-reactive (i.e. is not broken down or adsorbed to the soil) is likely to be very conservative. Anderson (1979) reported the results of a sensitivity analysis of an advection model developed at Oregon State University and used to predict the quantity and quality of leachate produced by a sanitary landfill. The analysis demonstrated the importance of biodegradation and adsorption in determining the contamination detected.

> "Moderate degradation in the landfill removed essentially all of the contaminant, thereby producing virtually contaminant-free leachate. Moderate degradation occurring only in the soil below the landfill almost eliminated contaminants from the ground-water discharge, while weak degradation of the contaminant in the soil eliminated 86% of the contaminant ...In contrast, the results were relatively insensitive to changes in ground-water velocity and water-table fluctuations".

Finally, ground-water models are in general deterministic while the hydrogeologic and chemical processes they are developed to represent are generally stochastic.

How good, then is the output of current ground-water models for use in risk assessment? In contrast to Gaussian dispersion models, estimates of the magnitude of the errors associated with contaminant transport models are not readily available. Data can be found on numerical errors and on the differences between concentrations observed in the field and simulated during model calibration but there is a dearth of information on the magnitude of the errors that might be expected when models are used under "real world" conditions (few data, complex geology, etc.) to predict contaminant transport over extended time periods or distances. This paucity of data reflects the relatively recent arrival of ground-water modeling to the regulatory arena; and the difficulties in running and testing the application of the models has not been as routine or widespread as it has been for air pollution dispersion models. The systematic evaluation of ground-water models necessary to provide perspective on the appropriate role of the models in risk assessment and in the management of hazardous waste has not been done.

The most common error estimates that appear in the literature describe the ability of the models to reproduce existing conditions at a site (e.g., the existing contamination phases). Under the "best" conditions (for instance, during the research and development of a model, with adequate data) models can be expected to reproduce existing conditions relatively well. Table 6 presents the error estimates for several analytical and numerical models that have been field tested (Pope-Reid, 1982). The errors range from about 5 to 30 percent but unfortunately, neither the source of the errors nor the outputs to which they

Outputs	Analytic Verification	Application to Field Studies	Error Estimates	Advantages	Disadvantages	Model	Equation	Type	Method	Dimension	Input
Location of	N	N	10%	Inexpensive	Cumbersome graphical techniques for an isotropic media	Wilson	Mass transport of nonconservative chemical instead, saturated flow	Analytic	Graphical	2	F,M,C (mass load of point source relative concentration), decay constant
Concentration	Y	Y	5%	Low numerical dispersion, multicomponent system, solution stacking, designed for small computers	Saturated flow only, constant aquifer thickness, long computation time, requires groundwater velocities as input	Ahlstrom	Vertically averaged mass transport	Computer	Random walk	1 or 2	*B,F,C, reaction expression or equilibrium constants, decay constants
Relative Concentration moisture content	Y	Y	5%	Incorporates evaporative loss from surface, reactive chemical species		Davidson	Transport of water and adsorbed solutes to transient flow	Computer	Finite difference	1	B,M,F,C adsorption and desorption coefficients
Relative Concentration	Y	Y	30% error of concentration contour		Oscillating solutions, numerical dispersions, ignores non-diagonal terms in dispersions, ignores non-diagonal terms in dispersion tension	Goreghian	Mass transport in saturated homogenous media	Computer	Finite element	2	M,F,C (initial peak concentration and standard deviation)
Concentration, heads, drawdown ground-water velocities, mass balance	Y	Y	10%	Variable thickness, anisotropic pumping schedules, observation point printouts	Some oscillation, nonreactive solute, mass balance error, no nondiagonal dispersion tensor terms	Konikow	Mass transport in transient saturated flow	Computer	Finite difference method of characteristics	2	B,M,C storage coefficient
Pressure head, total head, moisture content	N	Y	9%	Continuous velocity field, non-diagonal K tensor terms, capillary forces, unsaturated flow, seepage face boundary conditions	Requires detailed soil parameter inputs	Yeh	Transient ground-water flow in saturated/un-saturated media	Computer	Finite element	2	B,M coefficient of compressibility of medium and water, h-9-k function
Relative concentration material flux	N	Y	10%	No numerical oscillation, non-diagonal dispersion tensor terms, sorption and decay, capillary forces, unsaturated flow, seepage faces, 12 optional numerical schemes	Detailed soil parameter inputs and requires separate program for flow variables	Yeh	Mass transport in transient saturated unsaturated flow	Computer	Finite element	2	F,C distribution coefficient, coefficient of compressibility

Notes:

*B: Boundary conditions, including initial head distribution, locations and types of boundaries.

M: Material constants, including hydraulic conductivity, porosity, compressibility of the medium, thickness of geohydrologic units.

F: Flow variables, including ground-water velocity and coefficient of hydrodynamic dispersion.

C: Concentration parameters, including initial concentration and boundary conditions.

Table 6. Error Estimates for Groundwater Models

apply are always indicated. They are believed to reflect both numerical/mass-balance errors as well as differences between simulated and observed values.[2]

In most contaminant transport models, mass balance errors alone are on the order of 10-15% (Anderson, 1979; Pope-Reid, 1982). Konikow and Bredehoft (1974) used a solute transport model to study the effects of irrigation on the distribution of dissolved solids in the ground-water of an alluvial aquifer in Colorado. The finite difference method was used for the dispersion equation. When the model was calibrated, it reproduced the dissolved solids concentration within 10% of the observed values about 80% of the time (Anderson, 1979).

The more interesting and more perplexing question for risk assessment concerns the accuracy of model predictions over time and space. Unfortunately, the error estimates given above for models that have been calibrated for specific sites provide limited insight into the ability of the model to predict the movement and concentration of ground-water contaminants. One would expect the agreement between observed and simulated values to be reasonably good when the model is calibrated. During calibration, the parameters of a model are adjusted until the model simulates observed field conditions as well as possible. The extension of the model to areas where contaminant plumes have not been found and studied is a very different problem. Here, all the uncertainties created by the difficulty of characterizing flow and contaminant transport chemical interaction, deposition and decay in complex three-dimensional systems come into play.

Even the calibration process may lead future predictions astray since the physical properties of the aquifer can become fitted parameters whose relationship to actual field conditions is uncertain; some authors have expressed the concern that different fitted values of hydraulic conductivity, for instance, can yield the same head distribution.

An additional problem is that, strictly speaking, predictions should be made for periods of time only as long as the period of observation for the aquifer or site. In practice, however, this advice is often not followed since historical data on ground-water flow or contaminant levels is available for relatively short periods of time and even then for few sites. For proposed sites these data are especially rare. Since ground-water contamination problems may span decades, there is a strong incentive to apply models to periods exceeding those for which there is adequate data.

To date, few authors have been willing to estimate the likely accuracy of ground-water model predictions for "real-world" applications. One modeler pessimistically estimated that under typical field conditions with few data, model predictions could be off by as much as two or three orders of magnitude.[3] In a recent editorial, Anderson (1983) argued restraint in the use of ground-water models:

> It is clear that models must be used in conjunction with field studies and in fact, field studies to help resolve the questions about dispersion and chemical reactions in the subsurface are in progress and in planning. Until the results of these studies are in, the promotion of ground-water models for contaminant transport applications should be viewed with extreme caution.

The time when ground-water modeling can be used routinely for ground-water contamination problems, even by qualified modelers, is still a good way off.

Concentration, Exposure and Dose

Environmental fate and transport models yield spatial and temporal fields of ambient concentration estimates. But hazard assessment typically requires as an input some estimate of either exposure or dose. Exposure analysis involves consideration of human activity patterns and dietary habits in conjunction with the estimation of ambient concentrations. Dose estimation goes one step further -- taking into account human metabolism.

Dose is a complex function of absorption, metabolism and excretion rates, which themselves are influenced by age, sex, stature and activity level. Although the details of absorption, metabolism and excretion are known for certain chemicals, for most, it is impossible to carry out a full analysis. Therefore, for risk assessment, a simplified estimate based solely on the concentration in air or water and volumetric rate of intake (liters of air breathed/day, liters water/day) is commonly made. All of the chemical inhaled or ingested is assumed to be absorbed. For carcinogenic risk assessment, dose rate is usually expressed in the form compatible with EPA carcinogenic potency estimates derived from animal bioassays involving uniform life-time exposure -- mg/kg bodyweight/day or mg/m^2 surface area/day.

The EPA has adopted the International Commission on Radiological Protection (IRCP) "reference man" as a basis for converting exposure to dose rate (Federal Register, November 28, 1983: Norman, Charles; EPA Exposure Assessment Group, July 1983). The "reference man", a composite of individuals in various countries and time periods, basically represents a Caucasian male, 20-30 years of age, of Western European or North American origins (ICRP), 1981). Table 7 presents the physical characteristics and estimates of daily air and fluid intake for the reference man.

The dose rate estimate, d (mg/kg per day), for the "reference man" would then be:

$$d = 0.33\ C_a + 0.028\ C_w \qquad (13)$$

where $C_a (mg/m^3)$ is the concentration of the contaminant in air and $C_w (mg/L)$ is its concentration in water. This dose rate estimate is obviously appropriate for estimating the risk faced by a hypothetical "reference man" exposed to identical airborne concentrations in all microenvironments and deriving his entire fluid intake from equally contaminated sources. Several conceptual generalizations would seem necessary to make this dose rate estimate useful for risk assessment.

First, it is likely that the risk faced by a biologically-average person is of more interest than the risk faced by the ICRP "reference man". Therefore, in principle, the metabolic coefficients 0.33 and 0.028 should be adjusted to reflect a weighted average of the values appropriate for various demographic (age - sex - and racial) groups. In practice, such an adjustment is not commonly made and is not believed to be impor-

Table 7. Physiological Parameters for ICRP Standard Reference Man

Height	170 cm
Body Weight	70 kg
Total Fluid Intake	1.95 L/day
Tap Water	0.15 L.day
Volume of Air Breathed/day	
8 hr. working "light" activity	9.6 m^3
8 hr. non-occupational activity	9.6 m^3
8 hr. resting	3.6 m^3
Total	23 m^3/day
% total air breathed at work	42

Source: International Commission on Radiological Protection, 1981. Report of the Task Group on Reference Man. Publication No. 23. Pergamon Press, New York, New York, 480 pp.

tant since any bias introduced by use of doses appropriate for a "reference man" is thought to be very small in comparison with the uncertainty typical in estimates of risk.

