DEVELOPMENT OF ENVIRONMENTAL-SUPPORT SOFTWARE TO DETERMINE AND COMPARE DOSE-RESPONSE EXTRAPOLATIONS

By

BRYCE K. SMITH

Bachelor of Science

University of Oklahoma

Norman, Oklahoma

1988

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 1994

DEVELOPMENT OF ENVIRONMENTAL-SUPPORT SOFTWARE TO DETERMINE AND COMPARE DOSE-RESPONSE EXTRAPOLATION

Thesis Approved

William F. McTernan

Thesis Advisor

Gregory G. Wilber

John M. Veazey

Thomas C. Collins

Dean of the Graduate College

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to Dr. Bill McTernan for his advice and encouragement throughout my studies in the graduate program. I also wish to thank Dr. John Veenstra and Dr. Greg Wilbur for serving on my graduate committee.

I wish to express my deepest affection, appreciation, and sincere thanks to my wife, Kristi Smith, for her moral support as well as her typing and proofreading skills. I wish to express to my son, Bryce-David Smith, all my love and let him know that I missed a lot during his small life while pursuing my degree, but I hope when he is older, he will realize I did it all for our family. I wish to thank my mother-in-law, Sharon Beumeler, for her valuable skills in creating my cover pages, helping in preparation of my documents, and all her home-cooked meals. I also wish to thank my father-in-law, David Beumeler, for his help in babysitting my son throughout my graduate studies.

I wish to thank my parents, DeLacey and Louise Smith, for their support throughout my school years. Without them I would not have been able to accomplish this goal. I wish to express sincere appreciation to my Grandparents, Orville and Felma DeLapp and Ruth Smith, for all the financial help they have given me through my years of schooling. I also wish to acknowledge my late Grandfather, Forrest Smith, who was always a great inspiration in my life.

I wish to extend a special acknowledgement to my late Grandfather, Orville DeLapp, who passed away on August 1, 1994. He was a dear friend, grandparent, and great inspiration to me throughout my life. God bless him and all of my family that have supported me through the years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The following was an effort to develop a computer program that: (1) presented the user with a valuable dose-response assessment tool, (2) helped the user in the process of environmental decision making, (3) was user friendly, and (4) required little computer knowledge to operate. The computer program that was developed, called Q-Risk, could be considered a "Decision Support System" which is defined as "an interactive data processing and display system used to assist in a concurrent decision-making process, and also conforms to the following characteristics:

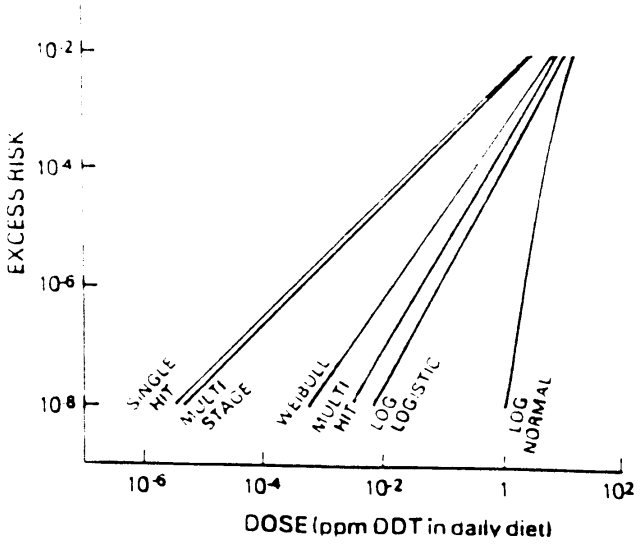
- it is sufficiently user-friendly to be used by the decision maker(s) in person.
- it displays its information in format and terminology which is familiar to its user(s).
- it is selective in its provision of information and avoids exposing its users(s) to an information overload." (Simons, 1985).

In the past, programs were written in less "user friendly" languages that did not supply the user with ample information to accurately address their concerns. Also, the user generally had to be extremely computer literate to operate these programs. Q-Risk was an attempt to bridge this barrier between

user and computer in the area of dose-response extrapolations. These extrapolations are made necessary when utilizing high-dose, short duration animal toxicity testing to determine potency or slope factors necessary to determine unit health risks from environmental contaminants. As such, this code is intended to be used within the Environmental Risk Assessment process.

Q-Risk was designed to aid the scientist in the "Toxicity Assessment" step by the incorporation of dose-response models for low-dose extrapolation of quantal bioassay data. Animal models are acting surrogates for humans subjected to high exposure levels to initiate a response. Then by the use of mathematical models the data are extrapolated to the low-dose region more typically found in environmental exposures. From this low-dose extrapolation a slope or potency factor is determined. The slope factor, in units of $(mg/Kg\text{-}day)^{-1}$ is multiplied by the dose in mg/Kg-day units to determine an incremental excess cancer probability. Comparison of high to low dose extrapolation for six dose-response models (Brown, 1984) is presented in Figure 1. A residual exposure producing "Acceptable incremental risk," (i.e. one in one million) can be determined from this figure. Table 1, the Goodness of fit statistics for the data used in Figure 1, presents how well the six various models fit the observed data where virtually safe dose is represented by VSD (Brown, 1984). The chi-square (χ^2) value shows how well the model fits the data. A high χ^2 value corresponds to a "good-fit" and subsequently a low p-value (probability).

FIGURE 1
COMPARISON OF HIGH TO LOW DOSE EXTRAPOLATION FOR 6 DOSE-RESPONSE MODELS



Source: *(Brown, 1984)*

TABLE 1
GOODNESS OF FIT FOR FIGURE 1

Extrapolation Model	VSD* (ppm DDT in daily diet)	Goodness-of-fit Statistic of Model to Observed Data χ^2	(d.f.)	P-value
Log normal	6.8×10^{-1}	3.93	(2)	0.14
Weibull	5.0×10^{-2}	3.01	(2)	0.22
Multihit	1.3×10^{-2}	3.31	(2)	0.19
Log logistic	6.6×10^{-3}	3.45	(2)	0.18
Multistage	2.5×10^{-4}	---------**		
Single-hit	2.1×10^{-4}	5.10	(3)	0.16

* 97.5% lower confidence limit on VSD computed by the likelihood method described in (22)
** no goodness-of-fit statistic since the number of parameters equals the number of data points

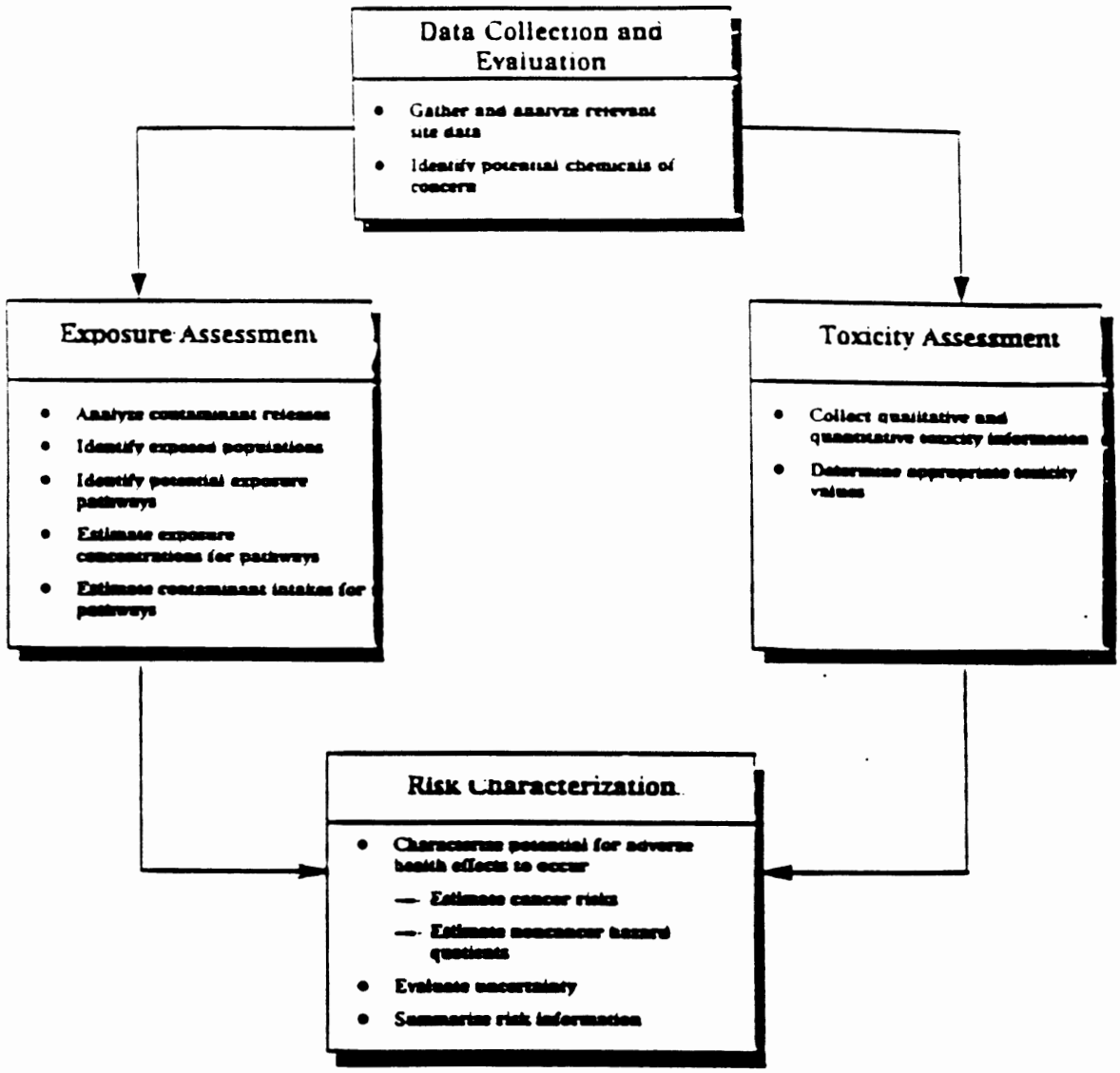
Source: *(Brown, 1984)*

To combine the science of low-dose extrapolation and computers, Microsoft* Quickbasic, version 4.5, was used to generate and compile code necessary to extrapolate laboratory toxicity data by means of several alternative formulae. Plots are produced comparing the unit risk (risk associated with its corresponding dose) associated with each calculation. This programming approach was chosen because of the ability to generate user friendly graphical screens, and to calculate lengthy algorithms with Quckbasic. The resultant graphics allows the end user, the risk assessment engineer or scientist, the opportunity to easily and visually compare toxicological extrapolations with a range of techniques. Q-Risk was designed to allow the user to choose between five dose-response models: Probit, Log-Logistic, Weibull, One-Hit, and Multistage. The Multi-Hit model was described but was excluded from computational applications due to its similar extrapolation characteristics with the Weibull model (Brown, 1984). Help screens were generated to guide the user in selecting an appropriate model, and to guide the user easily through the program.

Risk assessment is the process by which scientists "determine the nature and magnitude of risk associated with various levels and conditions of human exposure to a carcinogen and non-carcinogen." (Rodricks & Tardiff, 1984). The data could come from two separate sources: (1) epidemiological studies or (2) experimental data from animal studies (Fenner-Crisp, 1986). The epidemiological source is not practical because it either involves subjecting humans to the

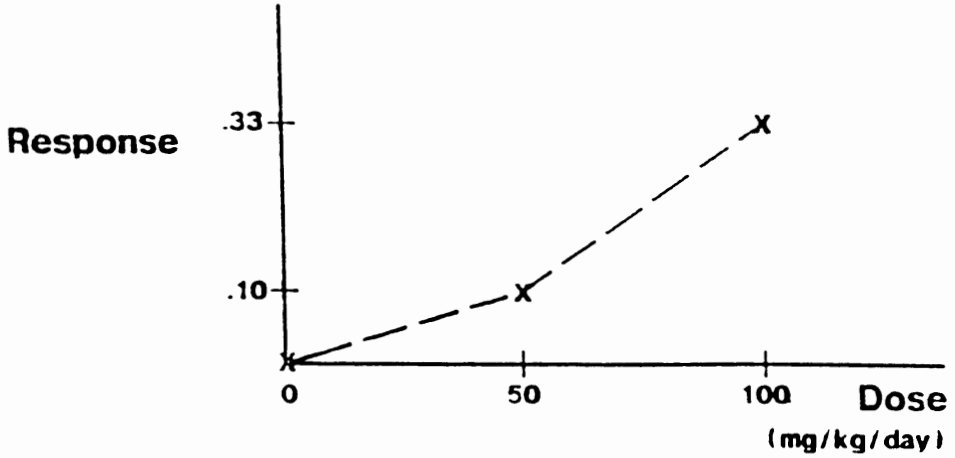
exposure of potentially hazardous chemicals or compiling worker exposure data which generally lack statistical rigor. Figure 2 represents the four steps involved in determining human health risk (EPA, 1989). The data collection and evaluation step "... involves gathering and analyzing the site data relevant to the human health evaluation and identifying the substances present at the site that are the focus of the risk assessment process," while exposure assessment "... is conducted to estimate the magnitude of actual and/or potential human exposures, the frequency and duration of these exposures, and the pathways by which humans are potentially exposed." "Toxicity Assessment for contaminants found at Superfund sites is generally accomplished in two steps: hazard identification (identifying which contaminants are hazardous) and dose-response assessment." Risk Characterization "... summarizes and combines outputs of the exposure and toxicity assessments to characterize baseline risk, both in quantitative expressions and qualitative statements." (EPA, 1989). Q-Risk was prepared to address the dose-response component found in the Toxicity Assessment element. A typical dose-response curve is represented by Figure 3A plots dose (mg/Kg/day) versus response (which could be a death or any adverse effect) in the observable range (Environ Corp., 1987). Also shown is the linear extrapolation from the typically high dose experimental range to the near origin region where chronic, long-term effects are expected to occur. Figure 3B plots the observed data and the Weibull and Multi-stage models as fitted to the data. At a dose of 50 mg/Kg/day a "threshold" is observed in the Weibull plot (indicated by the sharp increase in

FIGURE 2
THE BASELINE RISK ASSESSMENT PROCESS



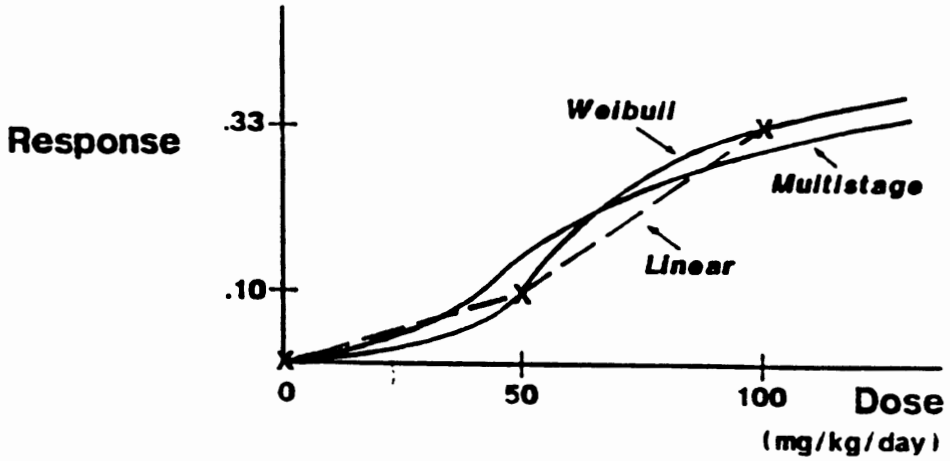
Source: (EPA, 1989)

FIGURE 3A
DOSE-RESPONSE CURVE-LINEAR



Source: (Environ Corp., 1987)

FIGURE 3B
MODEL COMPARISON



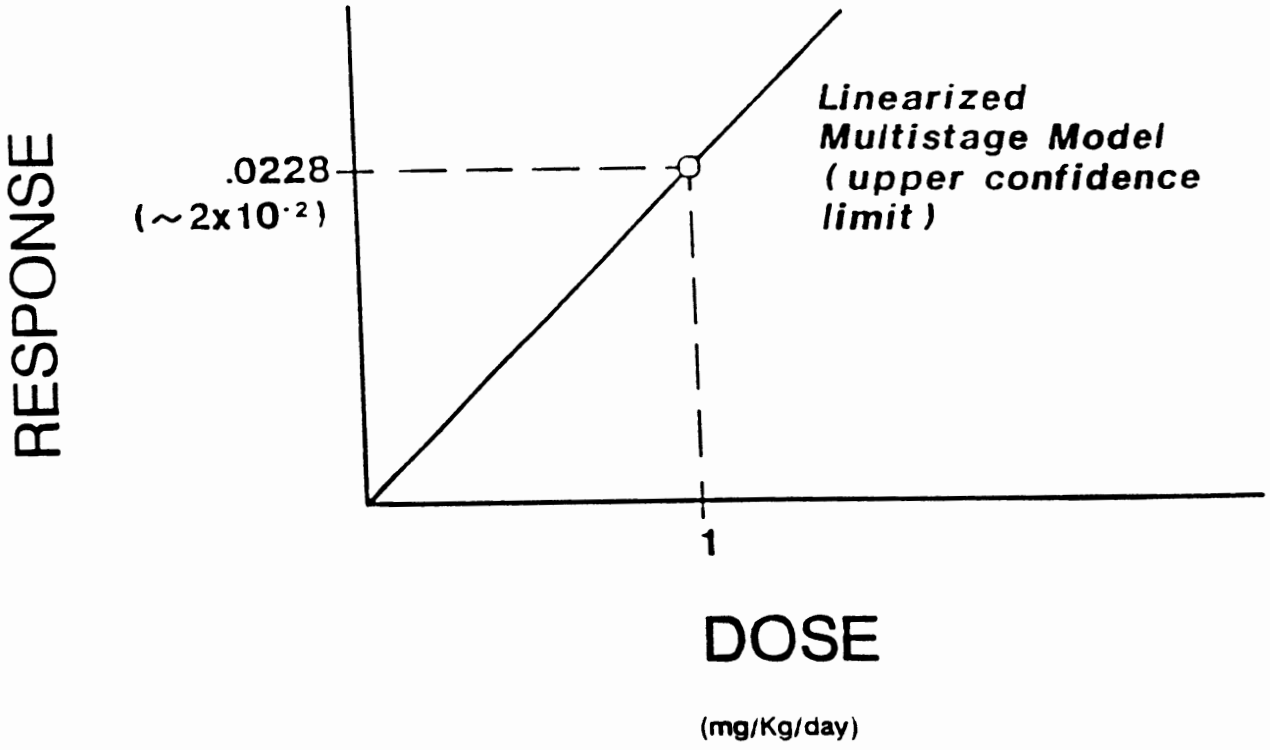
Source: (Environ Corp., 1987)

slope). This threshold can be used to assume a point below which greater safety or lowered risk is observed. The other two models linearize the data near the origin, thereby perfecting greater unit risks for comparable doses. The current administrative position of the USEPA is that thresholds do not exist for carcinogens. The main reason for this is to be more conservative in determining a dose relating to acertain risk. This allows safer protection for the population as a whole. To assign a threshold to a general population would be a gross mis-interpretation, because of the vast genetic variability within a population. This latter view is shared by most toxicologist. Also, to obtain thresholds, research would take enormous resources and even if found would be suspect (Rall, 1978).

Figure 4 plots dose (mg/Kg/day) versus response (Environ Corp., 1987) and represents the method by which EPA extrapolates lab data to low-dose regions. A risk is selected by assigning a unit exposure dose of 1 mg/Kg/day. In the plot the corresponding risk is 0.0228, which means that incremental cancer risk per unit dose of chemical is 2.28×10^{-2}. This translates to one incremental cancer in 44 potential exposures. In general, regulatory levels of one in one million exposures are considered "acceptable." The USEPA recommends either the linear Multistage or the One-Hit model equation to estimate the risk associated with high carcinogenic risk levels (EPA, 1989).

USEPA has established computerized data bases such as the Toxic Substances Release Inventory, the Chemical List, Information Pointer System, MIXTOX, and the Integrated Risk Information System (IRIS) to provide agency

FIGURE 4
SLOPE FACTOR



Source: (Environ Corp., 1987)

estimates of these slope factors and related information (Shoeny, 1991). These data bases provide estimates of low-dose extrapolation from single models.

In summary, all of the models employed fit the high-dose data within acceptable statistical ranges. Extrapolation to low-doses, however, shows significant variation, as observed in Figure 1. Using one incremental incidence of cancer per million of population as exposed an acceptable incremental risk illustrates some of the uncertainties associated with model selection. While essentially endemic to the current state of epidemiological knowledge, this uncertainty reduces confidence in the resultant assessments.

The following sections include descriptions of the assumptions inherent in each of the models, coding of the Q-Risk program, a report of the results from data analyzed with Q-Risk, a comparison and discussion of these results, and a conclusion stating what has been accomplished.

CHAPTER II

MATHEMATICAL MODELS USED FOR DOSE-RESPONSE EXTRAPOLATIONS

"Cancers are believed to be single cell in origin .... Of a large number of cells at risk in the individual organism, one undergoes certain changes that allow it to divide and grow into a tumor. Thus we can view the carcinogenic process as mechanistically single cell in origin even though, by the time a cancer is pathologically recognizable, very extensive changes may have developed. ... If the individual cancers arise from an original, single, "transformed" cell, then the statistical nature of the carcinogenic dose-response will be governed by the extreme tail of the "transformation" response distribution. The effect of this is to make virtually any process of discrete events approximately linear at low dose." (Crump, Hoel, Langley, and Peto, 1976). This means it would be linear in the sense that the slope would be equal to one and the shape of the dose-response curve would be linear and not convex or concave.

The exact mechanisms for most environmentally induced diseases are not fully understood. To bridge this knowledge gap, the environmental toxicologist employs dose-response testing where laboratory models (i.e. animals, protists, etc.) are subjected to the chemicals of concern at defined dosage levels and for specific time periods. Typically, these are high-dosage, relatively short duration

tests to optimize laboratory resources while providing information in a timely manner (Brown, 1984).

Most human exposures are chronic rather than acute in nature. These involve low-doses over extended periods. As such, laboratory models can not be applied directly to predications on human systems. To accomplish this conversion from acute to chronic exposures, mathematical extrapolations from the testing region to the typical exposure levels are attempted. The models available for this are of two types:

- mechanistic
- tolerance distribution

Mechanisitic models assume that for carcinogenesis to occure a normal cell must be exposed to a certain number of hits by a toxicant. Tolerance distribution models assume that each individual has a unique exposure level or tolerance to a toxicant. Both type has several functional forms available to make these extrapolations. Selection is dependant upon either the underlying biological mechanisms of disease initiation or with fitting data with various statistical distributions (i.e. normal, log-normal, or Weibull).

Extrapolations to low-dose regions result in either low-dose linearizations, as indicated in the previous quote (Crump et al, 1976), or in the formation of a concentration threshold below which a response (disease) will not occur. Rall (1978) states "Many diseases resulting from exposure to foreign chemicals are delayed in their onset and, to some extent at least, are irreversible. That is, if

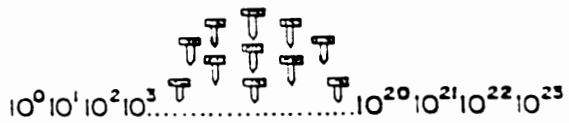
the chemical is removed, the disease continues to progress, or at least not regress. Typical are the diseases called cancer." (Rall, 1978). Figure 5 presents this threshold theory of chronic irreversible toxic effects. Part (a) of Figure 5 illustrates the concept of assigning a concentration above which a deleterious effect is observed in any animal or human and below which there is no effect observed (i.e. a threshold). Part (b) illustrates the uncertainty of which the concentration should represent a threshold. Part (c) illustrates the point that if a threshold is assigned to a particular person or part of a population, then the question is to whom and when is this threshold applied. Figure 5 illustrates the threshold, but does not illustrate this important question.

There are six commonly used mathematical dose-response models for this high-to-low dose extrapolation in animal test subjects. These are the Probit, Log-Logistic, Weibull, One-hit, Multi-hit, and the Multi-stage models (Brown, 1984). Dose-response refers to the response of a subject to various levels of a stimulus (dose). The response may be quantified in terms of the number of tumors, birth defects, deaths, etc. observed. The dose concentration may be quantified in terms of dietary percent or volumetric concentrations ingested, inhaled or dermally contacted. The source of exposure can be either by air, food, soil or other external stimulus.

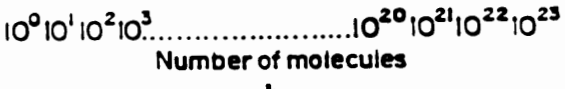
Once the data have been extrapolated to the low-dose region an "acceptable incremental risk" can be calculated corresponding to a specific low level of response. Typically, for carcinogens, 95% upper bound confidence level

FIGURE 5

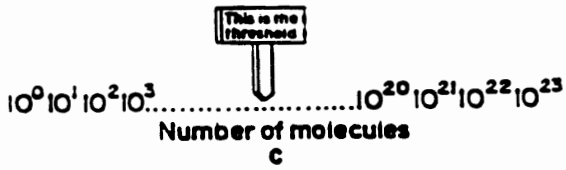
THRESHOLDS FOR CHEMICALS THAT CAUSE CHRONIC IRREVERSIBLE DAMAGE



a



b



c

Source: (Rall, 1978)

or percentile is determined about the unit response (USEPA, Means, 1989). This value, termed a slope or potency factor is used in conjunction with chemical exposure levels to calculate probabilities of incurring excess cancers.

The first three models are considered tolerance distribution models, while the last models are mechanistic based models.

Tolerance distribution models

Tolerance distribution models assume that each individual in an exposed group has a unique level of tolerance to a toxicant, the level of dose below which the toxicant is ineffective in producing an effect (i.e. the threshold). The only difference among the three tolerance distribution models is the assumption regarding the mathematical character of the distribution of response frequency. These models (Probit, Log-Logistic, and Weibull) possess a common assumption that there is a specific dose at which a subject will produce a quantal response (Brown, 1984). Above this concentration, a response is certain; below it there is no response. This is considered the subject's tolerance (Brown, 1984).
Equation 1 gives a mathematical expression of the frequency distribution of tolerances, f(D) (Brown, 1984). This frequency distribution can be thought of as the range of tolerances for a population

$$f(D) = \partial P(D)/\partial D \qquad (1)$$

where:

$\partial P(D)$ = Partial derivative as a function of dose (D)
∂D = Partial derivative in terms of dose (D)

and ∂D (the difference between the doses corresponding to each subjects tolerance level within the population) is small. This represents the proportion of subjects whose tolerances lie between D and D + ∂D. If all the subjects have a tolerance below or equal to an exposure dose, D_0, then all of them will produce a response. The proportion, $P(D_0)$, that represents the total population responding is represented by equation 2,

$$P(D_0) = \int f(D)\partial D, \qquad (2)$$

where the integral is evaluated in the range $0 \leq D \leq D_0$. If it is assumed that all the subjects would respond to a considerably large dose level, then equation 2 becomes:

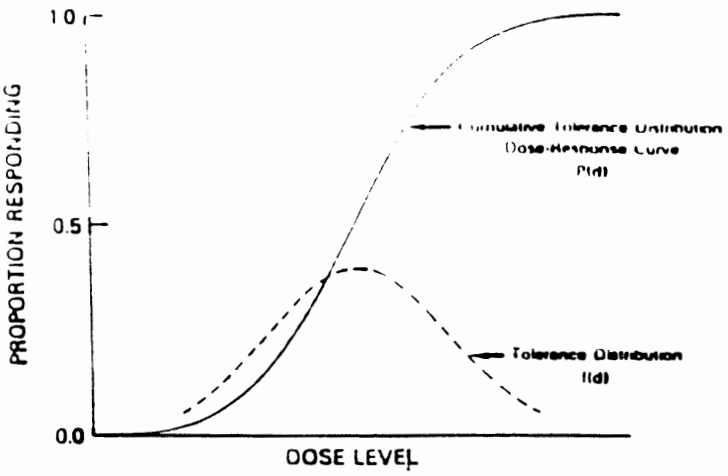
$$P(\infty) = \int f(D)\partial D = 1, \qquad (3)$$

where the integral is evaluated in the range $0 \leq D \leq \infty$. Figure 6 compares a tolerance frequency distribution, f(D), with its similar cumulative distribution, P(D). This shows that the dose-response can be viewed as being represented by the function P(D) for a whole population or a randomly selected individual (Brown, 1984).

Most often the frequency distribution of tolerances is skewed to one side as seen in Figure 7. This figure illustrates the frequency of response versus the concentration for the tolerance concentrations of a population (Finney, 1971). When a common logarithm transformation is applied to the scale of measurement (i.e., expressing the tolerances in terms of the common logarithm of concentrations), the distribution can resemble the Gaussian or normal

FIGURE 6

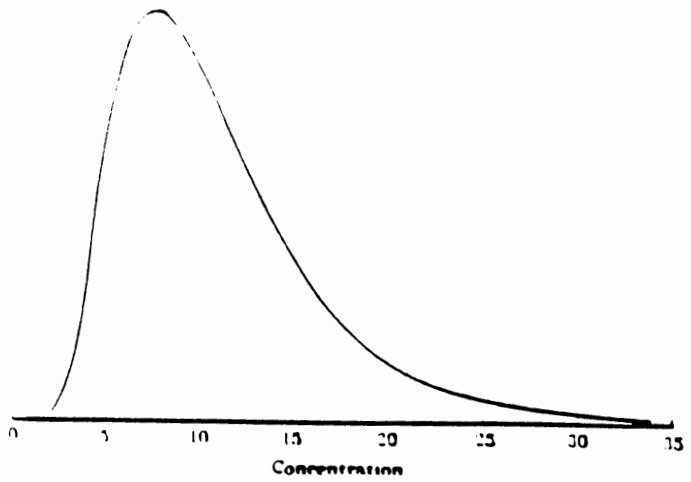
RELATIONSHIP BETWEEN TOLERANCE DISTRIBUTION AND DOSE-RESPONSE CURVE



Source: (Brown, 1984)

FIGURE 7

EXAMPLE OF SKEWED FREQUENCY DISTRIBUTION OF TOLERANCES



Source: (Finney, 1971)

distribution as seen in Figure 8 (Finney, 1971). Figure 8 plots the logarithm of the concentration versus the frequency of response to produce a normal distribution curve. The significance is that they illustrate the use of log transformation to fit data to a symmetrical tolerance distribution.

Probit (Log-Normal) Model. Equation 4 presents the probit or log-normal model for the tolerance frequency distribution (Finney, 1971). "Gaddum proposed to measure the probability of response on a transformed scale, the normal equivalent deviate (or N.E.D)" (Finney, 1971). N.E.D is represented by the dose corresponding to probability in a norma distribution with mean zero and variance one (N.E.D = y).

$$P(D) = (1/\sigma(2\pi)^{0.5})\exp\text{-}(\log 10D - \mu)2/2\sigma 2)\partial D, \quad (4)$$

where:

σ = standard population deviations
μ = mean
∂D = partial derivative in respect to dose
P(D) = probability of a response as a function of dose
$-\infty < \log_{10}(D) < +\infty$ (Finney, 1971).

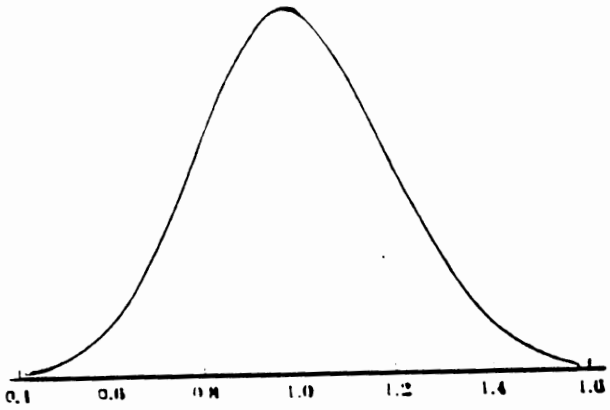
The dose-response function, P(D), is represented by equation 5 (Food Safety Council, 1980). Equation 5 is a result of the integration of equation 4. Y represents the response metameter which is a result of the probability log transformation and Y + 5 is the probit of P. (Food Safety Council, 1980)

$$Y = P(D) = \Phi[(\log(D) - \mu)/\sigma] = \Phi(\alpha + \beta\log(D)) = \alpha + \beta\log(D)$$

$$Y = \Phi^{-1}(P(D)) = \alpha + \beta\log(D) \quad (5)$$

FIGURE 8

NORMAL FREQUENCY DISTRIBUTION FOR THE LOGARITHM OF THE TOLERANCE CONCENTRATIONS



Source: (Finney, 1971)

where:

$\Phi(x)$ = Standard normal integral from $-\infty$ to x
α = $-\mu/\sigma$ (referred to as y-intercept)
β = $1/\sigma$ (referred to as slope)

The final equality in Equation 5 replaces the parameters μ and σ with α and β.

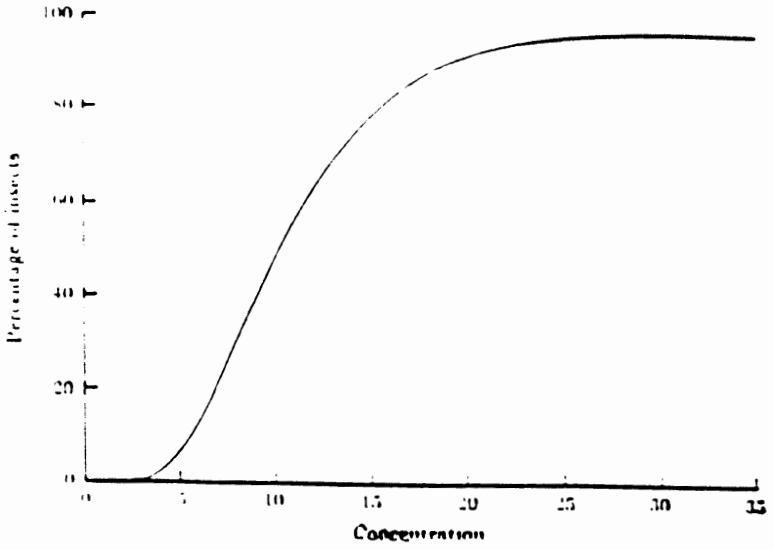
The idea of incorporating population statistics to determine tolerance distributions or dose response functions was first introduced by Gaddum (1933) and then by Bliss (1934) (Brown, 1984). Bliss (1934) looked at the effectiveness of a poison to kill Aphis rumicis L. Bliss observed an asymmetrical S-shaped curve when dosage was plotted directly against response, and stated that a common logarithmic plot of the dosage versus response in "probits" might have to be done to show a uniform dose-response distribution. Bliss suggested that the response interval be from 0.01% to 99.99%. This interval would then be transformed into a range of probits from 0 to 10, with 50% equaling 5 probits.

Probit transformation originated with psychophysical investigators. Their problem was quantifying the effect of stimulus on human subjects whose statements were measured as "right or wrong" or "greater than or lesser than" answers.