Concerns that the 70 kg male may not adequately represent more susceptible or sensitive individuals in the population has led to the use of the 10 kg child in some risk assessments. Children have a higher air and fluid intake on a body weight basis than adults (Severn, 1983). However, this is relevant only if one is interested in the distribution of risk among the population, or if potencies are strongly dependent upon age at exposure, as is believed to be the case for at least one carcinogen, ionizing radiation.[7] For most carcinogens, data on potency as a function of age at exposure are unavailable.

To estimate the distribution of risks for various age-groups in the population, it is also necessary to rely on a model of the dynamics of risk. Two such models have been advanced in the area of radiation carcinogenesis. The absolute risk model assumes that after a latency period, 1 (yr), thought to be about 2 years for leukemia or bone cancer and 10 years for most solid tumors, incremental annual risks are constant throughout a plateau, or expression-period, p (yr). The plateau period is now thought to be 25 years for leukemia and bone cancer and the remainder of lifetime for most solid tumors. The relative risk model assumes that after the latency period incremental annual risks increase roughly in proportion to baseline cancer risks.

These models provide a basis for projecting cancer risks in future time periods and also for estimating the contributions of

doses received in past time periods to current cancer risks. They give an indication of the appropriate averaging-time for cancer dosimetry. If the models are correct, and applicable to other carcinogens, then the appropriate averaging times are roughly 25 years for leukemia and bone cancers and about 55 years for most solid tumors.[8]

Similarly, the assumption that 100% of the contaminant inhaled or ingested is absorbed and reaches the target tissues would seem quite crude. Absorbtion, metabolism, and excretion strongly influence the dose actually reaching the target tissue. Although there is a compelling need for the incorporation of these rates into quantitative risk assessment, the necessary supporting data are lacking for most chemicals. Since 100% absorption does not occur for all chemicals and because metabolism and excretion of a substance is likely to decrease the doses actually reaching the target tissue, the use of 100% absorption assumption usually overestimates the dose.

It might seem that this would lead to overestimation of risk. But, this is not always the case. What is important is the correspondence between the dosimetry used in the derivation of potency estimates and the dosimetry used in risk assessment. If potency has been estimated from analysis of past human exposures, no bias will result if similar measures of dose or exposure were used in both components of the analysis. If, on the other hand, potency has been estimated from analysis of animal data, the accuracy of the animal analogy becomes important. As

long as the relationships between amount of contaminant inhaled or ingested and dose to target tissues are similar for man and the test species, no bias will result.

Recently, a great deal of attention has been devoted to indoor air pollution. It has been estimated that the average person in the U.S. spends more than 75% of the time indoors. Again, it would seem that dose estimates based upon outdoor concentrations of toxic pollutants might yield biased risk estimates. However, the existence and extent of bias depends entirely upon the relationship between outdoor and indoor concentrations. For many pollutants the only difference in indoor and outdoor concentrations is due to the capacitance effect of structures. For these pollutants, although pulses may be damped and there may be lags in the time patterns of exposure, long-term indoor averages are virtually identical to long-term outdoor averages. Although under non-linear does response curves with short biological averaging times this could lead to bias, in most circumstances it would not. For other pollutants, deposition on surfaces and/or reaction with structural materials and furnishings may lead to depletion of contaminants. For these, indoor concentrations might be substantially less than those outdoors causing risk estimates based upon outdoor exposures to be biased upwards.

III. HAZARD ASSESSMENT

Potency Estimation

The final step in risk assessment is to apply estimates of, the potency of a given substance in humans. The methods for estimating which differ depending on the source of data and on the form of toxicity, are discussed briefly below.

Epidemiologic studies and animal bioassays are the two primary sources of potency estimates. Because they eliminate the need to extrapolate results from animals to humans, epidemiologic studies are preferable to animal studies. However, there are relatively few compounds for which valid epidemiologic data exist. Less than thirty of the 70,000 chemicals in commercial use in the U.S. have been definitely associated with cancer in humans (Tomatis, 1978 in NRC, 1983). The human data for other toxic effects, such as teratogenicity and neurotoxicity are even more limited.

When epidemiologic data are available they must be used cautiously. Many epidemiologic studies require retrospective analysis of the health effects of occupational exposures to toxic compounds. In these studies the comparability between workers and the general population is a source of uncertainty.

In occupational epidemiology, a critical issue is selection of an appropriate control group. In addition, the comparability of exposures of workers and members of the public is at issue. A working population is generally exposed to much higher concentra-

tions than the general population albeit for 8 hours a day and for less than lifetime. Retrospective epidemiologic studies typically have little data on the nature and level of historical exposure of the study populations and must therefore develop associations on the basis of current exposure measurements. For diseases with long latency periods, like cancer, current exposures may be poor surrogates for the relevant measures of dose, especially if exposures have changed appreciably over the years. In general, the use of imprecise measures of exposure tends to bias risk estimates towards zero. Prospective studies (studies which measure current exposures and monitor incidence of disease forward in time) eliminate this problem but are expensive, time consuming and as a result are far less common.

Inherent limitations of epidemiologic studies further restrict their usefulness in risk assessment. Difficulties in obtaining and following up a large enough study population to be able to detect an effect, in defining exposed and unexposed populations, in describing the nature and levels of exposures over the study period and in controlling for exposures to confounding factors -- factors which are associated with both the exposure and the disease (e.g. smoking) generate considerable uncertainty in the existence and the strength of an effect detected (Weinstein, 1979). Rarely can the magnitude of that uncertainty be estimated. In addition, the long latency periods between exposures and the appearance of statistically detectable

effects and the time and expense involved in large epidemiologic studies make it unpractical for most decisions requiring risk assessment to await the outcome of epidemiologic studies (Weinstein, 1979).

Toxicity testing in animals is far more common than epidemiologic studies. Approximately 7000 substances have been investigated in animal bioassays, of these, 1500 are reported to be carcinogenic (Maugh, 1978 in NRC 1983). However, there are several sources of uncertainty in the use of animal data as a basis for human potency estimates. First, the doses administered to the animals are typically much higher than those encountered in the environment. Therefore, models must be relied upon to provide estimates of animal potency at low dose rates. Second, the effects in genetically homogeneous populations must then be extrapolated to a heterogeneous human population. Furthermore, although the National Cancer Institute has developed standardized designs for carcinogenesis bioassays, problems in design and execution of the studies can greatly affect confidence in the observed results.

Two critical steps in the development of potency estimates are low dose extrapolation and, for animal data, interspecies comparison or scale up of results.

Low Dose Extrapolation

Both epidemiological and animal studies typically involve health effects of exposure to concentrations that are typically a few orders of magnitude greater than those encountered in the general environment. In order to estimate the risks of long-term

exposures to low concentrations, it has been necessary to develop methods for extrapolating from the existing experimental data.

Several mathematical models have been developed for this purpose. All of the models widely used in regulatory applications are non-threshold models -- models for which there is no dose below which the risk is assumed to be zero (Anderson, 1983). The choice of this class of models has been based on prevailing theories of carcinogenesis and on current practical obstacles to identifying thresholds for carcinogens. A widely held belief is that most forms of carcinogenesis involve interaction with, and irreversible damage to DNA, a process for which there is theoretically no threshold dose. Even if thresholds exist, it is not currently feasible to design experiments capable of detecting them.

One of the earliest procedures for low dose extrapolation was developed by Mantel-Bryan in 1961 (Hogan and Hoel, 1982). The Mantel-Bryan method is based on the assumption that the relationship between the logarithm of dose and the probability of response is approximately described by the cumulative normal distribution. The Mantel-Bryan estimate of an upper bound on risk at low doses is found by extrapolating along a line of slope one[9] from an upper confidence limit (99%) on the proportion of animals observed with tumors at a given exposure level to the dose level of interest. The Mantel-Bryan procedure is no longer commonly used because it often does not fit the data well in the experimental dose range, is not well supported by any biological theory of carcinogenesis, and (although it is inherently

conservative) often gives less conservative results than more recently developed models. (Hogan and Hoel, 1982).

More common are the so called "hit" models which assume that carcinogenesis involves a finite number of interactions or "hits" of the substance with the target tissue before an identifiable tumor develops. The most basic of the hit models and one which has been used in risk assessment by EPA is the one-hit model. The model, which assumes that only one dose related stage is necessary to induce cancer, has the following mathematical form:

$$P(d) = 1 - \exp(-\beta d) \qquad (14)$$

where P(d) is the projected risk at dose level, d,; β is the unknown model parameter and d the expected number of hits. In the low dose region this becomes $P(d) \approx \beta d$, a simple linear model. The one-hit model is not considered to be as flexible as the other linear non-threshold models to be discussed. Because it only has one parameter, β, it often is not able to fit the experimental data well; particularly if the data have strong upward curvature. (Crump and Howe, 1980).

The gamma multi-hit model, developed by Cornfield and Van Ryzin, assumes that the initiation of cancer requires a series of k hits and incorporates spontaneous cancer incidence using what is known as Abbot's Correction Factor (Hogan and Hoel,

1982). The most general form of the model, in which k can assume non-integer values, is:

$$P(d) = \int_0^d \beta^k t^{k-1} \exp(-\beta t)\, dt/\Gamma(k) \qquad (15)$$

where β and k are model parameters and $\Gamma(k)$ is the gamma function. Cooper (1983) has derived a form of the model appropriate for integral numbers of hits:

$$P(d) = 1 - \exp(\beta d)\left[1 + d + \frac{1}{2!}(\beta d)^2 + \ldots + \frac{1}{k!}(\beta d)^k\right] \qquad (16)$$

where β and k are model parameters. The parameters in this version of the model are readily interpretable; βd is the expected number of hits from a dose d, and k is the number of hits required to initiate a tumor. While the gamma multi-hit model is more flexible for fitting experimental data, persistent doubts about the model have precluded its widespread acceptance (Hogan and Hoel, 1982).