As indicated in Figure 9 the rate of increase of response per unit of dose is minimal in the zero and 100 percent ranges, but is sharp between the lower and upper responses. This produces a sigmoid curve as seen in Figure 9, which is derived from the skewed frequency tolerance distribution presented in Figure 7 (Finney, 1971). Figure 9 is an example of a plot of percentage responding against

FIGURE 9

SIGMOID CURVE DERIVED FROM FIGURE 7



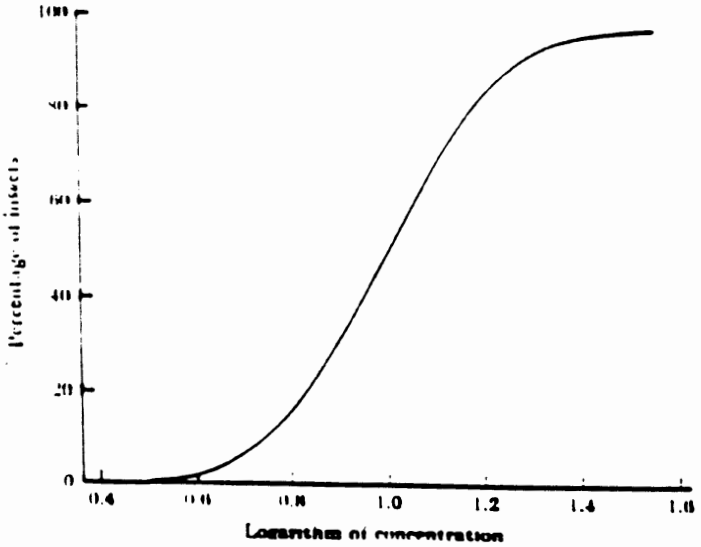
Source: (Finney, 1971)

dose. When the doses were transformed to the common logarithms, the tolerances became normally distributed as seen in Figure 10, which plots the logarithm of concentration versus percentage of insects affected (Finney, 1971). This shows that dose approaches zero at infinitely small values, but is limited at infinitely high doses. It is limited at infinitely high doses because all subjects will produce a response (shown by the uper "flat" portion of the s-curve.

Gaddum proposed the transformation of response to the normal equivalent deviate (N.E.D.). This is represented by Y where Y + 5 equals the probit of the response (Food Safety Council, 1980). Figure 11 shows the effect on the frequency of response by this probit transformation (Finney, 1971) while Table 2 gives the resulting probit corresponding to each percent mortality (Bliss, 1935). This table is useful in transforming a percent response into the corresponding probit. For example for a percent response of 10 the corresponding probit is 3.7184 (Bliss,1935).

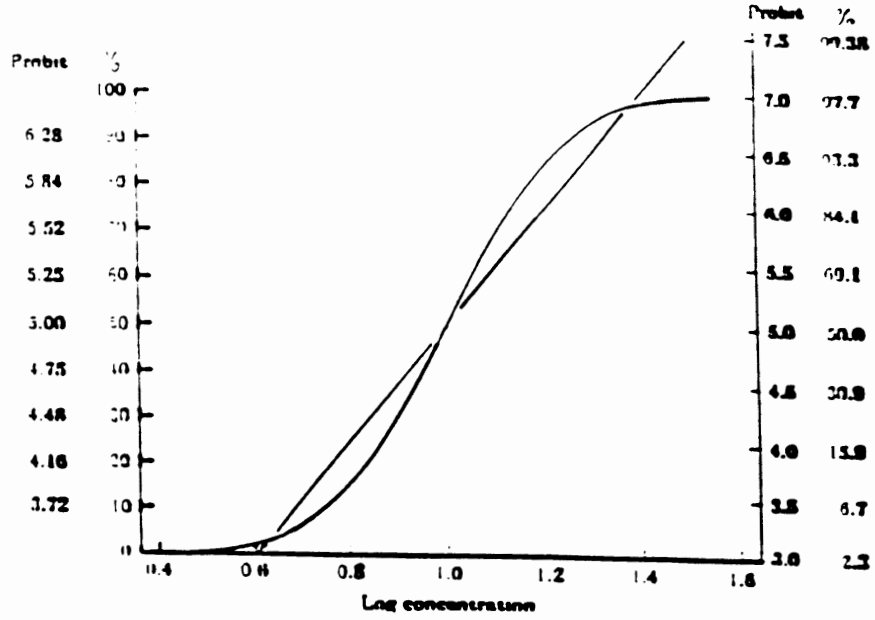
This model was originally incorporated in the area of drug standardization, where the responses in the 5 to 95% range were of most interest in assessing the potency of drugs. Therefore no threshold was assumed for the individual tolerances (Food Safety Council, 1980). When measuring the response directly, when a delay between the time of exposure and a response was observed, the tolerance dose could be overestimated (Finney, 1971). Historically this model was used for dose-response interpolation (observable range) instead of extrapolation (outside observable range) (Brown, 1984). Mantel and Bryan

FIGURE 10
NORMAL SIGMOID CURVE DERIVED FROM FIGURE 8



Source: (Finney, 1971)

FIGURE 11
EFFECT OF PROBIT TRANSFORMATION



Source: (Finney, 1971)

TABLE 2

PROBITS OR PROBABILITY UNITS FOR TRANSFORMING THE SIGMOID DOSAGE-MORTALITY CURVE TO A STRAIGHT LINE

	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
0	—	1·9908	2·1218	2·2322	2·3470	2·4242	2·4879	2·5427	2·5911	2·6344
1	2·6737	2·7096	2·7429	2·7738	2·8027	2·8299	2·8556	2·8799	2·9031	2·9251
2	2·9463	2·9665	2·9859	3·0046	3·0226	3·0400	3·0569	3·0732	3·0890	3·1043
3	3·1192	3·1337	3·1478	3·1616	3·1750	3·1881	3·2009	3·2134	3·2256	3·2376
4	3·2493	3·2608	3·2721	3·2831	3·2940	3·3046	3·3151	3·3253	3·3354	3·3454
5	3·3551	3·3648	3·3742	3·3836	3·3928	3·4018	3·4107	3·4195	3·4282	3·4368
6	3·4452	3·4536	3·4618	3·4699	3·4780	3·4859	3·4937	3·5015	3·5091	3·5167
7	3·5242	3·5316	3·5389	3·5462	3·5534	3·5605	3·5675	3·5745	3·5813	3·5882
8	3·5949	3·6016	3·6083	3·6148	3·6213	3·6278	3·6342	3·6405	3·6468	3·6531
9	3·6592	3·6654	3·6715	3·6775	3·6835	3·6894	3·6953	3·7012	3·7070	3·7127
10	3·7184	3·7241	3·7298	3·7354	3·7409	3·7464	3·7519	3·7574	3·7628	3·7681
11	3·7735	3·7788	3·7840	3·7893	3·7945	3·7996	3·8048	3·8099	3·8150	3·8200
12	3·8250	3·8300	3·8350	3·8399	3·8448	3·8497	3·8545	3·8593	3·8641	3·8689
13	3·8736	3·8783	3·8830	3·8877	3·8923	3·8969	3·9015	3·9061	3·9107	3·9152
14	3·9197	3·9242	3·9286	3·9331	3·9375	3·9419	3·9463	3·9506	3·9550	3·9593
15	3·9636	3·9678	3·9721	3·9763	3·9806	3·9848	3·9890	3·9931	3·9973	4·0014
16	4·0055	4·0096	4·0137	4·0178	4·0218	4·0259	4·0299	4·0339	4·0379	4·0419
17	4·0458	4·0498	4·0537	4·0576	4·0615	4·0654	4·0693	4·0731	4·0770	4·0808
18	4·0846	4·0884	4·0922	4·0960	4·0998	4·1035	4·1073	4·1110	4·1147	4·1184
19	4·1221	4·1258	4·1295	4·1331	4·1367	4·1404	4·1440	4·1476	4·1512	4·1548
20	4·1584	4·1619	4·1655	4·1690	4·1726	4·1761	4·1796	4·1831	4·1866	4·1901
21	4·1936	4·1970	4·2005	4·2039	4·2074	4·2108	4·2142	4·2176	4·2210	4·2244
22	4·2278	4·2312	4·2346	4·2379	4·2412	4·2446	4·2479	4·2512	4·2546	4·2579
23	4·2612	4·2644	4·2677	4·2710	4·2743	4·2775	4·2808	4·2840	4·2872	4·2905
24	4·2937	4·2969	4·3001	4·3033	4·3065	4·3097	4·3129	4·3160	4·3192	4·3224

Source: (Bliss, 1935)

(1961) however, proposed a method for obtaining a "virtual safety" dose of carcinogenic compounds by low-dose extrapolation using the probit model. In this method, every agent was considered carcinogenic. A 1/100 million response for calculating a "virtually safe" dose was suggested. It was stated that extrapolation to low-dose levels based on various dose-response data could lead to overestimation of risk, because the tumor occurrence and dose relationship in the low-dose region might be different than that in the observed region. To avoid this overestimation the use of a low slope (i.e. equal to one) value from the observed data was suggested. They suggested that a slope of one probit per common logarithm be used. The statistical assurance level was set at 99 percent. Control data to check for spontaneously occurring responses were also employed. When spontaneous rates are rather low, the "safe" dose determined would not be considerably affected. It was also suggested that responses be observed over wide ranges of stimulus and that statistical variations in large sampling sizes be considered negligible.

Mantel and Bryan et al. (1975) proposed an improved method to that investigated in 1961. An attempt was made to improve procedures to allow for spontaneous response rates, combining data from wide dose ranges, and calculating a combined "safe" dose from various data sets. The results from hypothetical experimental data sets revealed that the combined "safe" doses were considerably higher than those of the independent data sets.

Schneiderman and Mantel (1975) observed that experiments with large data sets with few responses produced a higher "safe" dose than those from similar smaller data sets. One major disadvantage of the Mantel-Bryan method is that it lacked biological credibility. A zero dose did not correspond to zero response. Brown (1984) did not propose this method for valid estimates of lowdose risk. Therefore, the dose-response curve did not have any biological support (Guess and Crump, 1976). The Mantel and Bryan dose-response function is represented by Equation 6.

$$P(d) = P(0) + (1 - P(0))\ \Phi\ (a + b \log_{10} d), \qquad (6)$$

Where:

P(0) = response due to background (Guess and Crump, 1976).
d = dose
P = dose-response function to be estimated
Φ = standard normal distribution function
a = curve fitting parameter
b = curve fitting parameter (referred to as the slope) assumed to be one
(Guess & Crump, 1976).

Although overestimation of the parameters a, b, P(0) were chosen in the high-dose range for this method, the increased risk over background approached zero at a rapidly decreasing rate in the extrapolated region (Guess and Crump, 1976). When the probit model was applied to low-dose extrapolations of vinyl chloride fed rats, a "safe" dose of approximately 500 times that of the one-hit model (described below) was produced (Guess and Crump, 1976). Guess and Crump (1978) analyzed data from animals exposed to vinyl chloride, DDT, dimethylnitrosamine, and ionizing radiation. They observed that in low-dose

extrapolations of four sets of data the extremely flat (probit-like) dose-response curves in the low-dose region fit the data worse than those linear curves (one-hit and Multi-stage) in the same region. Guess and Crump (1976) proposed that large animal experimental data could produce ". . . valid lower confidence curves on dose that decrease with decreasing dose at a faster than linear rate.". Presently the use of confidence intervals with the linear multi-stange and one-hit models is being used in place of the conservative estimates of the Mantel and Bryan parameters and slope of 1 (Hanes and Wedel, 1985). This results in the production of "safe" dose levels which could be met by industries as opposed to those practically near zero (Guess & Crump, 1976).

Log-Logistic Model. The log-logistic model is also called the growth function, autocatalytic curve, or the logit function as it was developed from chemical kinetic theory (Brown, 1984). The resulting curve is sigmoidal in shape (Berkson, 1944) and has been used to assess the potency of drugs (i.e. the L.D. 50; dose at which 50% of subjects will die) as compared to the probit model (Wilson and Worcester, 1943). Berkson (1944) stated that the term logistic was developed in 1920 by Pearl and Reed, who used the model for the description of population growth. The function itself is similar to the normal distribution, but fits the data from physicochemical phenomena better (Berkson, 1944). Equation 7 represents the logistic function.

$$P(D) = 1/[1 + \exp\text{-}(a + b \log10(D))], \quad (7)$$

where b > 0 (Brown, 1984). The logistic function has an advantage of giving a better fit with large data sets over the log-normal model (Berkson, 1944). Table 3 summarizes Berkson's comparison between the logit and the probit models. This shows on the basis of chi-square results, that either the results are the same or the logistic appears to have a slight advantage. The only result showing a distinct advantage (large difference between chi-square values) of the logistic is the Murray data. This may indicate that with a large sampling group the logistic is favorable.

Weibull Model. The assumption of this model is that the distribution of response as a function of dose follows the Weibull distribution (Hallenbeck, 1988), which previously has been utilized for the modeling of time to failure of electrical and mechanical devices (Hanes and Wedel, 1985). The model assumes a tolerance of the dose of a carcinogen for each subject (Hanes and Wedel, 1985) and is represented by equation 8 (Hallenbeck, 1988).

$$P(D) = 1 - \exp\text{-}a(D)^b \qquad (8)$$

where:

P(D) = probability of response as function of dose (D)
a = curve fitting parameter (y-intercept) - when linearized
D = dose
b = curve fitting parameter (slope) - when linearized

This model does have biological credibility because the probability of response at zero dose equals zero (Hallenbeck, 1988). In the low-dose region the curve is linear for b = 1, concave for b < 1, and convex for b > 1 (Food Safety Council, 1980).

TABLE 3

COMPARISON OF THE LOGISTIC AND THE NORMAL CURVE IN THE ESTIMATION OF DRUG POTENCY

Series	Number of dosages	Total observations	Logistic		L.D. 50		Sum of weighted squared deviation χ^2	
			a	b	Logistic	Normal	Logistic	Normal
Woodard and others (18)	5	50	5.0842	8.0459	5.54	5.62	2.77	3.01
Chen and others (2)	8	80	−15.8464	18.3321	0.0026	0.0913	7.77	8.04
Bliss, series I (1)	6	175	70.9035	39.8781	60.0	60.1	1.09	1.12
Bliss, series II (1)	6	187	74.4771	41.8241	60.4	60.1	4.88	5.03
Garwood (6)	5	200	11.0581	3.8767	0.001401	0.001421	1.75	2.89
Murray, Table I, female flies (9)	7	3,121	17.0358	7.2080	218.9	218.8	2.39	3.38
Murray, Table I, male flies (9)	11	5,495	10.3299	5.6597	66.9	66.8	5.89	11.48

Source: (Berkson, 1944)

Another general formula for the Weibull model is given by equation 9.

$$P = 1 - \exp-(\alpha + \beta x^m), \quad (9)$$

where:

x = dose
P = probability of response
m = estimated parameter
α = estimated parameter
β = estimated parameter (Carlborg, 1980).

Parameter α represents the background incidence rate, and the excess risk over background in the low-dose range can be given by βx^m (Carlborg, 1980). The Virtually Safe Dose (VSD) at a 1/1,000,000 risk over background can be calculated by equation 10 (Carlborg, 1980).

$$VSD = ((10^{-6})/\beta)^{1/m} \quad (10)$$

Weighted least squares are used to estimate the parameters by linearizing the model (Carlborg, 1980).

$$Y = -\ln(1-P) = \alpha + \beta x^m, \quad (11)$$

The weight (W) for an observed value of Y is given by equation 12.

$$W = nQ/P, \quad (12)$$

where Q = 1 - P and n represents the number of subjects at risk. A trial-and-error would need to be performed on the parameter m (Carlborg, 1980) if a linear weighted least-squares program was used. The Weibull model parameters given in equation 8 can be estimated by linear regression of data sets of three points or greater (Hallenbeck, 1988).

Carlborg (1980) calculated VSDs for varying values of m. These are presented in Table 4 and Figure 12. Figure 12 plots the dose versus tumor rate for varying values of the parameter m. Carlborg (1980) proposed that it is the parameter m that determines the VSD in the low-dose range. Table 4 and Figure 12 both show that the parameter m determines the VSD at low concentrations. The sharp increase in VSD resulting from an increase in the parameter m supports this observation.

The tolerance distribution models are based on the assumption that when the response is quantal the frequency will depend on the concentration of the toxicant. The tolerance varies among subjects within the population due to the biological variability. Therefore, it is convenient to consider frequency of distribution of tolerances throughout the population (Brown, 1984).

Mechanistic Models

Mechanistic models are based on the assumption that for a normal cell to become cancerous a certain number of "hits" by a toxicant is required. These models, unlike the tolerance distribution models, do not have biological credibility (this means that at zero dose a zero response is obtained).

Brown (1984) states, "A number of dose-response models have been suggested on the basis of assumptions regarding the mechanism of action of the toxic agent upon its target site. The "hit" or mechanism of action is the basis of the mechanistic theory. The "hit theory" rests upon the suggestion that a response is produced by the cell after being exposed to certain number of hits by

TABLE 4

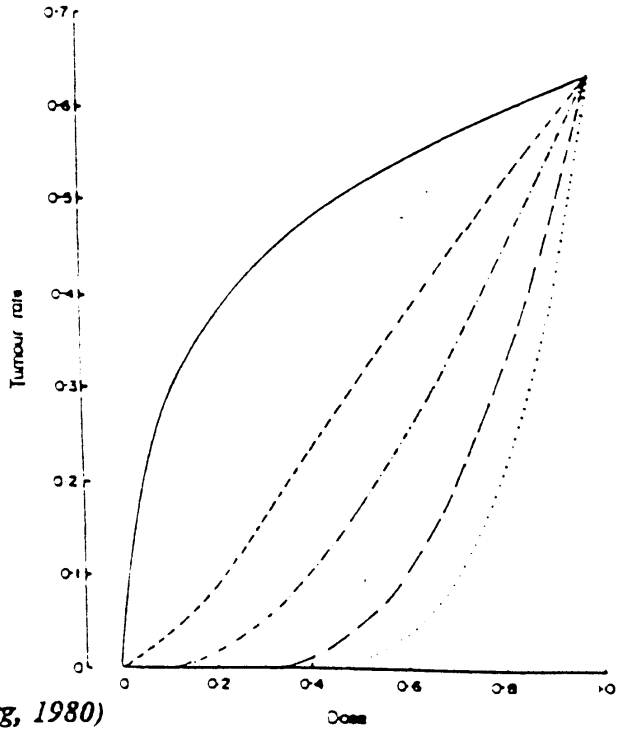
THE RELATIONSHIP BETWEEN THE VIRTUALLY SAFE DOSE (VSD): (REPRESENTING A RISK OF 1:10⁶) AND THE VALUE OF THE PARAMETER M IN THE WEIBULL MODEL

Weibull model parameter m	Virtually safe dose
0·5	$3{\cdot}2 \times 10^{-10} \times TD_{0{\cdot}63}$
1·5	$1{\cdot}0 \times 10^{-4} \times TD_{0{\cdot}63}$
2·5	$4{\cdot}0 \times 10^{-3} \times TD_{0{\cdot}63}$
3·5	$1{\cdot}9 \times 10^{-2} \times TD_{0{\cdot}63}$
4·5	$4{\cdot}6 \times 10^{-2} \times TD_{0{\cdot}63}$
5·5	$8{\cdot}1 \times 10^{-2} \times TD_{0{\cdot}63}$
6·5	$1{\cdot}2 \times 10^{-1} \times TD_{0{\cdot}63}$

$TD_{0{\cdot}63}$ = Dose that produces a tumour rate of 0·63.

FIGURE 12

GRAPHS OF THE WEIBULL FOR VARIOUS VALUES OF THE PARAMETER M



Source: (Carlborg, 1980)

Graphs of the Weibull model for various values of the parameter m: 0·5 (——), 1·5 (------), 2·5 (– – – –), 4·5 (— — —), 6·5 (· · · · · ·)

the toxic substance or by a certain number of stages of change (Hallenbeck, 1988). Four postulates upon which the "hit theory" is based are (Brown, 1984):

> "(1) the organism has some number M of "critical targets" (usually assumed to be infinitely large);
>
> "(2) the organism responds if m or more of these critical targets are "destroyed";
>
> "(3) a critical target is destroyed if it is "hit" by k or more toxic particles;
>
> "(4) the probability of a hit in the low dose region is proportional to the dose level of the toxic agent, i.e., Prob(hit) = λd, $\lambda > 0$."

<u>One-hit Model.</u> Iverson and Arley produced one of the first quantitative theories of carcinogenesis, which became known as the "one-hit" model (Brown, 1976). Equation 13 represents the one-hit model (Food Safety Council, 1980).

$$P(D) = 1 - \exp(-\lambda D), \quad (13)$$

where:

P(D) = probability of response as a function of dose (D)
D = dose
λ = curve fitting parameter (slope) - when linearized

and $\lambda > 0$. λD represents the number of effective hits of an offending chemical and is taken to follow a Poisson distribution as a function of dose (Rai and Van Ryzin, 1979). This means that the model assumes a toxic effect occurs after a single effective hit is received (Rai and Van Ryzin, 1979). The parameter λ, is considered the slope of the curve at the origin (Refer to equation 14) (Rai and Van Ryzin, 1979).

$$\lim_{d \to 0} \frac{P(d)}{d} = \lambda \qquad (14)$$

This shows that as the limit of dose-response function approaches a dose of zero the parameter λ equals the slope of the dose-response curve. The EPA currently uses the one-hit model for risk evaluation and disregards the concept of threshold (Wardlaw, 1985). "Radiation experience has been cited as the best evidence for the one-hit model, even though the action of genotoxic carcinogens differs from that of radiation. The pattern of responses seen in the induction of genetic mutation, which are likely involved in the cancer process, also suggests that the one-hit model may be valid . . . This model is the most conservative in terms of setting the VSD" (Wardlaw, 1985).

The Food and Drug Administration (FDA) also uses the one-hit model for risk evaluation of toxic chemicals (Maxim and Harrington, 1984). The FDA used the one-hit model with a 99 percent confidence interval as a safety factor for calculation of the VSD for polychlorinated biphenyls (PCB) concentrations in fish (Maxim and Harrington, 1984).

In a report by a subcommittee on estimation of risks of irreversible, delayed toxicity to the Department of Health, Education, and Welfare Committee (DHEW), the one-hit model was recommended for low-dose extrapolation of incidence data (Hoel, et al., 1975). The one-hit model was also recommended for risk assessment by the BEIR Report (Hoel, et al. 1975). They found that the "... use of the linear extrapolation from data obtained at high doses and dose rates may be justified on pragmatic grounds as a basis for risk estimation" (Hoel, et

al., 1975). The one-hit and multistage models become approximately linear at low dose levels. This linearity is important that these are conservative models (Brown, 1984). this means that the linearity in the low dose region produces a higher risk at a particular dose than does the convex curve.

Multi-Hit Model. Rai and Van Ryzin (1979) proposed a generalized multi-hit model based on a stochastic biological basis. This model assumes that a response (i.e., cancer) will be induced by series of "k" hits over a fixed period of time. Equation 15 represents the probability estimate of the toxic response occurring given a multi-hit mechanistic assumption of cancer initiation and propagation (Rai and Van Ryzin, 1979). This equation represents the probability of a response occurring if the number of fixed hits over a period of time follows a Poisson distribution with expectation θd for dose d (Rai and Van Ryzin, 1979).

$$P(d) = P\{X >= k\} = \sum \{(\theta d)^i \exp^{(-\theta d)}/i!\} =$$

$$\int (t^{k-1}\exp^{(-t)}/(k-1)!)dt, \qquad (15)$$

where:

P(d) = probability of response
k = number of hits
t = time
θd = expectation of number of hits
d = dose
$0 \leq t \leq \theta d$ (Rai and Van Ryzin, 1979).

The last equality is a result of "... repeated integration by parts." (Rai and Van Ryzin, 1979). Equation 15 is rewritten to form equation 16 (Rai and Van Ryzin, 1979).

$$P(d) = P(d;k,\theta) = \int g(t)\partial t, \qquad (16)$$

where $0 \leq t \leq d$. Equation 17 represents the function g(t) (Rai and Van Ryzin, 1979).

$$g(t) = \theta^k t^{k-1} \exp^{(-\theta t)} / \Gamma(k),\ 0 < t < \infty, \qquad (17)$$

where:

θ = scale parameter
k = shape parameter
t = time

The gamma function $\Gamma(k)$ is represented by equation 18 (Rai and Van Ryzin, 1978).

$$\Gamma(u) = \int t^{u-1} e^{-t} \partial t \qquad (18)$$

where $0 \leq t \leq \infty$ and $u > 0$. This produces a statistical interpretation of the model. The scale parameter represented by θ^{-1} and the shape parameter by k ($k > 0$) are used to fit data to the dose-response model. Equation 19 represents the response in the low dose region (Rai and Van Ryzin, 1979).

$$\lim (d \to 0) \{P(d)/d^k\} = \theta^k / \Gamma(k+1) = c > 0 \qquad (19)$$

where:

k = curve fitting parameter (referred to as "number of hits")
P(d) = probability of response
θ = scale parameter
Γ = gamma function

At low dose (near zero), response is represented by equation 20 (Rye and Van Ryzin, 1979).

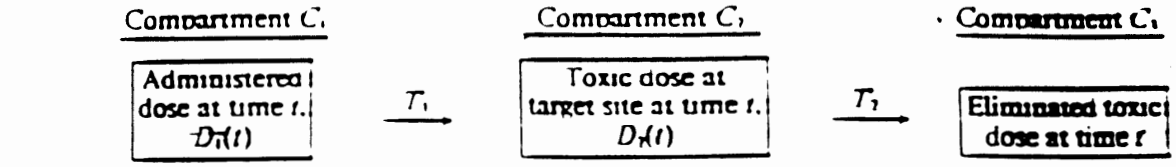
$$P(d) \doteq cd^k \qquad (20)$$

At k = 1 the model becomes the one-hit model. At $k < 1$ the curve is concave (gives higher estimate of risk) and at $k > 1$ convex (gives lower estimate of risk) (Rai and Van Ryzin, 1979). This means that the risk estimate is dependant on the number of hits required for carcinogenesis. At low doses the logit model and multi-hit model are similar, and at high doses the multi-hit model is similar to the probit model (Food Safety Council, 1980). The Weibull model, a tolerance distribution model, also has similar extrapolation characteristics to the multi-hit model in which the tolerance distribution is gamma (Brown, 1984).

Rai and Van Ryzin (1987) also proposed a multi-hit dose response model that incorporated non-linear kinetics. The incidence of spontaneous background response, when incorporated into the model, produced four parameters. They used maximum likelihood estimation to estimate these four parameters. They investigated three animal carcinogenicity bioassays that produce, respectively, concave, linear, and convex dose-response curves in the observed region (Rai and Van Ryzin, 1987). Figure 13 reveals the model from a compartmental point of view. This figure shows that a dose ($D_1(t)$) administered at time t in compartment one is transformed by an outgoing process, T_1, to an internal toxic dose ($D_2(t)$) at the target organ in the second compartment. Next this toxic dose is converted into a nontoxic dose by another outgoing process, T_2, into compartment three. Rai and Van Ryzin stated that the transformation process for any single compartment "... is said to follow dose-dependent Michaelis Menten nonlinear kinetics if

FIGURE 13

DOSE-RESPONSE MODEL INCORPORATING NONLINEAR KINETICS



T_1 = first process
(outgoing from compartment C_1 = incoming to target site)

T_2 = second process
(outgoing from compartment C_2 = outgoing from target site)

Compartmental model.

Source: (Rye and Van Ryzin, 1987)

$$D'(t) = - [(bD(t))/(c+D(t))], \; b>0, \; c>0 \qquad (21a)$$

where D(t) is the dose concentration at time t in the single compartment and D'(t) is the first derivative of D(t) with respect to t. The constant b is the maximum rate of change and c is the Michaelis-Menten constant, i.e., the dose concentration in the compartment at which the rate of change is 1/2(b)" (Rai and Van Ryzin, 1987). The dose response model proposed by Rai and Van Ryzin (1987) based on nonlinear kinetics is represented by equation 21b.

$$f(D) = 1 - \exp-(\alpha + \lambda D^{B}) \qquad (21b)$$

where:

f(D) = probability of response
α = curve fitting parameter
λ = curve fitting parameter
β = curve fitting parameter

Multi-stage (Armitage-Doll) Model. The processes involved in carcinogenesis are transformation and growth. One or more changes in a normal cell that enable it to form a tumor is called transformation. When the cell duplicates into multiple cells and produces a family of cells called clones, it is termed growth (Whittemore and Keller, 1978). The onset of carcinogenesis is caused by carcinogens (i.e. chemicals or viruses) (Whittemore and Keller, 1978). As stated in Whittemore and Keller (1978) Iverrsen and Arley (1952) proposed the earliest quantitative theory of carcinogenesis, which suggested that the normal cells were transformed to cancer cells in one stage. Equations 22 and 23 describe this theory.

$$\partial p_0/\partial t = -\lambda(t)p_0(t), \qquad p_0(0) = 1 \qquad (22)$$

$$\partial p_1/\partial t = \lambda(t)p_0(t), \qquad p_1(0) = 0 \tag{23}$$

where:

$p_0(t)$ = probability that a cell is normal at time t
$p_1(t)$ = the probability that the cell is transformed at time t
$\lambda(t)$ = the transition probability rate

The multi-stage theory proposed by Muller and Nordling suggests that a cell can produce a tumor only after passing through k number of mutations (Whittemore and Keller, 1978). Figure 14 is a schematic of the k-stage theory of transformation where cells start as normal cells at stage zero and are transformed at the kth stage (Whittemore and Keller, 1978). A cell has the probability q of having one mutation in any year, but it cannot have more than one. (Whittemore and Keller, 1978). Equation 24 represents this assumption and resembles a binomial distribution (Whittemore and Keller, 1978).

$$q^{k-1}(1-q)_{a-k}(a-1)\ldots(a-k+1)/(k-1)! \tag{24}$$

where:

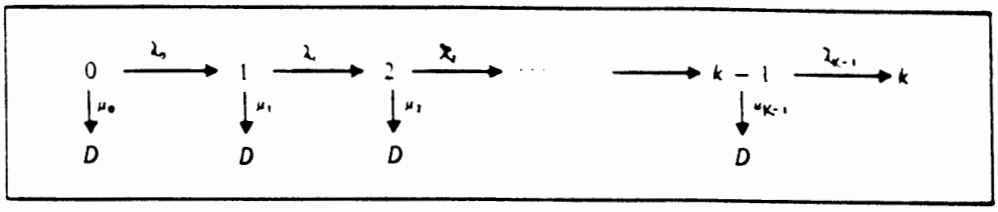
q = probability of having one mutation in any year (carcinogen studies are performed over a 70 year period)
k = number of mutations
a = year of mutation

This equation produces the transformation probability rate per cell in the a_{th} year.

As stated in Whiittemore and Keller (1978) Armitage and Doll (1954) in light of the work of Muller and Nordling (1953) proposed that ". . . k changes have different transition rates $\lambda_i(t)$, i = 0, ... ,k-1, ... and they must occur in the

FIGURE 14

SCHEMATIC REPRESENTATION OF THE K-STAGE THEORY OF TRANSFORMATION



Source: (Whittemore and Keller, 1976) [42]

order 0,...,k-1." Armitage and Doll further assumed that the effect of the agent at some of the stages was additive to effects induced by external stimuli at those stages. This caused a lower power than k (stages) for D (dose)." (Food Safety Council, 1980). Equation 25 presents the formula for this latest set of assumptions from Crump, Hoel, Langley and Peto (1976). This equation assumes additivity at all stages (Food Safety Council, 1980),

$$P(D) = 1 - \exp\{-\sum \alpha_i D^i\}, \ \alpha_i \geq 0 \qquad (25)$$

where:

P(D) = probability of response
α = estimated parameter
D = Dose
i = number of stages
$0 \leq i \leq \infty$ and α_i is nonnegative.

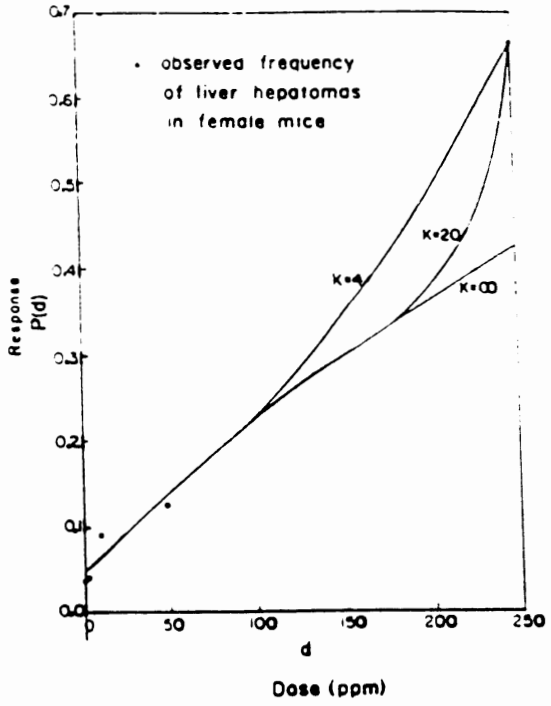
Guess and Crump (1976) proposed a method of estimating the parameters in the Armitage and Doll Model by maximum likelihood estimation. They found that the lower order coefficients of the k = 4 curve were similar to those of the k = ∞ and k = 9 (Guess and Crump, 1976). Figure 15 illustrates these findings, where the observed frequencies of the various kth stages in the extrapolated region match almost exactly. Dose is parts per million of DDT fed to femal and male mice and resonse is the percentage of mice exhibiting tumors. The observed frequencies of the various kth stages vary only in the high dose region.

Correction For Backkground Response

Two methods are commonly employed to correct for background response at zero dose. That is, there is a base level of disease incidence associated with

FIGURE 15

COMPARISON OF OBSERVED FREQUENCY OF RESPONSE VARYING K STAGES



Source: (Guess and Crump, 1976)

any population. Toxicity testing focuses upon incremental increases in cancer or other effects. These background-response-correction methods offer alternatives to make null these non-specific effects. The first method, termed "Abbott's correction," assumes an independent action between the stimulus and the background (Brown, 1984). Equation 26 represents this assumption,

$$P(D) = P_0 + (1-P_0)P^*(D), \quad (26)$$

where:

$P^*(D)$ = dose-induced probability of response
$P(D)$ = probability of response
$P_0(D)$ = probability of response due to background
(Brown, 1984).