The Weibull model has the following form:

$$P(d) = 1 - \exp(-\beta d^m) \qquad (17)$$

where β and m are model parameters (FSC, 1980). The model is linear when m=1, concave when $m<1$ and convex when $m>1$. To date the Weibull has not found widespread use in risk assessment for environmental cancer.

The last of the models to be discussed is the one that is primarily used by the EPA for risk assessment (Anderson 1983).

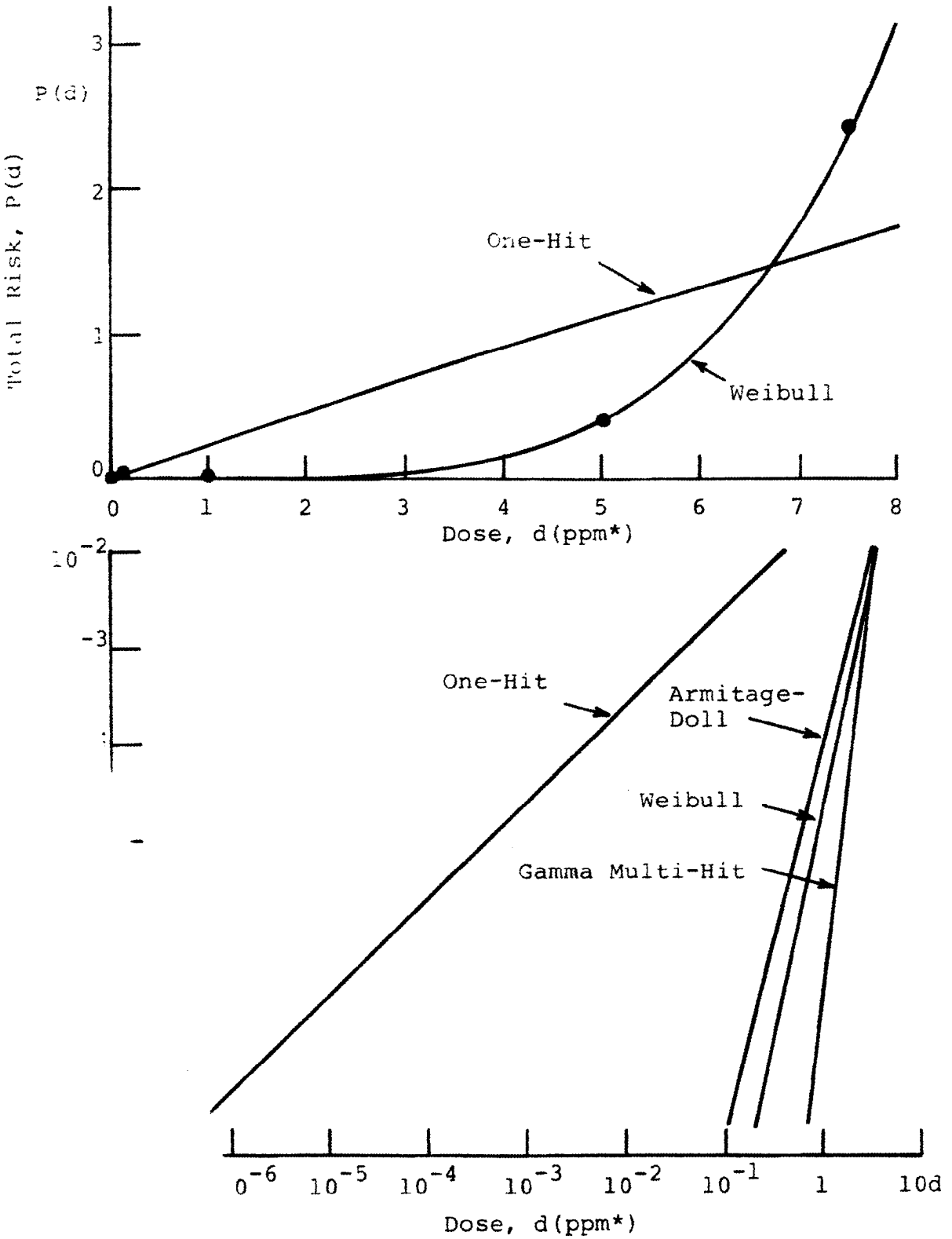
The multistage model, developed by Armitage and Doll (1961) and later modified by Crump (1980), assumes that carcinogenesis occurs as a series of stages or events. Each event or stage is assumed to be independent and additive and their rate of occurrence is linearly related to dose:

$$P(d) = 1 - e^{\{-(\beta_1 d + \beta_2 d + \cdots + \beta_k d^k)\}} \quad (18)$$

where k is the unknown number of stages. At low doses the model can be approximated by $P(d), = \beta_1 d,$ a simple linear model. The upper 95% confidence limit on β_1 commonly is used to develop an upper bound estimate of potency. Use of the multistage model has been defended on the basis of the model's biological plausibility and flexibility in fitting data in the experimental dose ranges. Critics, however, have questioned the relevance of the model for carcinogens for which interaction with DNA does not appear to be a critical step -- a problem with all of the "hit" models (Hogan and Hoel, 1982).

Given current understanding of process carcinogenesis, there is no scientific basis for determining which model is most appropriate for low dose extrapolation. The models in most cases fit the data in the experimental dose ranges equally well. The problem is that at low doses the estimate of risk is strongly dependent upon the choice of model.

Figures 1 and 2 illustrate the wide variation that exists among risk estimates given by various models at low doses. At doses in the experimental range the models give similar estimates of risks. However, the variability among estimates of risk



Low Dose Risk Estimates for Saccharin Under Various Dose-Response Models

lor et al., Toxic. Applied Pharmicology, 29, 154, ract 200, 1974

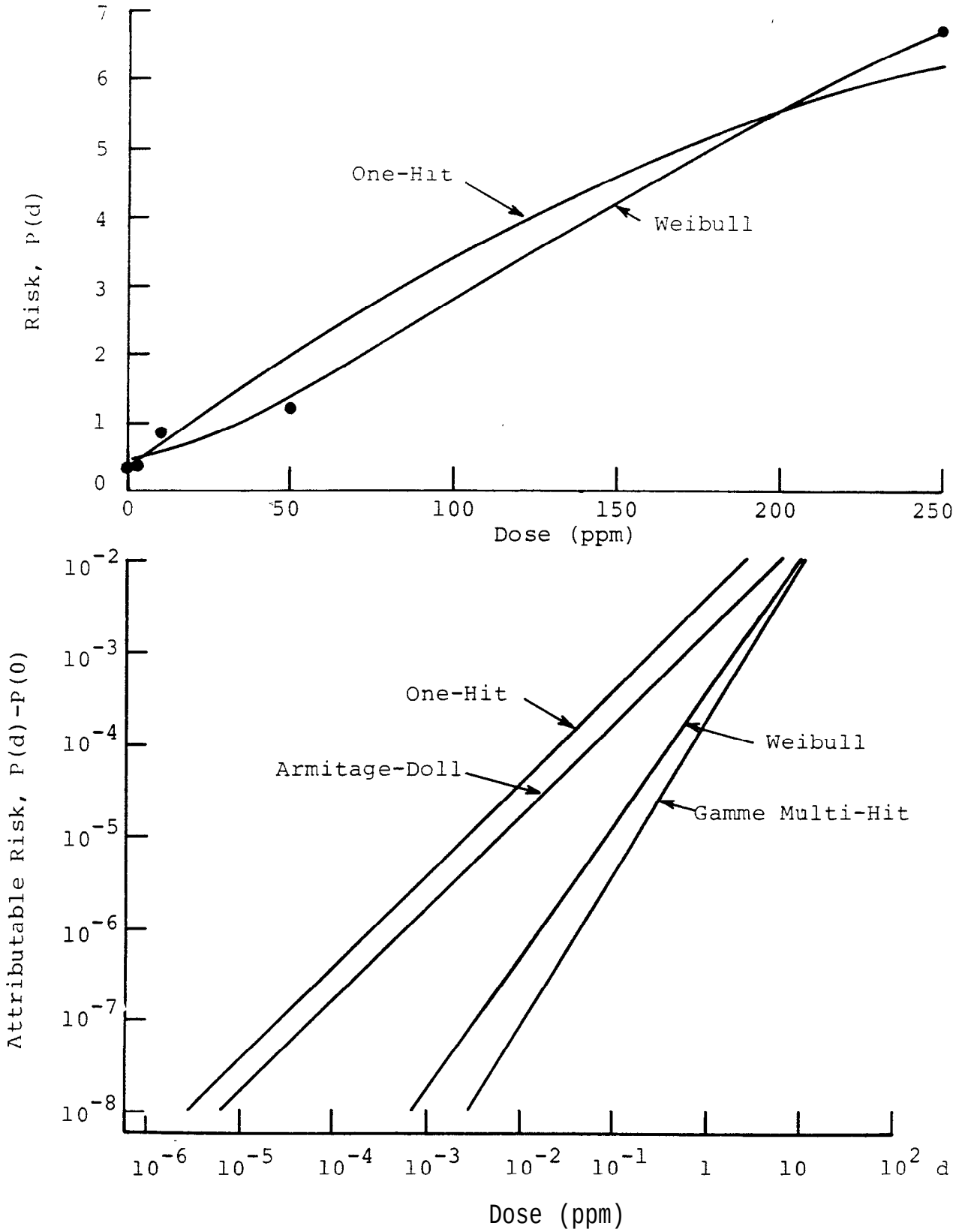


Figure 2. Lose Dose Risk Estimates for DDT Under Various Dose-Response Models

Source: Tomatis et al., Int. J. Cancer, 10, 489, 1972

increases dramatically with decreasing dose. At risk levels of significance for regulatory purposes (10^{-3} - 10^{-6} lifetime risks) the associated dose levels may vary by several orders of magnitude.