This equation corrects the probability of response based on the independent background assumption. The second method proposes that the stimulus/background relationship is always additive and the overall probability of response will be a linear combination of the experimental and background chemical doses. This is illustrated by equation 27,

$$P(D) = P^*(D+D_0) \quad (27)$$

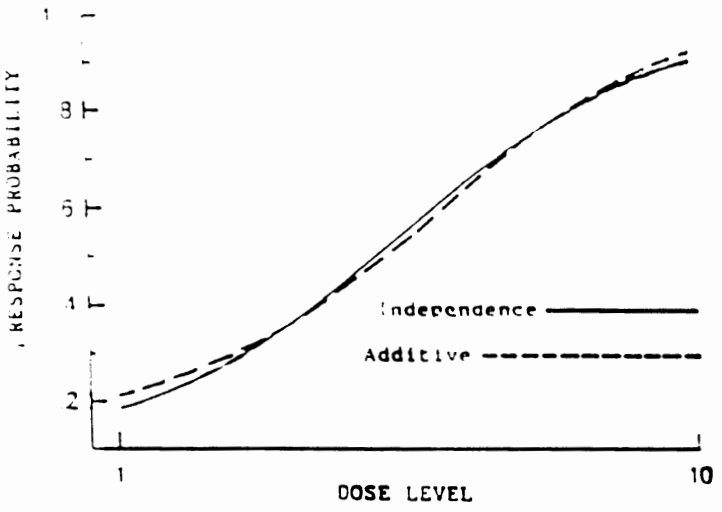
where:

D^0 = some unknown background dose

Brown reports that "... both assumptions lead to identical mathematical models for overall response rates when the assumed dose-induced model is either the single-hit or multistage" (Brown, 1984). Figure 16 graphically shows the difference between the additive and independent assumption of background response using the log-logistic model by plotting dose versus response probability

FIGURE 16

COMPARISON OF LOG-LOGISTIC DOSE-RESPONSE MODELS ASSUMING INDEPENDENT AND ADDITIVE BACKGROUND



Source: (Brown, 1984)

(Brown, 1984). This figure shows both correction assumptions describe data equally well. Table 5 presents data showing the vast difference between the two assumptions in the low-dose region using the log-normal model at various doses (Brown, 1984). Hoel found that "... low dose linearity prevails except when the background mechanism is totally independent of the dose-induced mechanism." (Brown, 1984).

Each of these correction methods will introduce specific biases. In order to standardize this approach a decision was made to remove this background before the dose-response data was modeled.

Summary

As shown in Figure 1 and discussed in the preceding information, these five dose-response models can generate vastly different results. Professional scientists and engineers require techniques which allow comparisons between alternative formulations whenever environmentally critical decisions are to be made. Similarly, students can benefit from techniques which allow quantification and subsequent comparisons among often arcane theoretical material. Q-Risk is an attempt to aid all of these audiences with these problems.

TABLE 5

EXCESS RISK [P(D)-P(O)] FOR LOG-NORMAL DOSE-RESPONSE MODEL ASSUMING INDEPENDENT AND ADDITIVE BACKGROUND

Dose (D)	Type of Background	
	Independent	Additive
10^{0}	4.0×10^{-1}	4.0×10^{-1}
10^{-1}	1.5×10^{-2}	5.2×10^{-2}
10^{-2}	1.6×10^{-5}	5.2×10^{-3}
10^{-3}	3.8×10^{-10}	5.1×10^{-4}
10^{-4}	1.8×10^{-16}	5.1×10^{-5}

*P(O) = 0.1; log normal model slope = 2 from (34)

Source: (Brown, 1984)

CHAPTER III

MATERIALS AND METHODS
Q-Risk Computer Program

Development of Program

This program was developed to aid the scientist in extending dose response data. The main focus of the program was to provide a "user-friendly" computer code to aid in risk calculations and serve as a tutorial in some of the areas of environmental decision making. The program incorporated the Probit, Weibull, Log-logistic, One-hit, and Multi-stage dose-response models and was structured to allow users with minimal computer knowledge to spend more time completing data analysis than learning how to operate a computer or specific, more complex codes.

To accomplish this, Microsoft® QuickBasic, version 4.5, was used to generate the basic code. This complex language was chosen due to its graphic capabilities and abilities to do reiterative calculations. Version 4.5 is also equipped with its own compiler so that "stand-alone" executable files could be made. These "stand-alone" executable files allow the user to be able to run the code under Microsoft® or related DOS without having to use interpretive QuickBasic, allowing easier, quicker and more universal operation on practically any IBM®-compatible home computer.

The program was designed with tutorial screens explaining the various dose-response models, calculations, and graphing. Menus are generated to allow easy access to any part of the program. Data files and output files (containing estimated parameter values and model information) are written during execution for future use. Graphs of the original data and extrapolated data are generated for the user to do dose-response determination. Axis ranges of the graphs can be chosen by the user to allow some flexibility in plotting. Parameters for each model are calculated instantly by simple, directed keystrokes. The program will analyze a minimum of 3 and a maximum of 30 dose-response data points.

Code For Model Parameter Calculations

Method Of Estimation. The method chosen to estimate the curve fitting parameters for the tolerance distribution models was linear regression. This was chosen over maximum likelihood estimation because of the ease of calculation, coding of the program sequence, and processing time. The following equations, 28 through 30, were used for the linear regression calculations.

$$A \text{ (y-intercept)} = \Sigma y - B*\Sigma x/n \tag{28}$$

$$B \text{ (slope)} = n*\Sigma xy - \Sigma x*\Sigma y/(n*\Sigma x^2 - (\Sigma x)^2) \tag{29}$$

$$r \text{ (correlation coeff)} =$$

$$n*(\Sigma xy - \Sigma x\Sigma y)/(\sqrt{[n*\Sigma x^2 - (\Sigma x)^2][n*\Sigma y^2 - (\Sigma y)^2]}), \tag{30}$$

where:

n = number of data points.

y = represents response data point

x = represents dose data point

These equations were used to calculate the curve-fitting parameters for the tolerance distribution models by linear regression using the equation of the line (y = mX + b). The A and B parameters would correspond to the curve fitting parameters in the linearized tolerance distribution equations as described below.

The One-Hit and Multi-stage model parameters were estimated using Gauss-Jordan elimination instead of the alternative method maximum likelihood estimation, because of the the ease of calculation, coding of the program sequence, and processing time. Although linear regression could have been used on the One-Hit model, Gauss-Jordan elimination was used because of the need to normalize the parameter estimation for the mechanistic models. That is, to apply the same method of parameter estimation for each of the mechanistic models. For a detailed description of the Gauss-Jordan elimination method see Appendix A (Equations 38-43).

Probit Model. The probit model parameters were estimated using the equation found in Hallenbeck (1988) (Equation 32). This equation is in the linear form and is derived from Equation 31 after log transformation and linearization as discussed by Hallenbeck (1988).

$$P_e = 1/(2\pi)^{0.5} \int \exp\{-z^2/2\}dz \qquad (31)$$

$$z = b \log_{10} D + a \qquad (32)$$

where:

z = standard normal variate
a = $-\mu/\sigma$ (μ = population mean of $\log_{10}$ D

σ = population standard deviation of $\log_{10}$ D)
b = $1/\sigma$.

The standard normal variate was calculated using a probit data file similar to that previously presented in Table 2 (Finney, 1971) that related probits to their corresponding percent response. The standard normal variate was calculated using equation 33 (Food and Safety Council, 1980).

$$z = \text{Probit} - 5 \quad (33)$$

Once the probability of response was converted and the common logarithm of dose calculated, the curve parameters a and b and the regression coefficient were calculated.

Log-logistic Model. Equation 34 was transformed into a linear form (equation 35), and linear regression was used for calculating the curve parameters (Hallenbeck, 1988). See Appendix A for linear transformation.

$$P_e = 1/(1 + e^{-(a + b*\log D)}) \quad (34)$$

$$-\ln [(1-P_e)/P_e] = a + b*\log D \quad (35)$$

where:

P_e = probability of response
a = curve fitting parameter (y-intercept)
b = curve fitting parameter (slope)
D = dose

Weibull Model. The original model equation (equation 36) was transformed into a linear form (equation 37), and linear regression was used for calculating the curve parameters (Hallenbeck, 1988). See Appendix A for linear transformation.

$$P_e = 1 - \exp\text{-}(aD^b) \quad (36)$$

$$\ln[-\ln(1-P_e)] = \ln a + b*\ln D \quad (37)$$

One-Hit Model. The one-hit model is represented by equation 44 (Hallenbeck, 1988). Parameter λ was calculated using Gauss-Jordian elimination to solve a least-squares polynomial fit of n data pairs. The polynomial is set to the first degree (p = 1), thereby assuming that "cancer" was produced in one stage.

$$P(D) = 1 - e^{-(\lambda D)} \quad (44)$$

The linear equation (equation 45) used for estimation of the parameter, λ_1 was a transformation of equation 44. Y in equation 38 is represented by -ln (1-P_e), and only the first two coefficients are determined as the degree of the polynomial was set equal to one in conjunction with the one-hit assumptions.

$$-\ln (1-P_e) = \lambda_1 D + \lambda_0 \quad (45)$$

Multi-Hit Model. The multi-hit model was not included in the Q-Risk program because of its similarity of extrapolation characteristics to the Weibull model (Brown, 1984).

Multi-Stage Model. The multi-stage model parameters were estimated by the Gauss-Jordan code sequence. The user was given a choice of choosing up to a fifth degree polynomial. The limit was based on two reasons: (1) Guess and Crump (1976) found that the low-order coefficients of a polynomial curve of degree 4 (K = 4) were the same as those for a polynomial curve of degree ∞ up to 9 significant figures, and (2) Whittemore and Keller (1978) stated that there

"... there is a lack of any direct experimental evidence that cancer occurs in more than two stages." The linear equation (equation 47) used for estimation of the polynomial coefficients was transformed from the original model equation (equation 46).

$$P(D) = 1 - \exp-(\Sigma \propto_i D^i) \quad (46)$$

$$-\ln (1-P_e) = \propto_0 + \propto_1 D + \propto_2 D^2 + \propto_3 D^3 + \ldots + \propto_k D^k, \quad (47)$$

where $\propto_i \geq 0$ and $0 \leq i \leq k$ (Food and Safety Council, 1980). See Appendix A for the linear transformation.

Coding Of Q-RISK'S Programs

Q-Risk was divided into ten separate programs. The entire code was divided into ten separate programs. This was done to facilitate compiling. The ten separate programs together occupy approximately 500,000 bytes of memory. Each compiled program was accessed from a central code responsible for displaying the user menus and graphing the results. The ten basic programs are listed in Table 6, which lists the function/description of each of these executable files. Figures 17A and 17B, Q-Risk flow charts, present program flow charts for the total code. The QRA.EXE program is the main program from which all subroutines and subprograms are called upon response from the user. QRA.EXE is executed by typing QRA at the disk drive prompt. The user has the option of exiting to DOS throughout the program. Information screens describing what the program does and models included are presented after the subroutines are called.

TABLE 6

FILES INCORPORATED INTO Q-RISK

FILE NAME	FUNCTION/DESCRIPTION
QRA.EXE	Main Program. Contains Main Menu and all subsequent menus, model descriptions, model limitations, model data requirements and graphing routine.
Subroutine OSU	Draws the letters "OSU." Called from QRA.EXE. Draws initial screens
INTRO.EXE	Lists models and brief introduction. Called from QRA.EXE.
Subroutine SCPDR	Sets the screen coordinates and resolution for graphing. Called from QRA.EXE.
ENTRY.EXE	Called from QRA.EXE. User is allowed to either input new data or use an existing data file. Creates DRI.DAT. Produces output file containing input data.
PROB.EXE	Called from QRA.EXE. Uses linear regression to estimate curve fitting parameters for Probit Model. Calls a PROB.DAT file for transformation of percent response. Produces output file PROBIT.OUT and data file PROBIT.DAT.
WEIL.EXE	Called from QRA.EXE. Uses linear regression to calculate curve fitting parameters for the Weibull model. Produces output file WEIL.OUT and data file WEIL.DAT.
LOGCAL.EXE	Called from QRA.EXE. Uses linear regression to calculate curve fitting parameters for the Log-logistic model. Produces output file LOGLOG.OUT and data file LOGCAL.DAT.
ONEHIT.EXE	Called from QRA.EXE. Uses Gauss-Jordan elimination sequence to calculate curve fitting parameter for One-Hit model. Produces output file ONEHIT.OUT and data file ONEHIT.DAT.
MULTSTG.EXE	Called from QRA.EXE. Uses Gauss-Jordan elimination sequence to calculate coefficients of the kth degree polynomial for the Multi-Stage Model. Produces output file MULTSTG.OUT and data file MULSTG.DAT.
Subroutine WELSR	Draws screen that displays author and program version.
Subroutine WELC3	Draws second welcome screen.
DATALIM.EXE	Program for displaying model information.
DATAREQ.EXE	Program for displaying model requirements for program.

Fig. 17A. - Q-RISK FLOWCHART PART 1

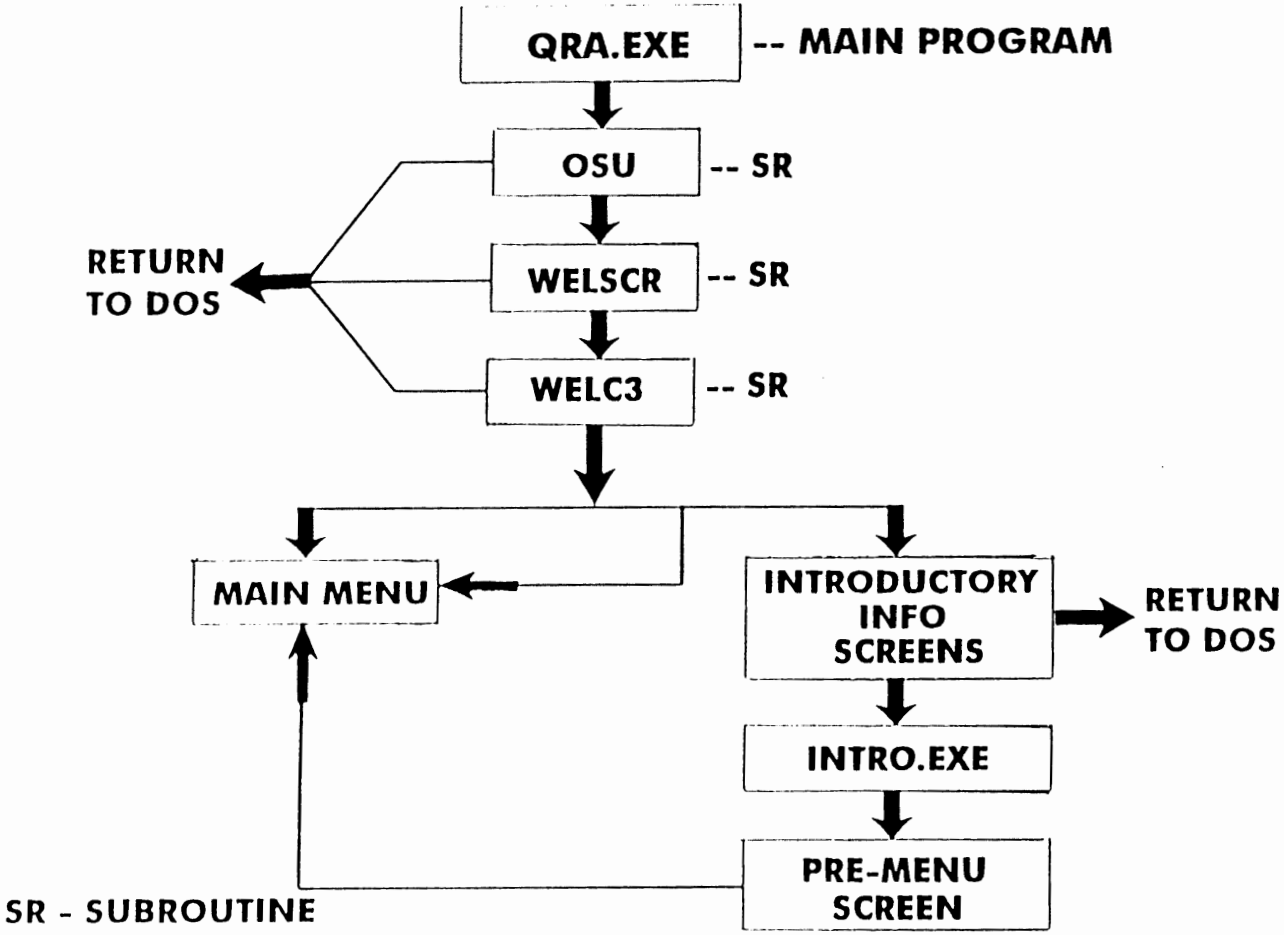
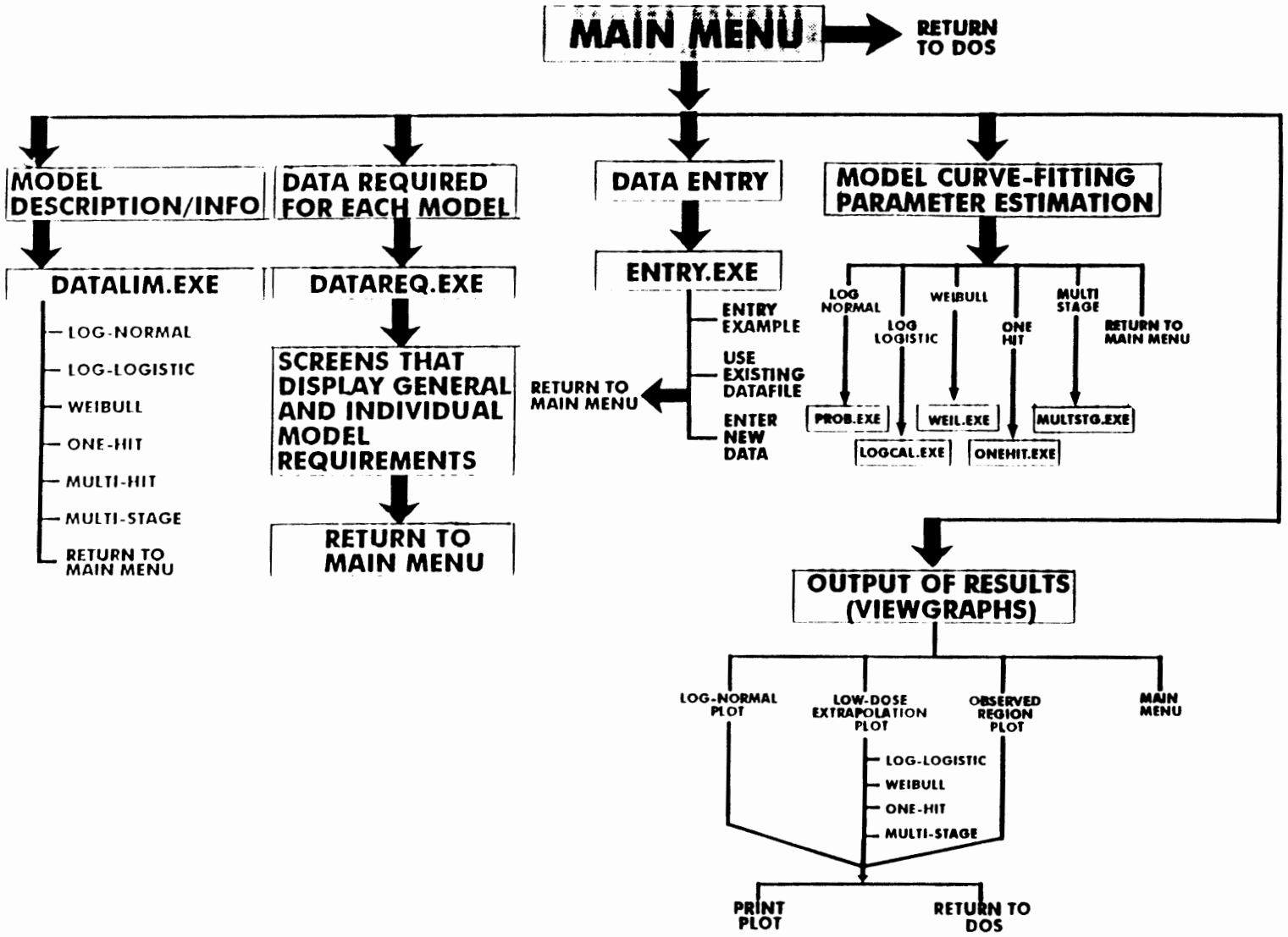


Fig. 17B. - Q-RISK FLOWCHART PART 2



Next, the main menu is displayed giving the user a list of functions to perform by pressing a function key.