One point concerning low-dose extrapolation must be appreciated. The attributable risk due to exposure is well known at two points: zero dose and in the range of experimental doses. Intuitively it is obvious that in an absolute sense the uncertainty surrounding an estimate of risk becomes quite small as the dose approaches zero. The absolute uncertainty about risk is also small in the center of the range of experimental doses. Quite the opposite is true if ratios of potency estimates are compared. As the dose approaches zero, an approximate upper bound on potency is given by β_1, the slope of the one hit model. However the lower bound on low dose potency near zero dose is zero. And the ratio of these potency estimates is infinite.

Choice of model is obviously a crucial issue in cancer risk assessment. Since there is no clear scientific basis for model selection some have suggested that central estimates of risk should be based upon a weighted average of the risks given by several plausible models:

$$P(d) = w_1 P_1(d) + w_2 P_2(d) + \ldots + w_n P_n(d) \qquad (19)$$

where $P_1(d)$ is the estimate of risk given by model 1, w_1 is the probability that model 1 is correct, and so forth. The weights of course would be subjective probability estimates. Harrison (1983) has noted that because the one-hit model typically gives risk estimates orders of magnitude above those given by the

multi-stage, Weibull and multi-hit models, the central risk estimate would reduce to:

$$P(d) \approx w_1 P_1(d) \qquad (20)$$

The practical difficulty in this approach lies in determination of the set of plausible models and the vector of subjective weights, $w_1, w_2, \ldots, w_n$. Decision analytic approaches would be required to estimate the weights (see, for example, Morgan, Henrion, and Morris (1981). The results would be sensitive to the composition of the group of experts chosen. And, therefore, the results might be difficult to defend as a basis for public policy. However if a central estimate of risk is required, there is no better alternative.

Recognizing the difficulty in determining the weights, $w_1, w_2, \ldots, w_n$, the EPA has taken the strategy of giving approximate bounds on potency, rather than a central estimate of potency and an estimate of the uncertainty in the potency estimate. The approximate bounds are zero and a 95% confidence interval estimate of $_1$ from (6), the multi-stage model.

Interspecies Extrapolation

One of the most uncertain steps in risk assessment involves inference of human potency, β_h, from estimates of potency derived from animal data. As we noted in an earlier paper (Evans et al., 1982):

> "The most common assumption to make, but one that is nonethe less difficult to use, is to assume that one species is like another. In Figures 3 and 4 we illustrate the problems with this assumption. Figure 3 is a rat; Figure 4 is a man,

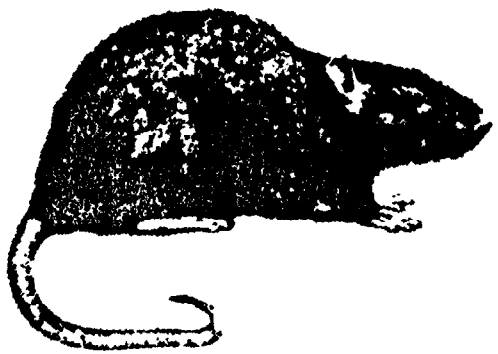


Figure 3. Rat

Figure 4. Homo Sapiens

> they do not look alike and we must be tolerant of the often expressed doubts by intelligent members of the public and explain ourselves well. In using a rat to tell us about men, we do not imply that a rat is an embryonic man (about to turn into a prince by a magic wand) nor that a man is an overgrown rat, but to use the known fact that some of the metabolic processes and cell structure are the same. Ultimately, however, the use of such an analogy must rest on data on other chemicals, accompanied by careful interpretation..."

The species most commonly used in cancer bioassays are rats and mice. As Table 8 indicates, there are many differences between rats, mice and men. Most estimates of human potency from animal potency have been based upon the assumption of <u>equal sensitivity</u> of both species. And the issue has been cast in terms of selecting the appropriate measure of dose. As Table 9 indicates, the estimates of human potency derived in this way are strongly dependent upon the measure of equivalent dose which is used to make the comparison. For example, estimates of human potency derived from studies of mice would be approximately 40 times lower if based on mg/(kg-day) than if based upon mg/kg. Similarly, human potency estimates from mice based on $mg/(m^2-day)$ would be about 3.5 times lower than those based upon mg/kg and 11.5 times higher than those based upon mg/(kg-day). Concerning this dilemma, Hogan and Hoe1 (1982) note:

> "In order to realistically choose among these competing dosage scales, it is necessary to compare actual human cancer risks derived from epidemiologic studies with the various animal-based estimates that would be produced with the different dosage scales under consideration. Unfortunately, very little data are available for making these types of comparisons."

Several attempts have been made to compare animal-based estimates of potency with those from epidemiology. A 1975 National Academy of Science (NAS) study of the health hazards of

Table 8. Body Weights, Life Expectancies and Ranges of Intake of Air, Water and Food for Rats, Mice and Men

	Mouse	Rat	Man
Weight (kg)	0.025	0.25	70
Lifelength (yr)	1.75	2.0	70
Air Intake (m3/day)	0.04	0.2	15
Water Consumption (L/day)	0.005	0.015	2.5
Food Consumption (kg/day)	0.005	0.015	1.5
Surface Area (m2)	0.0075	0.0357	1.8

Source: Crouch and Wilson (1979)
Friedrich et al. (1966)

Table 9. Ratios of Lifetime Dose (mg) Required to Produce a Unit Dose in Rat or Man to Lifetime Dose Required to Produce a Unit Dose in a Mouse

Measure of Dose or Dose Rate	Man	Rat	Mouse
mg/kg	2.8 x 103	10.0	1
mg/ (kg-day)	1.1 x 105	11.0	1
mg/ (m2-day)	9.6 x 103	5.4	1

pesticides compared animal-based and human estimates of potency for benzidine, chlornaphazine cigarette smoke, aflatoxin B_1, DES and vinyl chloride. Using the mg/kg measure of dose equivalence, the animal and human potency estimates were within a factor of ten of each other, with the exception of those for DES and vinyl chloride. The epidemiologic estimates of risks for vinyl chloride were a factor of 500 lower than those predicted on the basis of animal data. A reanalysis of these data by Hoel indicated that when either the mg/(kg-day) or $mg/(m^2\text{-day})$ measures of dose were used animal-based and epidemiologic estimates of human potency tended to agree within a factor of ten (Hogan and Hoel, 1982).

Figures 5 through 8 from Crouch and Wilson (1979) illustrate the results of an analysis which extends the NAS data set to include nine additional chemicals (acrylonitride, arsenic, benzene, chloroform, 3-3'-dichlorobenzidine, ethylenedibromide, lead acetate, saccharin and radiation. The figures, which are based upon mg/(kg-day), illustrate clearly the approximate nature of current models for interspecies extrapolation. Note that the best-fit lines of unit slope (dashed lines on Figures 5 and 6) indicate constant relative potencies between Osborne-Mendel and Fischer rats and those in B6C3F1 mice of 0.40 and 4.5 respectively. In general, the estimates of potency in rats based upon potency in mice appear to be within a factor of ten of the measured potencies in rats. The correlation between potency in

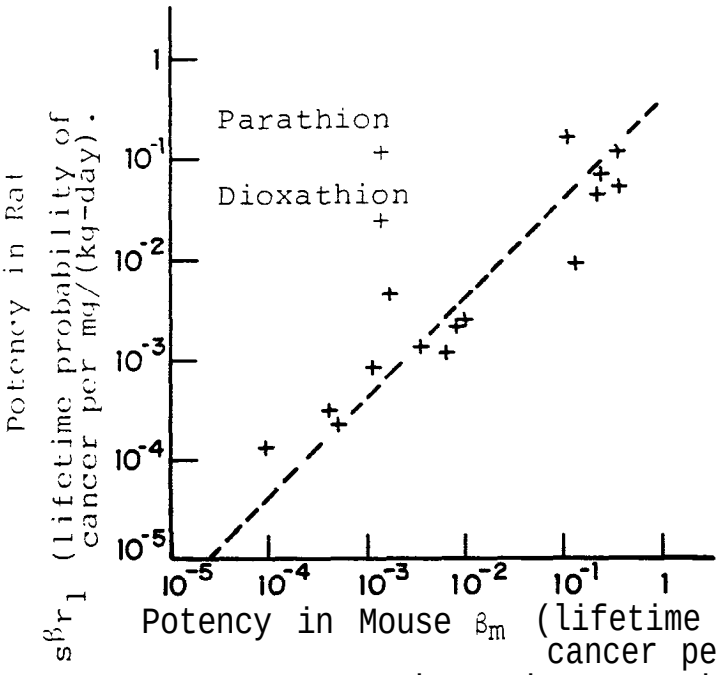


Figure 5. Carcinogenic Potencies in Osborne-Mendel Rat Versus B6C3F1 Mouse.

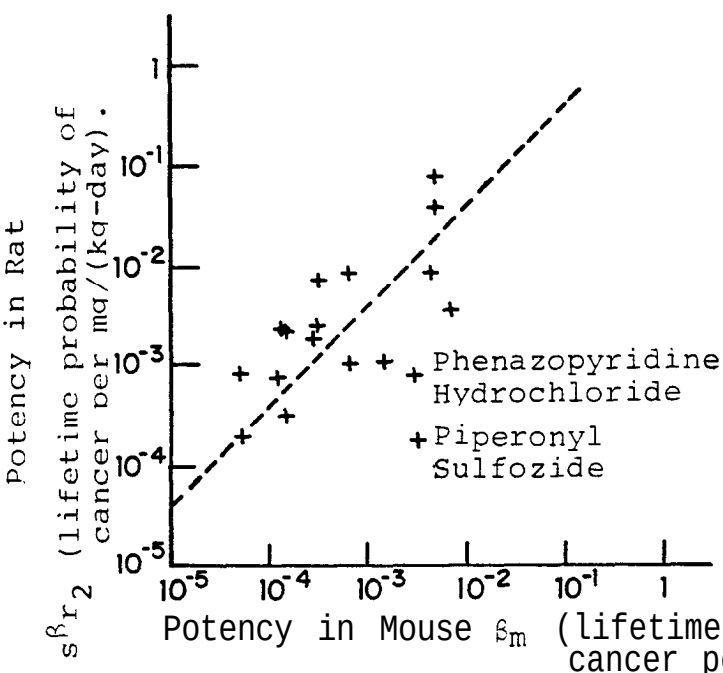


Figure 6. Carcinogenic Potencies in Fisher 344 Rat Versus B6C3F1 Mouse

Source: Crouch and Wilson (1979)

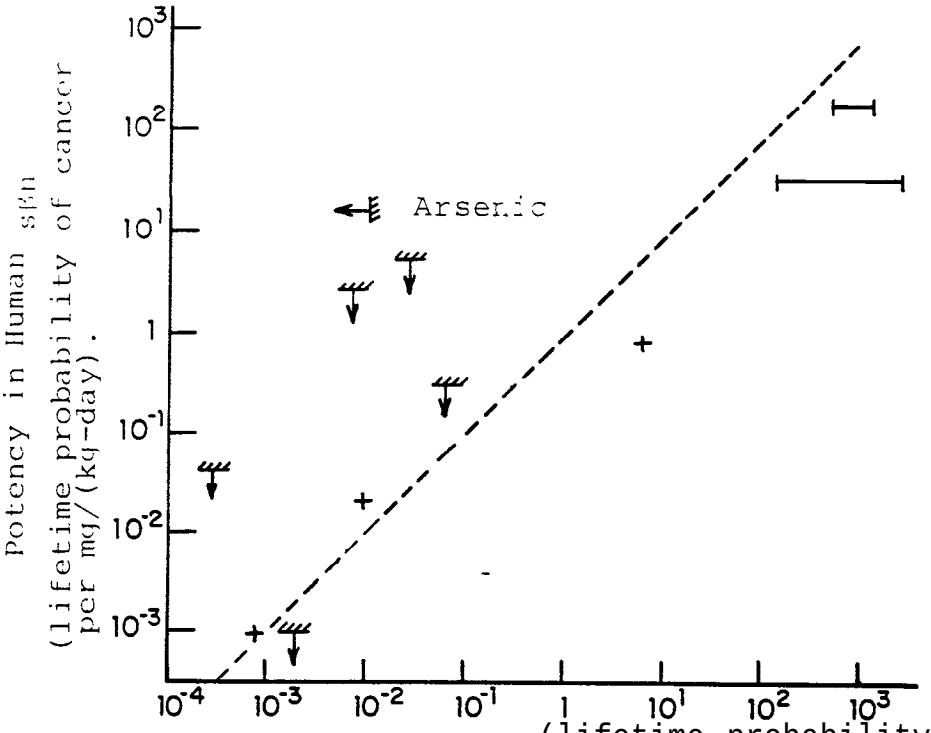


Figure 7. Carcinogenic Potency in Human Versus Rat

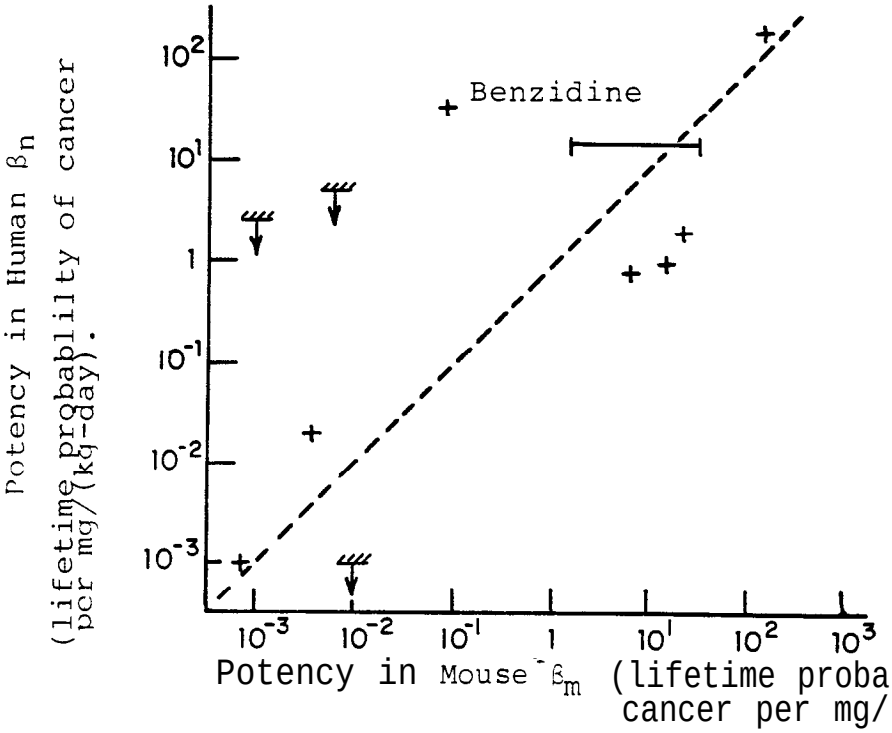


Figure 8. Carcinogenic Potency in Human Versus Mouse.

Source: Crouch and Wilson (1979)

humans and mice is not as good. In more recent work, Crouch and Wilson (1981) have suggested that it is appropriate to use unity as a central estimate of κ_{ha} (with dose expressed as mg/(kg-day)), and that an approximate estimate of $\ln K$ is 1.25.

DuMouchel and Harris (1983) have published an article describing an empirical Bayes' method for combining evidence from tests in several species. The new approach may eventually lead to more precise estimates of human potency. In an illustrative example, Harris and DuMouchel indicated that 95% confidence intervals for potency of diesel emissions could be reduced to a factor of approximately 9, i.e. $\ln 1.1$. However, at present it would seem that interspecies extrapolation is one of the weakest links in the risk assessment process.

Discussion

Attempts to assess quantitatively the risks to human health from chemical contamination resulting in exposures at low dose rats are fraught with uncertainty. The amount of uncertainty depends upon the contaminant, the pathways, the expected doses and dose rates, and the nature and extent of toxicological/epidemiological evidence.

Transport and dispersion models for non-reactive pollutants are better developed than those for pollutants which are chemically reactive or which are lost in transport due to deposition or absorption. Air pollution models are, in general, much better developed than ground-water models. Thus, estimates of the concentrations of conservative pollutants in ambient air are likely to be more precise than either estimates of the

concentrations of reactive pollutants in air or estimates of the concentrations of pollutants transported by ground-water.

Because interspecies extrapolation introduces uncertainty, risk estimates based on potency estimates derived from human data should, in principle, be more precise than those based upon potency estimates derived from animal data. However there are only about thirty contaminants for which human data exist. And the estimates of human exposures in these retrospective epidemiological studies are often so poor that the theoretical advantage in precision may be offset.

When expected doses and dose rates are orders of magnitude below those which were observed in epidemiology and/or bioassay, large uncertainties in estimates of low dose potency are introduced. The magnitude of these uncertainties increases dramatically as the differences between the expected environmental concentrations and the concentrations observed in epidemiology/bioassay increases.

The uncertainties in assessment of human health risks would not present a severe problem for policy analysis if their magnitudes could be estimated well. Techniques such as statistical decision analysis are well suited for policy analysis under uncertainty. And, if the magnitudes of the uncertainties in the components of risk could be estimated, methods for analysis of propagation of uncertainty could be used to derive estimates of the overall uncertainty.

However, as our investigation demonstrates, estimates of the uncertainties in many of the components are not widely available.

There are some estimates of the accuracy of air pollution transport models. And there are a few estimates of the precision of predictions of ground-water models. Data on about twenty chemicals provide rough estimates of the uncertainty introduced by interspecies potency extrapolation.

But there are many problems with these uncertainty estimates. They, in many cases, are not applicable to the situations of interest. For example, they may apply only to prediction of the transport of conservative pollutants. Or they may apply only to predictions of concentrations very close to the emissions source. And, in the case of interspecies extrapolation, they may not adequately reflect the variability of uncertainty and its dependence upon the specific contaminant, pathway, and test species involved.

Finally, the uncertainty introduced in low dose extrapolation is, at best, difficult to quantify. And it is virtually impossible to verify. The relationship between low-dose potency and the potency observed at high dose is dependent upon the choice of dose-response model. And this choice is subjective. No data or theory exist which unambiguously support the choice of a particular model.

Therefore, it would seem that in most cases overall uncertainties in risk assessment would be dominated by the uncertainty in determination of low-dose potency estimates and that, at present, it would be quite difficult to generate defensible estimates of the overall uncertainties due to the difficulty in estimating the uncertainty in determination of low dose potency.

NOTES

1. Variable aquifer material and hydraulic properties as a function of horizontal and vertical distance.

2. Personal comunication. William Rohrer, Senior Environmental Scientist, Pope-Reid Associates, Inc., St. Paul, Minnesota, July 1983.

3. Ibid.

4. See, for example, BEIR I (1972) and BEIR III (1980) reports of the National Academy of Sciences.

5. The average remaining length of life under the 1970 U.S. Life Table and age structure of the 1970 Census is about 45 years. Life expectancy at birth is approximately 70 years.

6. That is, for each tenfold reduction in dose, the probability of response is decreased by one standard normal deviate.

REFERENCES

"Air Quality Modeling and the Clean Air Act: Recommendations to EPA on Dispersion Modeling for Regulatory Applications", American Meterological Society, Boston, MA, 1981, p. 19.

Anderson, C. and J. Hushon. "Task Force Report on Health Effects Assessment. Prepared for Department of Energy" The Mitre Corp. MTR - 7966, McLean, VA, 1978.

Anderson, M.P., "Groundwater Modelling - The Emperor Has No Clothes", Editorial adopted from a talk delivered to the Minnesota Ground Water Association in May 1983.

Anderson, M. P. "Using Models to Simulate the Movement of Contaminants Through Ground-Water Flow Systems." Critical Reviews of Environmental Control, Vol. 9, No. 2.

Bass, A. "Modeling Long Range Transport and Diffusion", in Air Quality Modeling and the Clean Air Act: Recommendations to EPA on Dispersion Modeling for Regulatory Applications, Appendix E. American Meterological Society, Boston, MA, 1981.