The following are the functions available to the user:

- Model Information/Description
- Data Required For Each Model
- Data Entry
- Parameter Estimation
- Output Of Results

The "Model Information/Description" function provides a description of each model's assumption and general information. The "Data Required For Each Model" function explains what parameters are required for each model. The "Parameter Estimation" function estimates the parameters for each model after selecting the desired model. The "Output Of Results" function produces graphs of the original and extrapolated data after parameter estimation. Before the user can perform parameter estimations data must be entered or a data file selected by the user. The user must also perform the parameter estimation before selecting the "Output Of Results" option. After performing each option the user is given the choice of returning to the main menu or exiting to DOS. A Shift + Printscrn option is given to the user to allow them to print the plot.

QRA.EXE Program. Three subroutines are incorporated into the main program (see Table 6). This main program also includes routines for the "Main Menu" (from which all functions of the program are called), information screens, help screens, and graphing sequences.

Introductory Screens. Upon entry to QRA.EXE the user can go to the main menu or choose to view the introductory information screens that tell about Q-Risk. The program INTRO.EXE is executed upon selection of the latter option which displays these information screens.

Model Info/Data Requirements. The subprogram DATALIM.EXE is executed upon selection of the model description/information option from the main menu. This code sequence gives the user a list of models to choose from for information. The subprogram DATAREQ.EXE presents the user with screens explaining the data required for each model.

ENTRY.EXE Program. This option presented a help screen named "Data Limitations." This screen explains the limitations of the model in terms of the degree of the polynomial for the Gauss-Jordan elimination (see Appendix A for explanation of this method). The user could either input new data or use a previously created data file. The user was allowed to enter up to 30 dose-response data points. Since the models used in the program do not compensate for background response, the user was not allowed to input a response greater than zero for a corresponding dose of zero. Once the data are entered a screen was created to review and correct, if necessary, the input data. A data file is created once the user inputs the name they wish to call the file, called "NAME.DAT". This *.DAT file contains the original data points, number of data points, and the dose-response units of measurement and is named by the user. This file can be called for future use. Recall that this program must have

been executed by the user before any parameter estimation or plotting could be performed.

PROB.EXE Program. By selecting the parameter estimation option from the main menu and then the Probit Model option, PROB.EXE is executed. This program opens the previously created data file PROB.DAT, which contains the corresponding Probit values for the percent responses. The percent response was transformed into a standard normal variate, and then the curve fitting parameters were estimated using linear regression. Once the parameters are estimated, the viewer is given a screen displaying the parameters, the model equation, and the correlation coefficient. This saves to two output files named "PROBIT.OUT"and "PROBIT.DAT" containing the identical information as the screen as well as the estimated parameters. These files consisting of the input, output and parameter files can be subsequently manipulated by DOS editors or appropriate word processors.

LOGCAL.EXE Program. This program also called from the main program's menu through selection of the parameter estimation option generates the Log-Logistic Model option. The curve fitting parameters for the Log-Logistic Model are calculated using linear regression. Two output files are created called "LOGCAL.DAT" and "LOGLOG.OUT." The latter file contains the estimated parameters, the equation of the model, the estimated parameters, and the correlation coefficient. The *.DAT file contains only the estimated parameters.

WEIL.EXE Program. The WEIBULL Model option is called from the main program. The curve fitting parameters for the Weibull model were calculated using linear regression. A data file called "WEIL.DAT" is created which contains the estimated parameters, and an output file called "WEIL.OUT" containing the equation of the model, the estimated parameters, and the correlation coefficient is produced.

ONEHIT.EXE Program. The coefficients of the first degree polynomial for the One-Hit model were calculated using the Gauss-Jordan elimination sequence. The estimated coefficients for the first degree polynomial and the chi-square value for the model were written to a data file called "ONEHIT.DAT". The chi-square value gives a quantitative description of how well the model fits the data (i.e. the higher the value the better the model fits the data). The estimated coefficients for the first degree polynomial, the chi-square value, and the equation of the model were written to an output file called "ONEHIT.OUT".

MULTSTG.EXE Program. Prior to parameter estimation the viewer was given a help screen explaining the polynomial equation used and the selection of the kth (kth refers to the stage of cancer) degree of the polynomial. These parameter values, the chi-square value, and the model equation were written to an output file called "MULSTG.OUT". The parameters were written to a data file called "MULSTG.DAT".

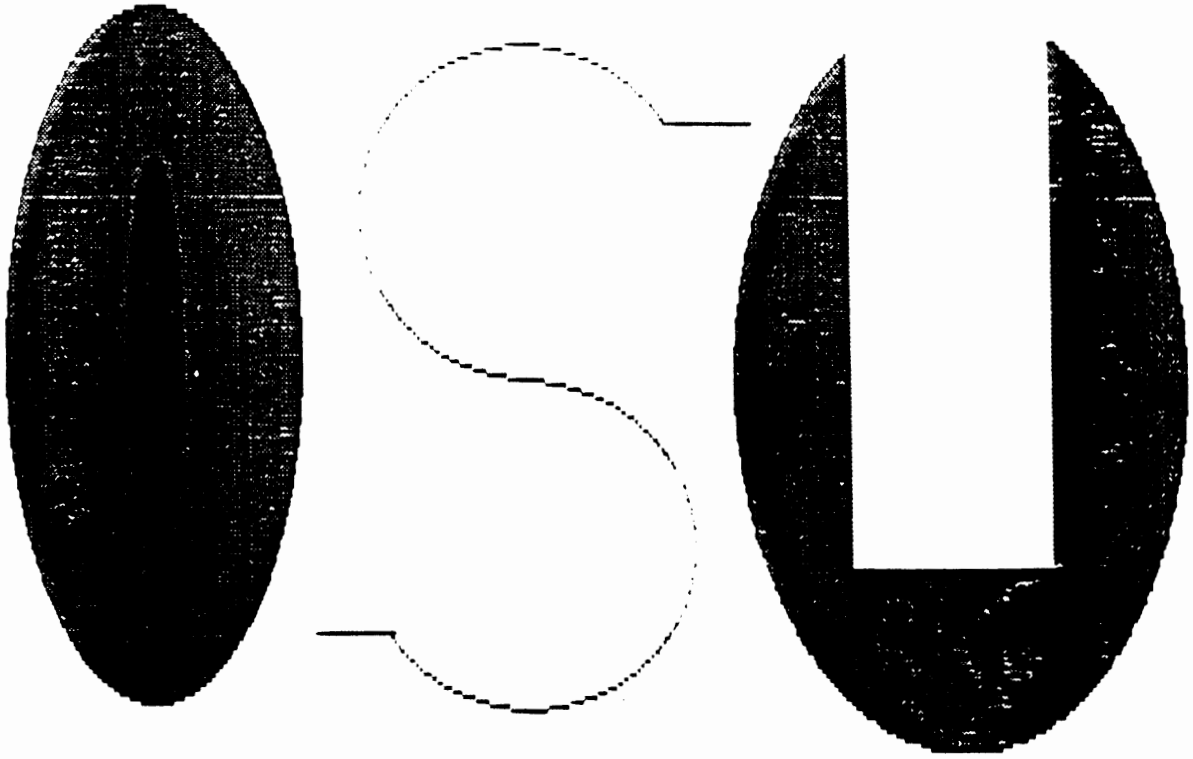
CHAPTER IV

RESULTS AND DISCUSSION
Q-RISK BASIC CODE SEQUENCE

Q-RISK Input AND Output

Figure 18 is the first screen that appears once Q-Risk is started. This screen is an emblem for the Oklahoma State University (OSU). Figures 19 through 30 represent welcome and information screens throughout the Q-Risk program. Figures 31 through 35 are examples of the selection menus found in the program. Figure 31 is the main menu from which all other functions are accessed. The model was programmed to allow the user to use a data file previously generated or to input original data. The user was allowed to name the file also with a *.dat file extension for later manipulation or review. Figures 36 through 40 are the screens which display the parameter estimate results for each of the models. These screens are produced by selecting the "Parameter Estimation" function and subsequently the function key for the corresponding model whose parameters are to be estimated. Figures 36, 37, 38, 39, and 40 are screens which lists the results of the linear regression for each model and the equation for that model. The parameter A is the y-intercept and B is the slope. The "Log Dose (#)" values represent X values and the "Transformed Response"

FIGURE 18
OSU EMBLEM SCREEN



(Press any key to continue or <ESC> to QUIT)