Bowne, N. "Preliminary Results from the EPRI Plume Model Validation Project - Plains Site", Electric Power Research Institute, Report No. EA-1788-SY, TRC, Wethersfield, CT, 1981.

Brown, S.M. et al. "Report to Electric Power Research Institute", 1981.

Casarett, L.J. and J. Doull (Eds.). "Toxicology, The Basic Science of Poisons", Second Edition, Macmillan Publishing Co., Inc., NY, 1980.

Committee on the Institutional Means for Assessment of Risks to Public Health. "Risk Assessment in the Federal Government: Managing the Process". National Academy Press, Washington, D.C., 1983.

Conversation: Dr. Chou Chen, EPA, Exposure Assessment Group. July, 1983.

Conversation: Charles Norman, EPA, Exposure Assessment Group, July, 1983.

Cooper, D. "Revision of the USNRC's Radiological Health Effects Models." Progress Report Number 2, May 1983. Department of Environmental Science and Physiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115.

Crouch, E. and R. Wilson. "Interspecies Comparison of Carcinogenic Potency," Journal of Toxicology and Environmental Health. 5:1095-1118, 1979.

Crouch, E. and R. Wilson. "Regulation of Carcinogens," Risk Analysis, Vol. 1, NO. 1, 1981.

Crump, K. and R. Howe. "Approaches to Carcinogenic, Mutagenic and Teratogenic Risk Assessment". Science Research Systems, 1201 Gaines Street, Ruston, LA, August, 1980.

Crump, K., H. Guess, and K. Deal. "Confidence Intervals and Test of Hypothesis Concerning Dose-Response Relations Incurred From Animal Carcinoginicity Data". Biometrics 33, 437-451, Sept., 1977.

Duan, N.A., "Micro-environment Types: A Model for Exposure to Air Pollution, SIAM Institute for Mathematics and Society Report No. 47, Department of Statistics, Stanford University, Stanford, CA, 1981.

Evans, J., H.Ozkaynak, and R. Wilson. "The Use of Models in Public Health Risk Analysis", Journal of Energy and Environment, 1(1): 1-20, 1982.

Faust, C.R. and J.W. Mercer. "Ground-water Modeling: Numerical Models". Ground-water, (July-August, 1980) vol. 18, no. 4, pp. 395-409.

Federal Register, Friday, Nov. 28, 1983, Vol. 45, No. 231. Part V. Environmental Protection Agency, pg. 79323-79357.

Fiering, M.B., R. Wilson, E.A.C. Crouch, E. Kleiman, and L. Zeise. "Statistical Distributions of Carcinogenic Risk", Energy and Environmental Policy Center, Harvard University, April 1982.

Fiksel, J.R. and K.M. Scow. The Assessment of Human Exposure and Health Risks Using Outputs of Environmental Fate Models. Arthur D. Little, Inc., Cambridge, MA, 1983.

Freireich, E., E. Grehan, D. Rall, L. Schmidt, and H. Skipper. "Quantitative Comparison of Toxicity of Anticancer Agents in Mouse, Rat, Hamster, Dog, Monkey and Man". Cancer Chemotherapy Reports. Vol. 50, No. 4, May 1966.

FSC. "Proposed System For Food Safety, Chp. 11, Quantitative Risk Assessment". FSC, 1725 K. St. NW, Washington, D.C. 20006, June 1980.

Freeze, R.A. and J.A. Cherry. Groundwater. Prentice-Hall, Englewood Cliffs, NJ, 1979.

Hanna, S.R., G.A. Briggs and R.P. Hosker, Jr. Handbook on Atmospheric Diffusion. Technical Information Center, U.S. Department of Energy, Chapter 13, 1982.

Hogan, M.D. and D.G. Hoel, "Extrapolation to Man," in Hayes, A.W. (Ed.) Principles and Methods of Toxicology, Raven Press, New York, 1982.

ICRP Publication 23 - Report of the Task Group on Reference Man. Pergamon Press, Inc., Maxwell House, Fairview Park, Elmsford, NY 10523, 1981.

JRB Associates. "The Establishment of Guidelines for Modeling Groundwater Contamination from Hazardous Waste Facilities", Discussion Draft, McLean, VA, 1982.

Konikow, L.F. "Role of Numerical Simulation in Analysis of Ground-water Quality Problems", The Science of the Total Environment (1981) vol. 21, pp. 299-312.

Lave, L. and R. Romer. "A Survey of Safety Levels in Federal Regulators", U.S. Nuclear Regulatory Commission. NUREG/CR-2226, Washington, D.C., 1981.

Mahoney, J.R. "Meteorological Aspects of Air Pollution". Industrial Pollution. Van Nostrand Reinhold Co., New York, N.Y., 1974, pp. Chapter - 15.

Mancuso, R.L., C.M. Bhumralkar, D.E. Wolf and W.B. Johnson. "The Exchange of Sulfur Pollution Between the Various Countries of Europe Based on the EURMAP Model", Preprint Volume, Third Symposium on Atmospheric Turbulence, Diffusion and Air Quality, Raleigh, N.C., 1979, pp. 345-354.

Mantel, N., and M. Schneiderman. "Estimating 'Safe' Levels, A Hazardous Undertaking". Cancer Research 35, 1379-1386, June 1975.

Mercer, J.W. and C.R. Faust. "Ground-water Modeling: An Overview". Ground-water (March-April, 1980) Vol. 18, No. 2, pp. 108-115.

Mercer, J.M. and C.R. Faust. "Ground-water Modeling: Mathematical Models". Ground-water (May-June, 1980) Vol. 18, No. 2, pp. 108-115.

----. "Ground-water Modeling: Applications". Ground-water (Sept.-Oct., 1980). vol. 18, no. 5, pp. 486-497.

Meyers, R.E., R.T. Cederwall, J.A. Storch and L.I. Kleinman. "Modeling Sulfur Oxide Concentrations in the Eastern United States: Model Sensitivity, Verification and Applications",

Preprints, Fourth Symposium on Turbulence, Diffusion, and Air Pollution. American Meterorological Society, Reno, NV., 1979, pp. 673-676.

Morgan, M.G., M. Henrion and S.C. Morris, "Expert Judgements for Policy Analysts," BNL-51358, Brookhaven National Laboratory, Upton, N.Y., 1981.

NRC - National Research Council. "Risk Assessment in the Federal Government: Managing the Process", Committee on the Institutional Means for Assessment of Risks and to Public Health, Commission on Life Sciences, NRC, National Academy Press, Washington, D.C., 1983.

Ott, W.R. "Models of Human Exposure to Air Pollution," SIAM Institute for Mathematics and Society Report No. 32, Department of Statistics, Stanford University, Stanford, CA, 1980.

Pasquill, F. Atmospheric Diffusion. Halsted Press, New York, 1974.

Pope-Reid Assoc., Inc. "Technical Review of Ground-water Models". Final Report. Prepared for Office of Solid Waste, U.S. EPA. EPA Contract No. 68-01-6404, 1982.

Ruff, R.E. "Evaluation of the Real-Time Air Quality Model Using the RAPS Data Base", EPA-600/4-80-013a, Vol. I, Overview by SRI International, Menlo Park, CA, 1980.

Seiler, F., Appendix to the Fluidized Bed Combustion HEED, HERAP Program, U.S. D.O.E., Washington, D.C., 1982.

Severn, D.J., K. Offutt, S. Cohen, W.L. Burnan and G.J. Burin. "Assessment of Ground-water Contamination by Pesticides", Discussion paper prepared for FIFRA Scientific Advisory Panel Meeting June 21-23, 1983. Hazard Evaluation Div. Office of Pesticide Programs, EPA, 1983.

Smith, M.E. "Transport and Diffusion Modeling and Its Application",in Air Quality Modeling and the Clean Air Act: Recommendations to EPA on Dispersion Modeling for Regulatory Applications. American Meterological Society, Appendix J, Boston, MA, 1981.

Spengler, J.D., G.D. Thurston and H. Ozkaynak. "Meteorological Aspects of Air Pollution", Course Notes. Harvard School of Public Health, Boston, MA, 1982.

Turner, D.B. "Atmospheric Dispersion Modeling, A Critical Review". Journal of the Air Pollution Control Association (May 1979) Vol. 9, No. 5, pp. 502-519.

Van Ryzin, J. "Quantitative Risk Assessment", Journal of Occupational Medicine, 22(5): 321-326 (1980).

PART 10

THE VALUE OF IMPROVED EXPOSURE INFORMATION IN BENEFIT-COST ANALYSIS OF TOXIC SUBSTANCES

John Evans

I. INTRODUCTION

Decisions about the control of toxics and hazardous wastes often must be made amidst great uncertainty. In many cases relationships between emissions and exposures, exposures and doses, and doses and health risks are poorly understood. In addition there are complex issues surrounding the valuation of risks to human health. The resulting uncertainty complicates decision making. One issue which often arises is how to determine when it is appropriate to collect additional information.

Intuition suggests that a decision maker should collect additional information only if the value of the information is greater than the cost of obtaining it. Unfortunately this principal is often ignored.

Statistical decision analysis provides a framework for decision making under uncertainty and a method for estimating the value of information. (See, for example, Raiffa, 1968.) This paper illustrates how this method can be used to estimate the value of improved estimates of exposure.

II. ANALYTICAL FRAMEWORK

Implicit in any framework for decision making under uncertainty is the notion that even the best decisions may have bad outcomes. Additional information may be valuable because it may reduce the likelihood of bad outcomes. Its value depends upon both the incremental costs of these bad outcomes and the reduction in the probability of their occurrence.

An example may clarify this point. Table 1 gives the control costs and health risks associated with three strategies for the control of toxic emissions from a hypothetical industrial source. Under the assumption that each unit health risk corresponds to a one unit social cost total costs are minimized by selection of control strategy B.

This problem did not involve uncertainty and the decision was simple to make. In contrast, consider the situation in which this same decision must be made on the basis of imperfect information concerning health risks. See Table 2.

Here $\hat{R}$ is an unbiased, but imprecise, estimate of the true health risk. ($\hat{R}$ is normally distributed with mean 1 and standard deviation s.) Figure 1 shows the total cost estimate for each strategy as a function of the health risk estimate. Analysis of the figure indicates that estimated total costs are minimized under the following decision rules. If $\hat{R}$ is less than 0.50 select strategy A. If it is between 0.50 and 1.40 select strategy B. Otherwise select strategy C.

Table 1. Hypothetical Control Costs and Health Risks without Uncertainty

Strategy	Control Cost	Health Risk	Total Cost
A	0	1.00	1.00
B	0.25	0.50	0.75
C	0.81	0.10	0.91

Table 2. Hypothetical Control Costs and Health Risks with Uncertainty

Strategy	Control Cost	Health Risk Estimate	Total Cost Estimate
A	0	$1.00\ \hat{R}$	$1.00\ \hat{R}$
B	0.25	$0.50\ \hat{R}$	$0.25 + 0.50\ \hat{R}$
C	0.81	$0.10\ \hat{R}$	$0.81 + 0.10\ \hat{R}$

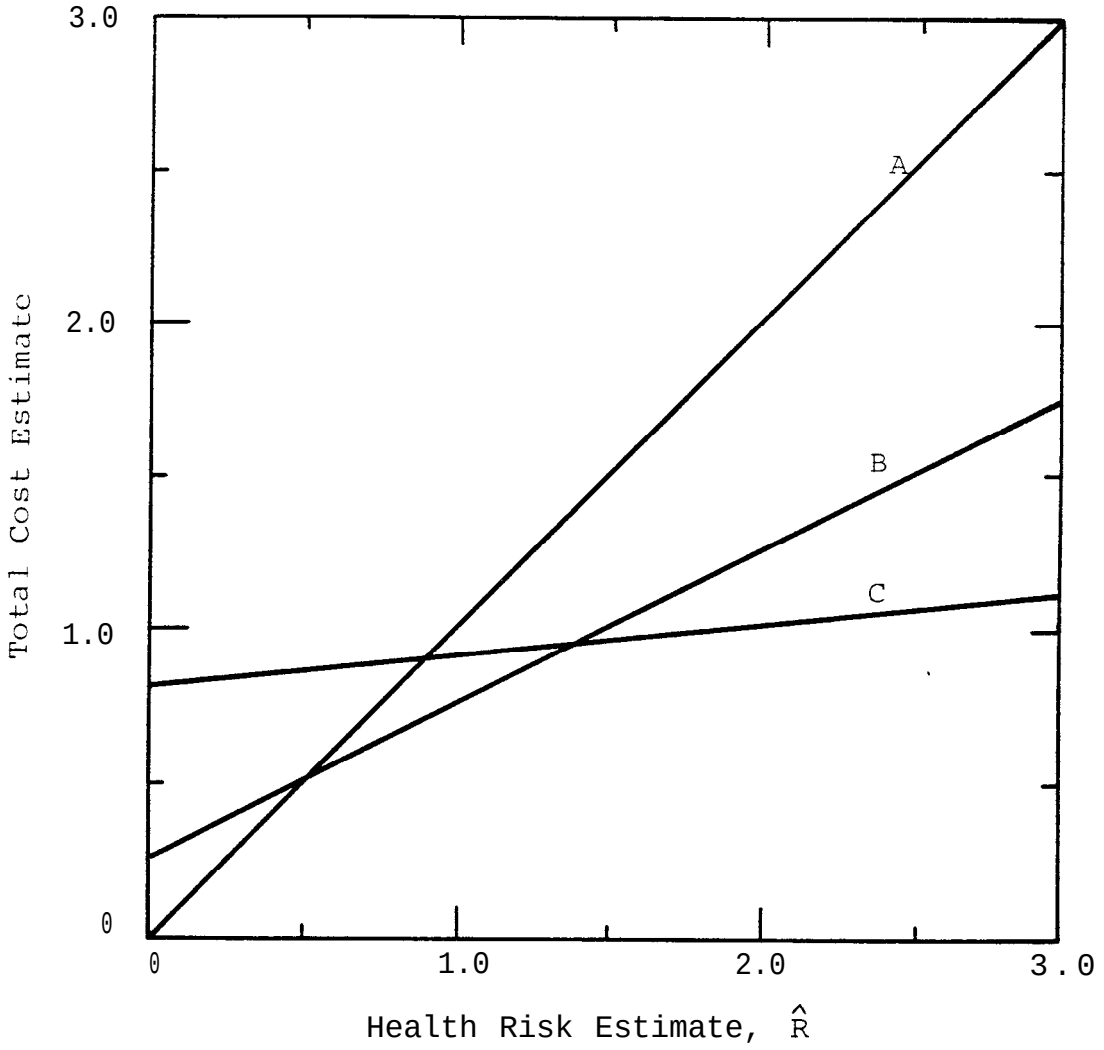


Figure 1. Dependence of Costs on Health Risks -- Hypothetical Case

If the true health risk were known, strategy B would be selected. Uncertainty in the estimates of health risks leads to random, and sometimes incorrect, decisions. The cost of this randomness may be determined by comparing the expected total cost of decisions made under uncertainty with the cost of decisions made under uncertainty would be:

$$E[TC] = P\{\hat{R} \leq 0.50\}1.00 + P\{0.50 \leq \hat{R} \leq 1.40\}0.75 + P\{\hat{R} \geq 1.40\}0.91 \quad (1)$$

The components of this equation are the probabilities and costs of choosing strategies A, B and C, respectively. With perfect information, strategy B would always be chosen, with a cost of 0.75. The difference between the expected total cost under uncertainty and the cost with perfect information is:

$$EOL = P\{\hat{R} \leq 0.50\}\ 0.25 + P\{\hat{R} \geq 1.40\}\ 0.16 \quad (2)$$

This quantity is known by decision analysts as the expected opportunity loss, EOL, or expected value of perfect information, EVPI. It is the most that a rational decision maker should be willing to pay to eliminate uncertainty.

The expected opportunity loss depends upon both the incremental costs associated with bad decisions and the probabilities of making bad decisions. These probabilities depend upon the amount of uncertainty in the health risk estimates. Figure 2 illustrates the dependence of the expected opportunity loss upon the degree of imprecision in health risk estimates. Notice the sensitivity of the expected opportunity loss to the standard deviation of the risk estimates.

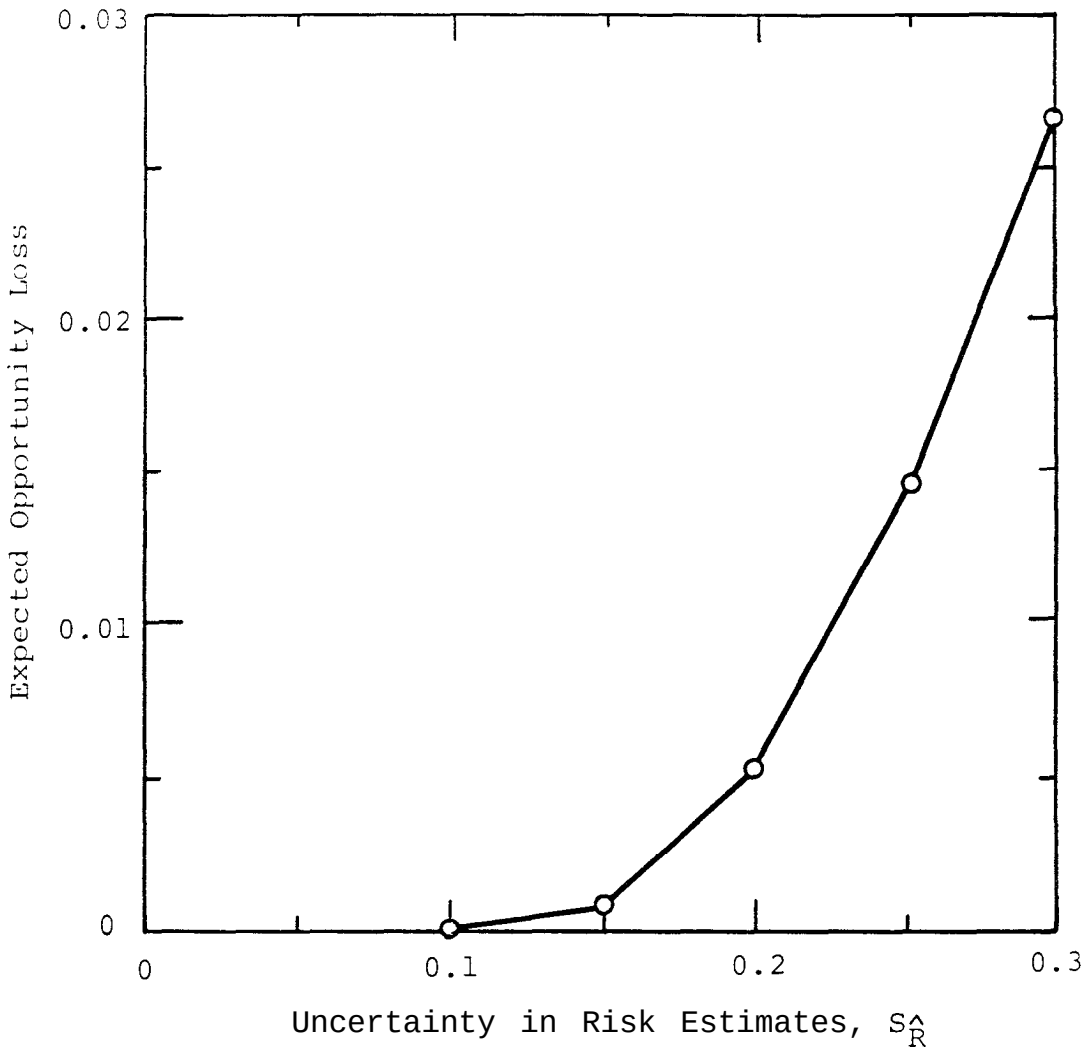


Figure 2. Dependence of Expected Opportunity Loss on Uncertainty in Risk Estimates -- Hypothetical Case

The example illustrates how uncertainty leads to incorrect decisions and to increased total costs. It also demonstrates that with information about the degree of uncertainty and the costs of incorrect decision making it is possible to estimate the expected value of perfect information.