FIGURE 19
FIRST WELCOME SCREEN

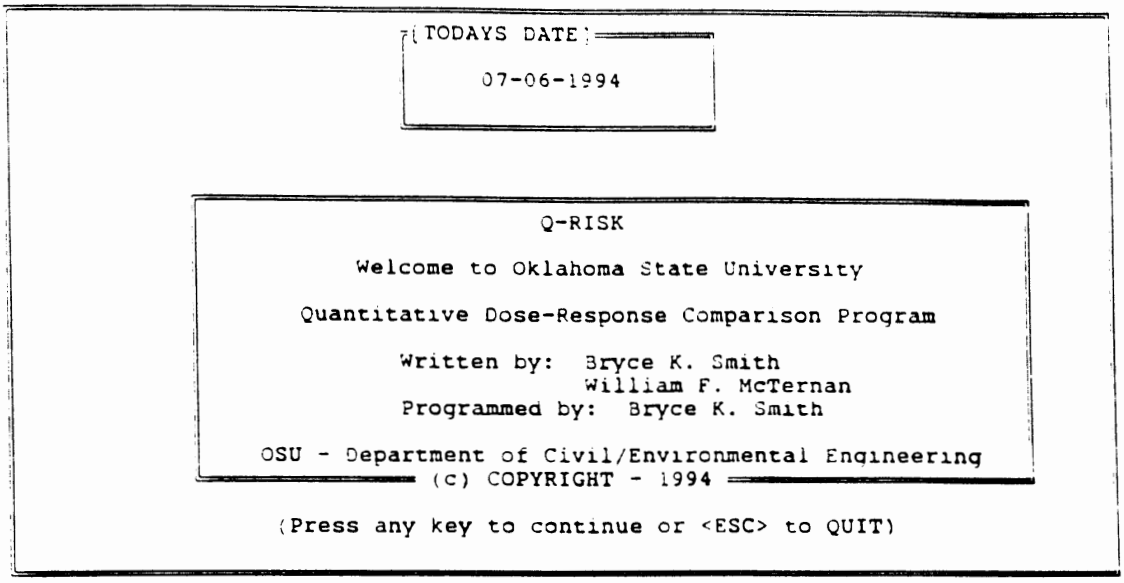


FIGURE 20
SECOND WELCOME SCREEN

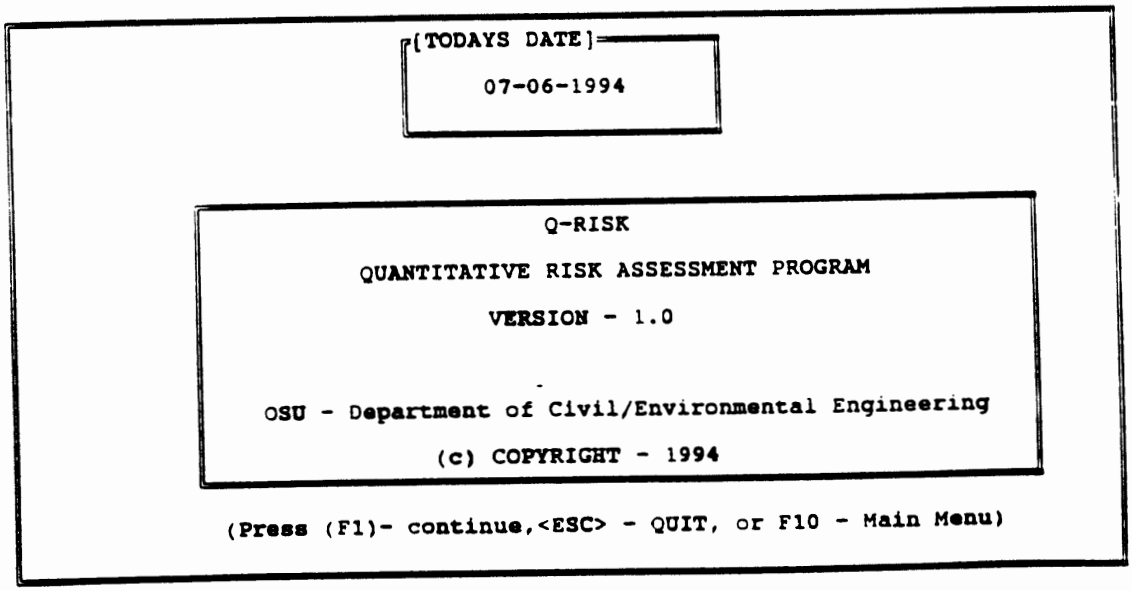


FIGURE 21
FIRST INTRODUCTORY INFORMATION SCREEN

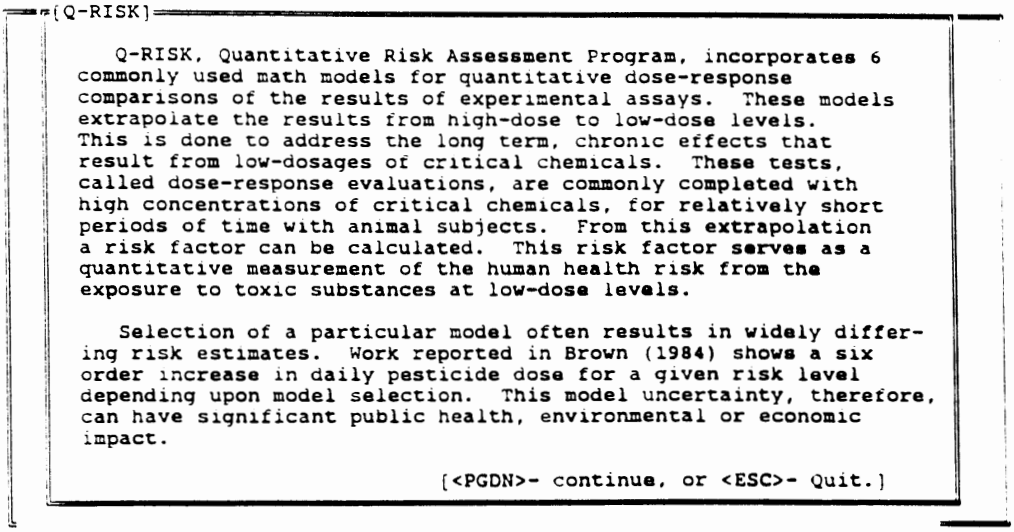


FIGURE 22
SECOND INTRODUCTORY INFORMATION SCREEN

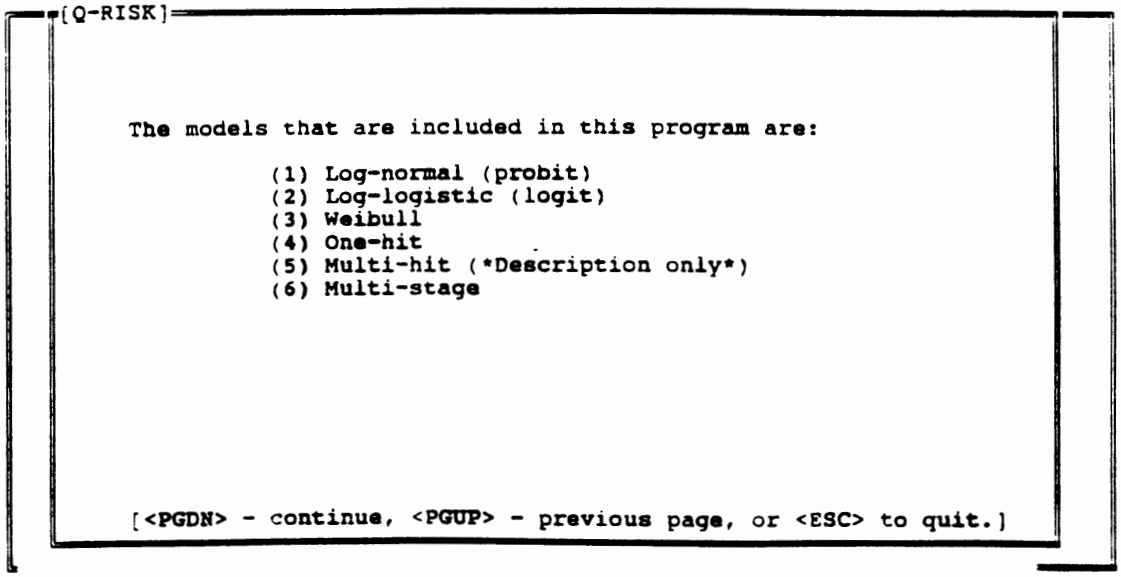


FIGURE 23
THIRD INTRODUCTORY INFORMATION SCREEN

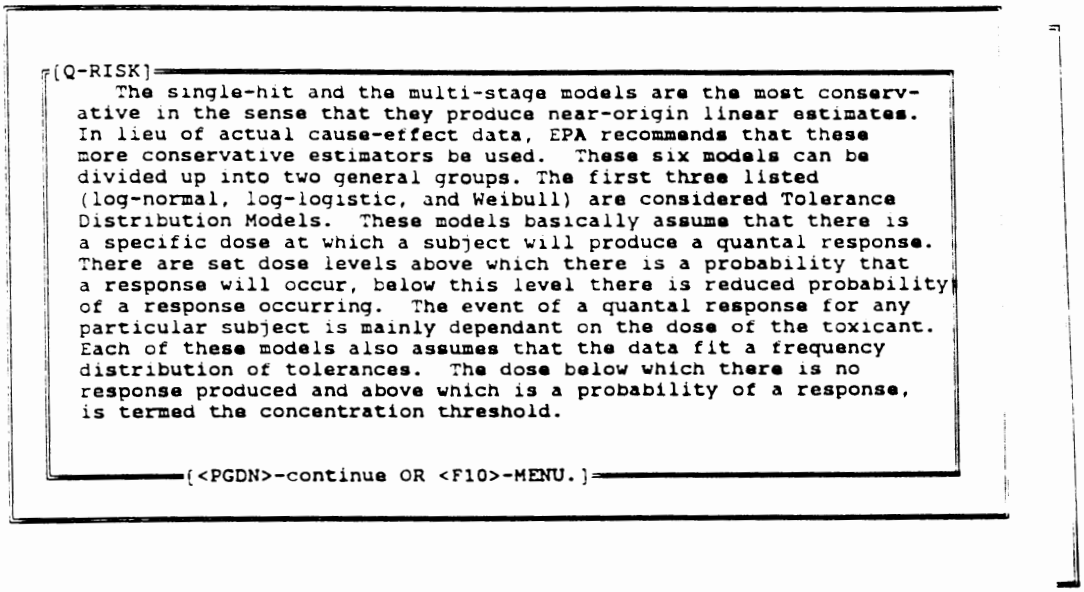


FIGURE 24
FOURTH INTRODUCTORY INFORMATION SCREEN

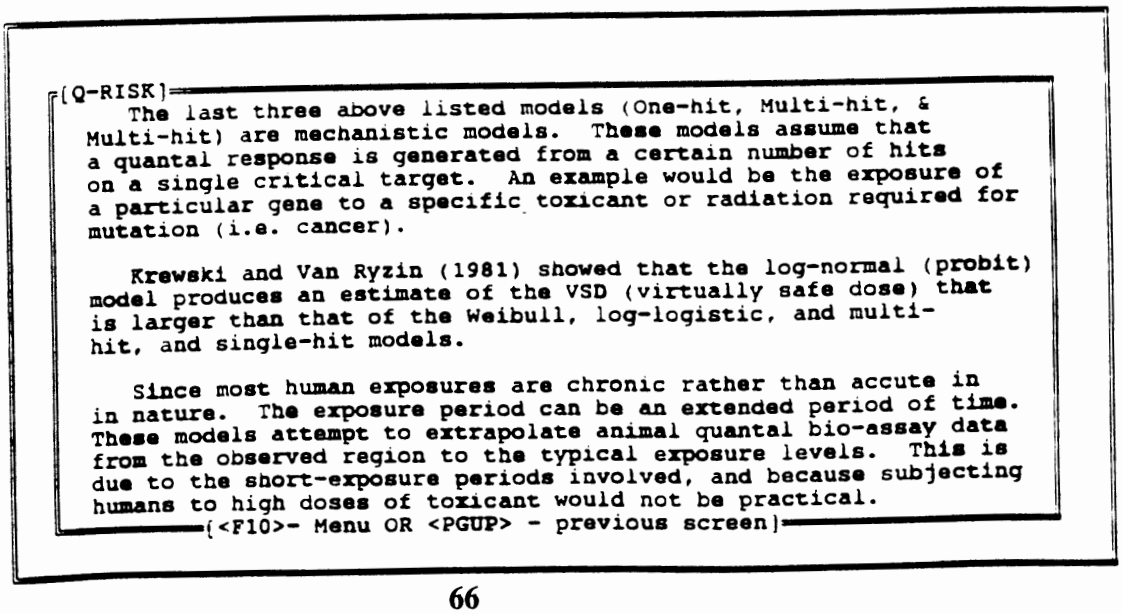


FIGURE 25
PRE-MAIN MENU SCREEN

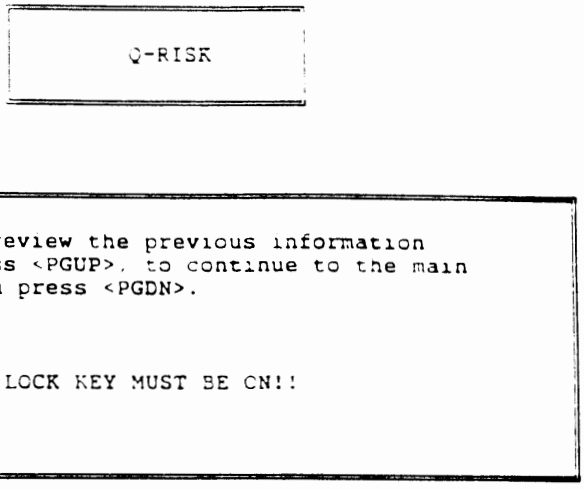


FIGURE 26
EXAMPLE SCREEN OF MODEL INFORMATION

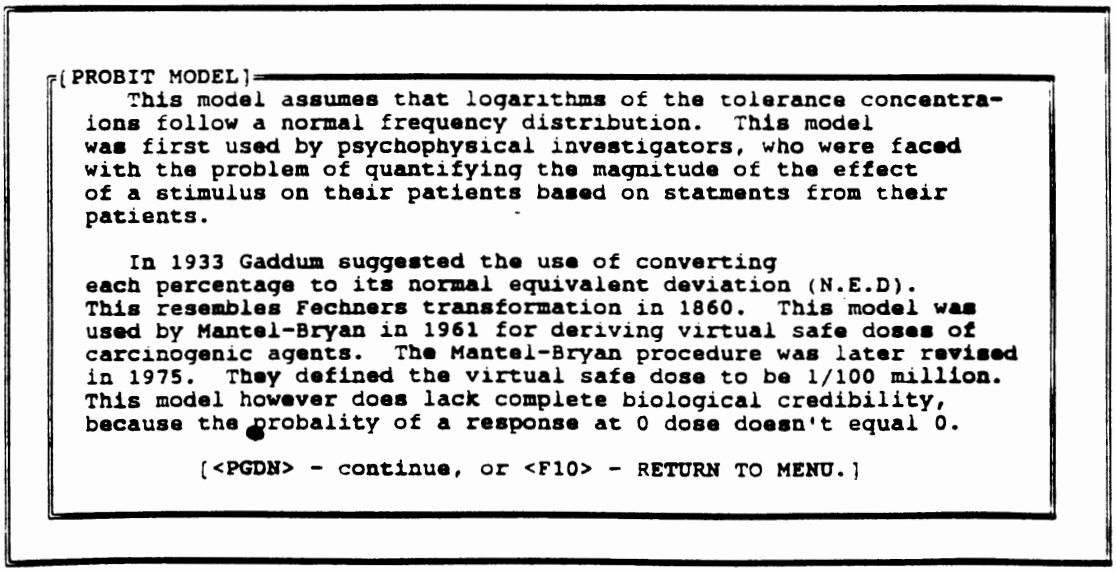


FIGURE 27
EXAMPLE OF DATA REQUIREMENT SCREEN

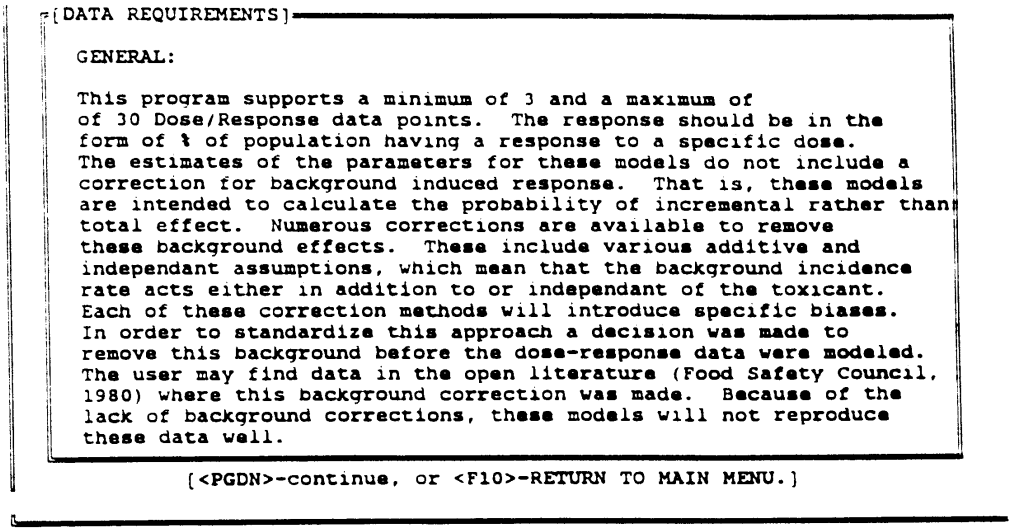


FIGURE 28
EXAMPLE OF DATA LIMITATION SCREEN

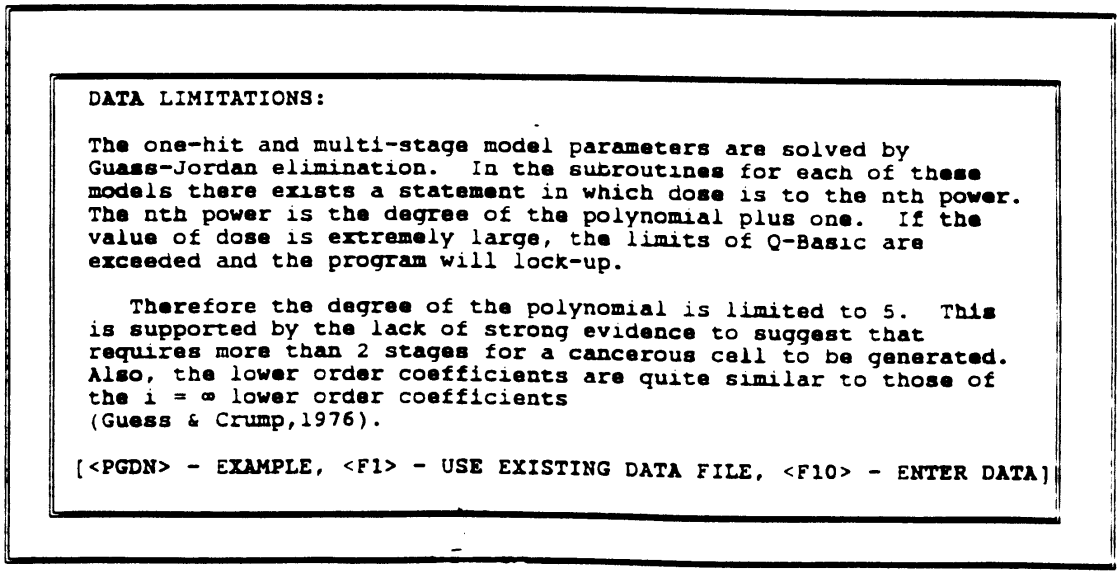


FIGURE 29
DATA ENTRY EXAMPLE SCREEN

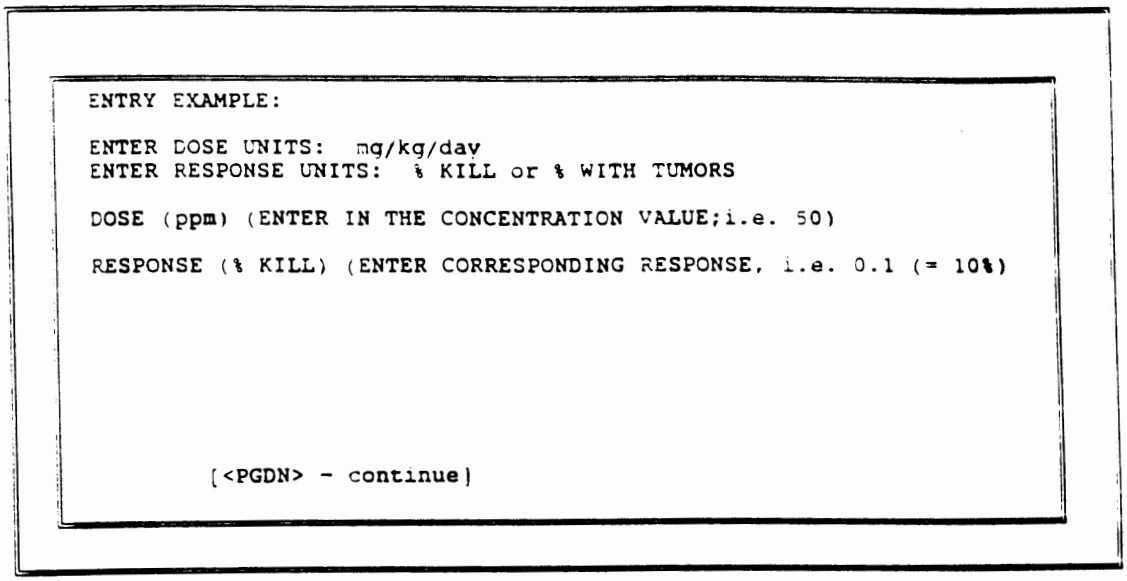


FIGURE 30
PLOTTING HELP SCREEN

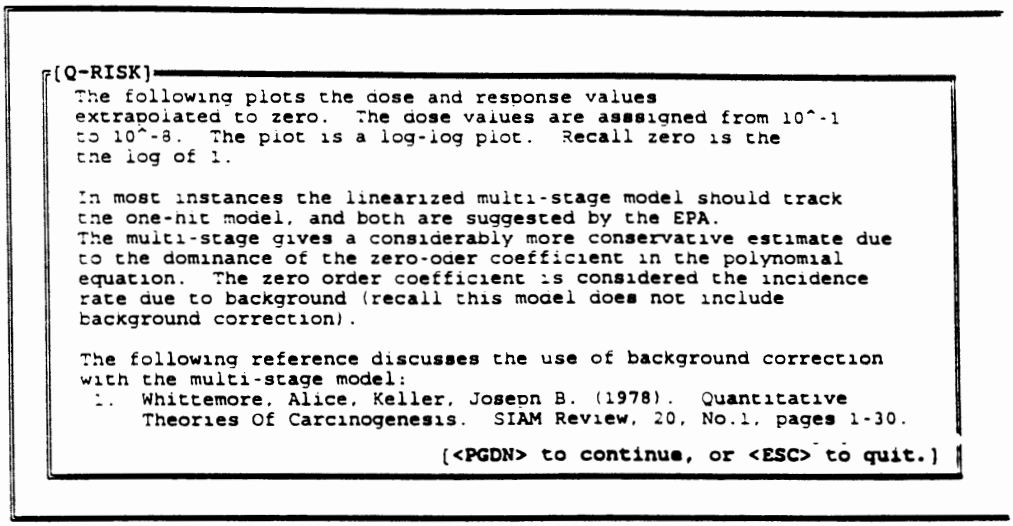


FIGURE 31
MAIN MENU SCREEN

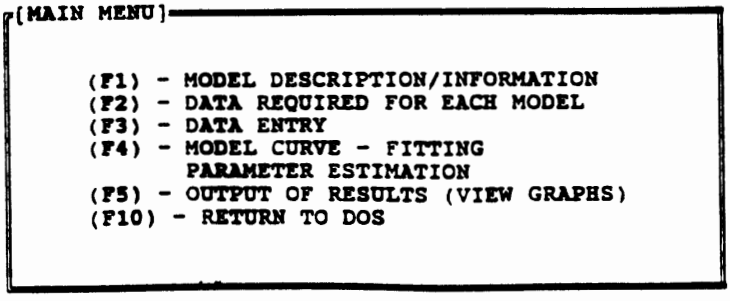


FIGURE 32
MODEL INFORMATION MENU SCREEN

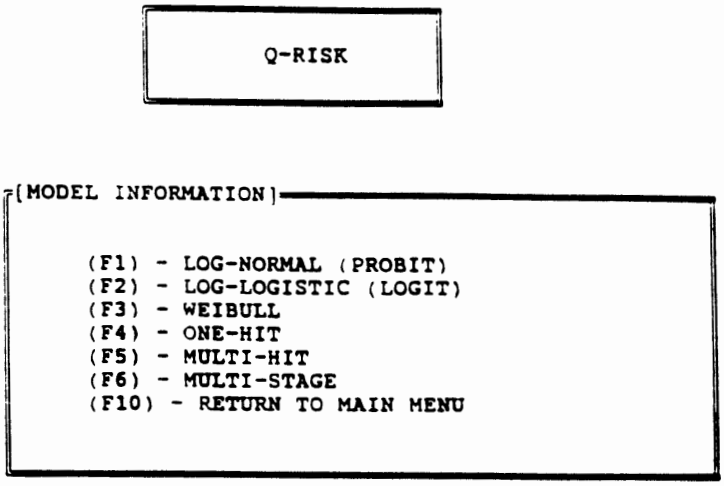


FIGURE 33
PARAMETER ESTIMATION MENU SCREEN

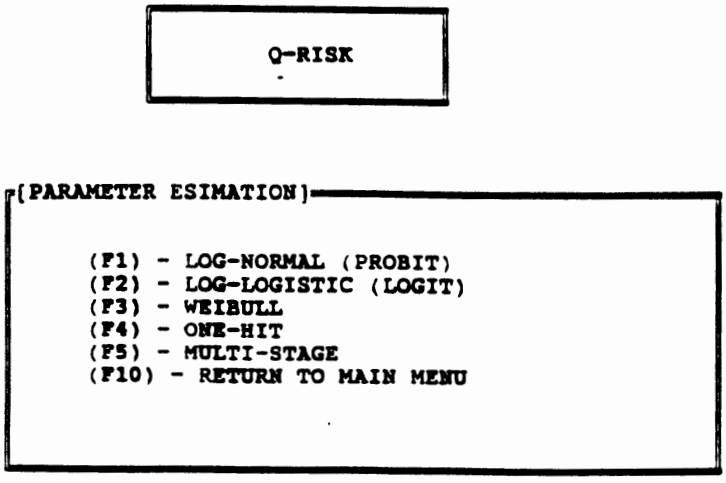


FIGURE 34
RESULTS MENU SCREEN

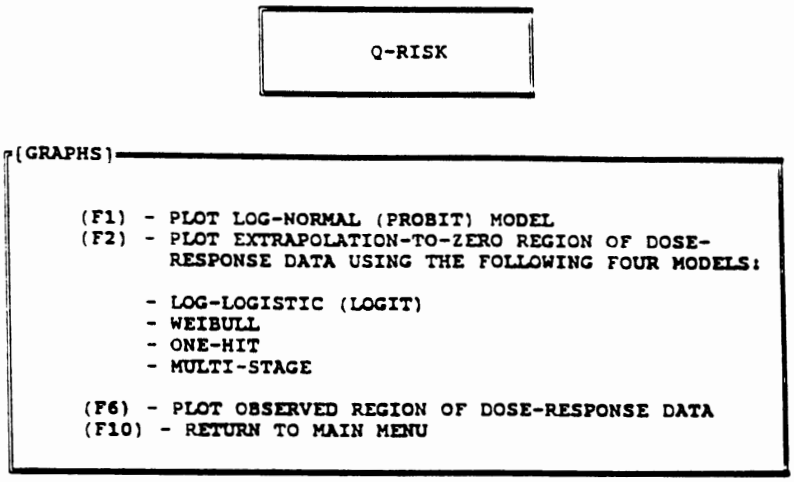


FIGURE 35
EXTRAPOLATED PLOT MENU SCREEN

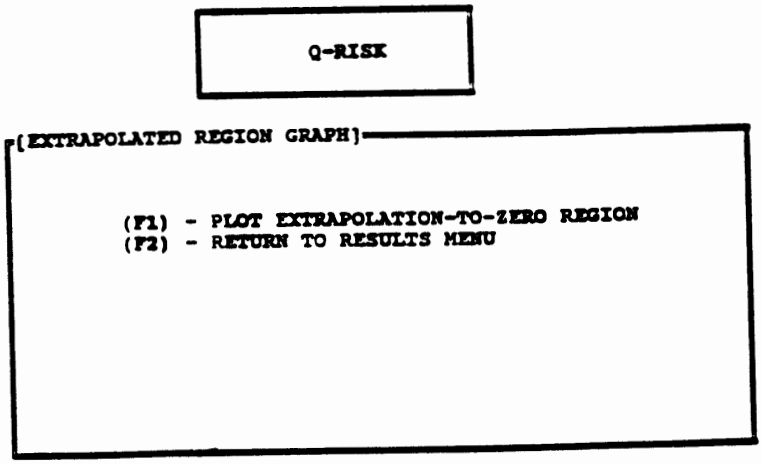


FIGURE 36
Q-RISK SCREEN FOR LOG-NORMAL PARAMETER ESTIMATE RESULTS

```
                    LOG-NORMAL (PROBIT) PARAMETERS

LINEAR REGRESSION RESULSTS
--------------------------

A (y-intercept) =          -2.25400
B (slope) =          0.97001
r (correlation coefficient) = 0.913136661


                LINEARIZED LOG-NORMAL (PROBIT) EQUATION:

                        z = B*LOG*(D) + A

              (Press any key to continue or <ESC> to QUIT)
```

FIGURE 37
Q-RISK SCREEN LOT-LOGISTIC PARAMETER ESTIMATE RESULTS

```
                    LOG-LOGISTIC PARAMETERS

LINEAR REGRESSION RESULTS
-------------------------
A (Y-INTERCEPT) =          -4.04844
B (SLOPE) =          1.73824
r (correlation coefficient) = 0.930654407


            LOG-LOGISTIC LINEARIZED EQUATION:

                    P(D) = 1/[1+EXP^-(A + B*LOG*(D))]

            (Press any key to continue or <ESC> to QUIT)
```

FIGURE 38
Q-RISK SCREEN FOR WEIBULL PARAMETER ESTIMATE RESULTS

```
                                WEIBULL PARAMETERS

LINEAR REGRESSION RESULTS
-------------------------

A =              0.02041
ln A (y-intercept) =            -3.8919
B (slope) =               0.64556
r (correlation coefficient) = 0.945852816


                                WEIBULL LINEAR EQUATION:

                                     P(D) = 1 - EXP^-A*D^B

                            (Press any key to continue or <ESC> to QUIT)
```

FIGURE 39
Q-RISK SCREEN FOR ONE-HIT PARAMETER ESTIMATE RESULTS

```
                                ONE-HIT PARAMETERS

Á( 0) = +4.31296229362E-03
Á( 1) = +4.31640958413E-03

X^2 (CHI-SQUARE) =     0.1770
DEGREES OF FREEDOM =   3


Á(0) = RESPONSE DUE TO BACKGROUND
Á(1) = COEFFICIENT USED IN THE ONE-HIT EQUATION

                                ONE-HIT EQUATION:

                                     P(D) = 1 - EXP^[Á(1)*D]
                            (Press any key to continue or <ESC> to QUIT)
```

FIGURE 40
Q-RISK SCREEN FOR MULTI-STAGE PARAMETER ESTIMATE RESULTS

```
                    MULTI-STAGE PARAMETERS

COEFFICIENT VALUES OF THE (i)+1 DEGREE POLYNOMIAL
-------------------------------------------------
α( 0) = +2.34617888927E-02
α( 1) = +7.42176035419E-03
α( 2) = -1.98967845790E-05
α( 3) = -2.12154918700E-06
α( 4) = +8.60481730314E-09


α(0) corresponds to the response due to background
α(i) corresponds to the coefficient of the (i)th stage.
α(i) is the coefficient in the equation below.
i = stage of the cell : i + 1 = degree of polynomial

               MULTI-STAGE EQUATION:

                    P(D) = 1 - EXP^Σ[-α(i)*D^(i)] , 0 < i < ∞

          (Press any key to continue or <ESC> to QUIT)
```

represents Y in the linear equation, $Y = mX + b$. In Figures 39 and 40 "$\alpha(\#)$" represents coefficients of the polynomial equation. Figures 41 through 43 represent each of the three plots Q-Risk generates.

Output Of Model

Output Files. The program automatically generates output files that can be viewed and printed under the DOS 5.0 or 6.x editors. These files contain the estimated parameters and the equations for each model. These output files are named corresponding to the name of the model with the file extension *.out (i.e. Weibull.OUT).

Graphs. Three graphs are generated for each simulation. The first plots the original data points while the second plots the $\log_{10}$ of each dose versus the standard normal variate. Figure 41 is an example of the first type of graph. This graph is obtained after estimating the parameters and selecting the "Output Of Results" function from the main menu. Figure 42 is an example of the second type of graph. This graph is produced by following the same steps as listed for Figure 41. The third graph plots the "extrapolated to zero" portion of the dose versus response curve as in Brown (1984) for the log-logistic, one-hit, Weibull, and multi-stage models. The response axis ranges from 10^{-1} to 10^{-9}. This allows the user to determine dose that corresponds to a certain unit risk. This would depend on what the user views as a VSD. Figure 43 represents this third type of graph and is produced from the main menu in the same manner as Figures 41 and 42. These three figures were created using the DDT exposure data given in