In more realistic cases, it will be possible to reduce, but not eliminate, uncertainty. A more appropriate measure of the value of information in these cases is the difference between the expected opportunity loss without the information and the expected opportunity loss with the information. In the sections which follow we develop an approach for addressing this more complex case and demonstrate the application of our approach.

III. PROPOSED APPROACH

Our approach for estimating the value of improved exposure estimates in support of environmental decision making involves five steps.

(1) Estimate the uncertainty in health risk estimates made on the basis of current exposure estimates.

(2) Calculate the expected opportunity loss associated with decisions made under the current level of uncertainty.

(3) Estimate the uncertainty in health risk estimates which would remain once exposure estimates were improved.

(4) Calculate the expected opportunity loss associated with decisions made under the reduced level of uncertainty.

(5) Estimate the value of improving exposure estimates by comparing (2) and (4).

To apply this framework to the problem of decision making for toxic air emissions one must consider the sources of uncertainty in health risk estimates. Five steps are involved in the estimation of health risks under alternative control strategies. One must estimate:

(1) the emissions expected under each control strategy,

(2) the contribution of emissions to ambient concentrations,

(3) the contribution of changes in ambient concentrations to changes in human exposures,

(4) the contribution of changes in exposures to changes in doses, and

(5) the contribution of changes in dose to changes in risks.

Each of these steps involves uncertainty. The efficiencies of various control strategies are not precisely known. There are uncertainties in the models used to predict transport and dispersion of pollutants. The behavior patterns of people are not well known and it is these patterns which determine the relationship between concentrations and exposures. Breathing rates, clearance parameters, and other factors which govern the dose received from a given exposure are somewhat uncertain. And finally, models of dose-response are subject to many uncertainties. For many pollutants the dynamics of dose-response and the functional form of the relationship between dose and response, are poorly understood. In addition, the parameters of dose-response models must often be estimated from very limited data.

To estimate the uncertainty in health risk estimates it is necessary to first evaluate the uncertainty in each of these elements. Once these have been evaluated, methods for the analysis of propagation of uncertainty may be employed to determine both the total uncertainty and the contribution of each element to the total. (See for example, Bevington, 1969.) For example, under a proportional model of risk, one of several methods may be used to analyze the propagation of uncertainty. With a proportional model the health risk, $\hat{R}$, is estimated as the product of potency, $\hat{\beta}$ and dose, $\hat{D}$. If the errors in β', and D are small (compared to their typical values), independent, and symmetric then Gauss' Law of Error Propagation gives:

$$\bar{R} = \bar{\beta}' \bar{D} \tag{3a}$$

$$\sigma_R^2 \quad \bar{\beta}'^2 \sigma_D^2 + \bar{D}^2 \sigma_\beta^2, \tag{3b}$$

where σ_R^2 is the variance of the distribution of risk estimates, σ_D^2 and σ_β^2, are the variances of the estimates of the total dose to the population and the potency of the chemical for the biologically average individual, respectively, and $\bar{R}$, $\bar{D}$ and $\bar{\beta}'$ are the mean values of these same quantitities. Alternatively, if the estimates of β' and D are thought to be distributed approximately lognormally around their medians, then:

$$R_m = \beta'_m D_m \tag{4a}$$

$$\sigma^2_{\ln\hat{R}} = \sigma^2_{\ln\hat{\beta}'} + \sigma^2_{\ln\hat{D}} \tag{4b}$$

where R_m, β'_m and D_m are the geometric means (or medians) of the distributions of estimates of risks, potency and dose and $\sigma^2_{\ln\hat{R}}$, $\sigma^2_{\ln\hat{\beta}}$, and $\sigma^2_{\ln\hat{D}}$ are the variances of the distributions of the natural logarithms of these quantities.

These methods would permit one to estimate the fraction of total uncertainty due to uncertainty in dose estimates, and therefore to estimate the value of improving dose estimates. More advanced methods, such as Monte Carlo simulation, could be used to analyze the propagation of uncertainty in more complex models of risk.

One additional complexity must be considered. Our analysis has been based upon the assumption that the form of the dose-response model is known and that the only sources of uncertainty are uncertainties in the dose estimates and in the potency estimates. This is not always the case.

The bulk of our knowledge about the risks associated with exposure to environmental carcinogens comes from either small rodent bioassay or occupational epidemiology. In both cases the doses and dose rates tend to be several orders of magnitude above those likely to be encountered in the ambient environment. This would not present difficulties for risk assessment if the shape of dose-response curves were known. However, neither theory nor empirical evidence provides unambiguous support for one model. (See, for example, Van Ryzin, 1980.)

To illustrate the difficulty this presents for risk assessment consider the following example. Imagine that only two models of dose-response were plausible: a proportional model and a k^{th} order model. See Figure 3. In order for these two models to give similar estimates of risk in the range of experimentally

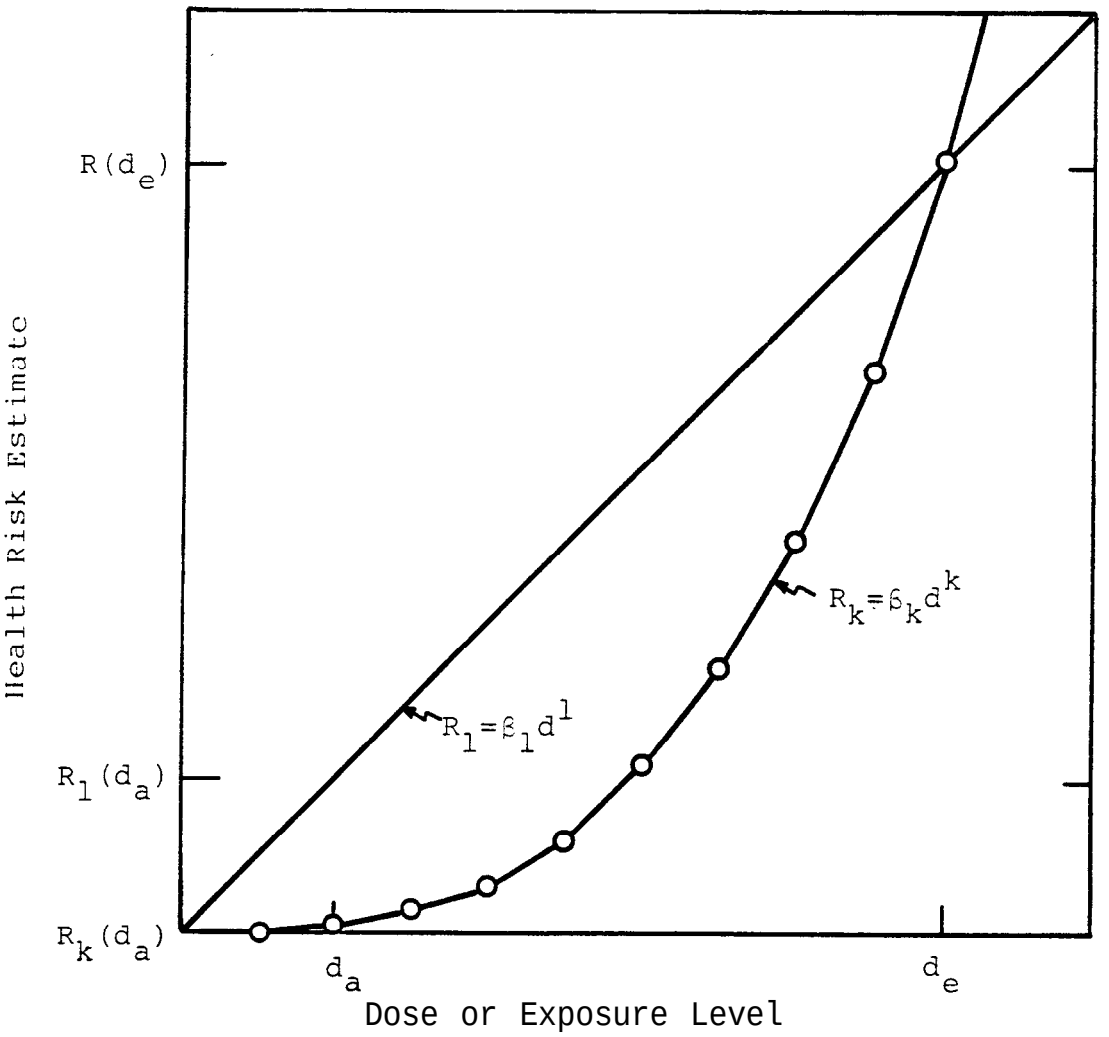


Figure 3. Dependence of Health Risk Estimate on Choice of Dose-Response Model $\left[d_a=d_e/10^P\right]$

observed doses, their parameters must be related. Thus their estimates of low dose potency and of risks at low doses are also related:

$$\frac{dR_1\ (d_a)/dd}{dR_k\ (d_a)/dd} = \frac{10^{p(k-1)}}{k} \tag{5a}$$

$$\frac{R_1\ (d_a)}{R_k\ (d_a)} = 10^{p(k-1)} \tag{5b}$$

Where p is the number of orders of magnitude that ambient doses are below experimental doses, and d_a is the level of typical ambient doses. It is obvious from these relationships that when ambient doses are several orders of magnitude below experimental doses, the uncertainty as to the form of the dose-response model may lead to quite large uncertainties in health risk estimates.

A simple modification to our approach for estimation of the value of improved exposure estimates accommodates this additional complexity. In the case of model uncertainty the probability density function for health risk estimates is generated using:

$$\text{pdf}\ (\hat{R}) = p_1\ \text{pdf}\ (\hat{R}_1) + \ldots + p_k\ \text{pdf}\ (\hat{R}_k) \tag{6}$$

which involves a sum of products of conditional probability density functions and estimates of the probabilities that each of k possible models is correct. The remainder of the analysis is unchanged.