FIGURE 41
Q-RISK PLOT OF DOSE VS. RESPONSE IN OBSERVED REGION

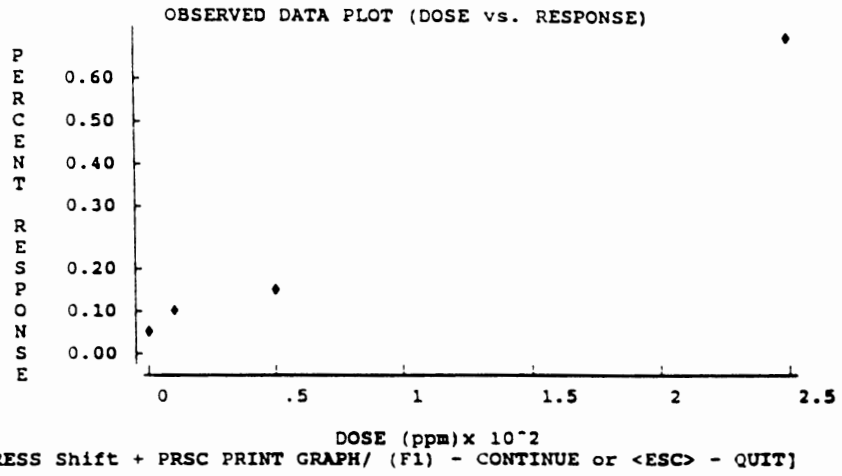


FIGURE 42
EXAMPLE OF Q-RISK PLOT FOR THE PROBIT MODEL

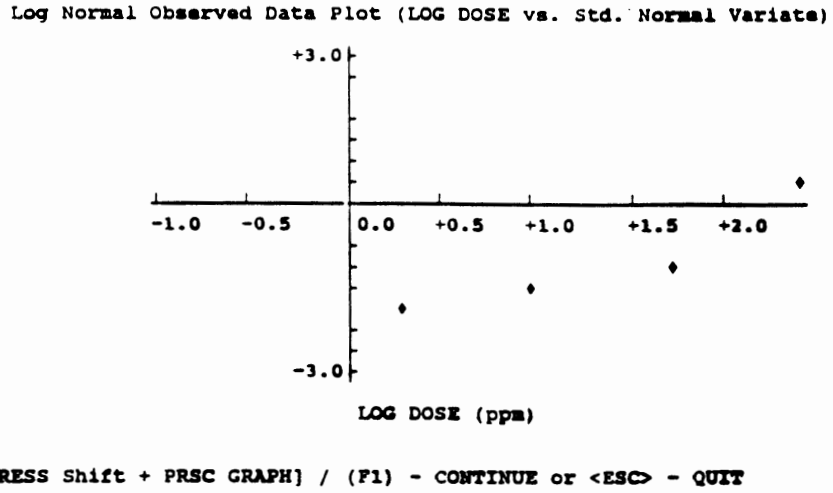
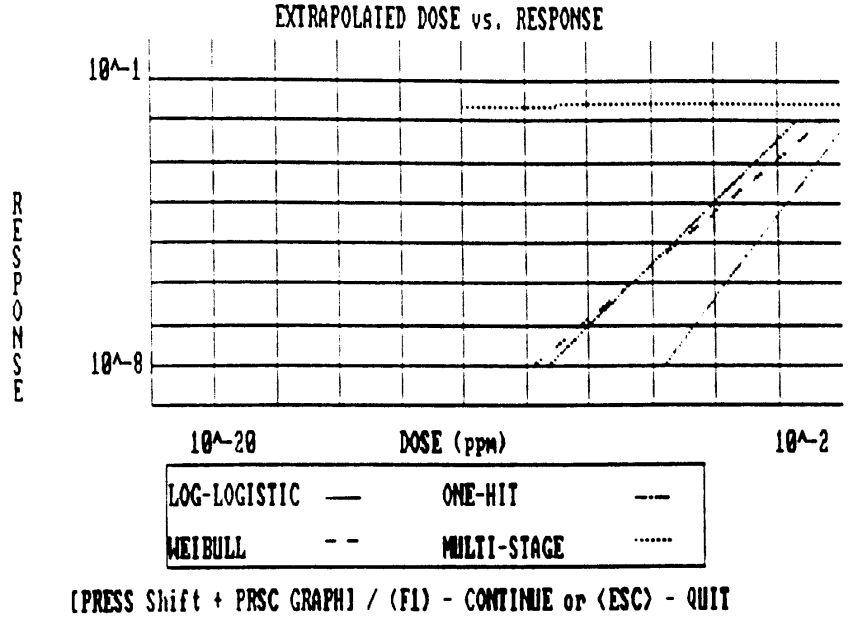


FIGURE 43
Q-RISK PLOT OF DDT DATA



Food Safety Council (1980). The first two graphs mentioned allowed for the user to alter the range of the axis.

Plotting. The graphing sequences allowed the user to either plot the original data points, the Probit $\log_{10}$ dose vs. standard normal variate, or the extrapolation-to-zero dose of the data. The x and y axes are automatically scaled and plotted upon selection of any of the three plotting function options. The selection, "Graph Low-Dose Extrapolation," features the extrapolation-to-zero dose of the data using the estimated parameters calculated from the log-logistic, Weibull, one-hit, and multi-stage models. Varying the x or y axis was not programmed into this sequence, because the region of interest will be displayed for every data set. The plot of the extrapolation-to-zero of the data was actually a log-log plot. Risk values of 10^{-2} and 10^{-8} were used to calculate the corresponding doses using the estimated parameters and then a line is drawn between them. These dose-response data pairs were then converted to the $\log_{10}$ and plotted. The plot of the response axis is from 10^{-1} to 10^{-9}. The user must have gone through the parameter estimation procedure for all models for this extrapolated region to be plotted, because the data files created from these estimation procedures are called to calculate the dose-response data pairs. The user was allowed to view the actual plot, and was given the option to print the plot by using the keys Shift + PrintScrn.

TABLE 7

DOSE-RESPONSE DATA FOR EXAMPLE 1

DOSE (mg)	RESPONSE
0.1	0.05
0.3	0.10
2.0	0.20

Source: (Hallenbeck, 1988)

Access of *.DAT Files and Creating of *.DAT and *.OUT Files

Upon execution of the aforementioned programs, the user chosen *.DAT file is accessed by opening the data file corresponding to the chosen name. Then the parameters used in the calculations are read from this file. The *.DAT files that are created are ASCII type files. The *.OUT files are text files. These are created upon completion of the parameter estimation sequence. The *.DAT files containing the parameters for each model are "zeroed" out upon entry to Q-RISK. This is done by erasing the *.DAT files created for each model. This is to avoid any incosistant comparisons.

Tutorial Screens

Q-Risk was coded with help/tutorial screens that explain each model (theory, uses, parameters, and limitations). These screens aided the user in making decisions (i.e., multi-stage polynomial degree selection) by providing them with this background information for these models.

Comparison Of Model Output To Literature Output

To test the validity of the model equations, the parameters estimation procedures, and the graphing of the fitted data, the Q-Risk program was executed using dose-response data points cited in different research publications. Then the output was compared to that from the original publications.

Example 1. Table 7 lists the original dose-response data points from Hallenbeck (1988). These data points were given as examples with Hallenbeck (1988). Data was entered into Q-Risk by selecting the data entry option. Next,

the parameters for the log-logistic, log-normal, and Weibull models were estimated by parameter estimation option. The Probit, Log-Logistic, and the Weibull models parameters were estimated with Q-Risk. Table 8 gives a comparison of the parameter values cited in Hallenbeck (1988) and those calculated by Q-Risk. The relative percent difference (RPD) assesses the precision of Q-RISK's parameter estimations. Usually, a RPD of less than 50 percent is considered acceptable between two data points. RPD is a quality control measure used in EPA SW-846 methods to assess precision of the analytical methods (EPA, 1986). See Appendix A for the RPD calculation. RPDs will show how precise Q-RISK estimates the parameters of these models as compared to literature values. This shows that Q-Risk estimates the parameters for these models with a great degree of precision.

Example 2. The Food Safety Council (1980) performed an investigation of the One-Hit, Multi-Hit, Weibull, Armitage-Doll, and Probit Model for use in the low-dose extrapolation of chronic cancer bioassay data. Table 9 lists the dose-response data for the substance DDT as described by Food Safety Council (1980). Mice were fed the pesticide DDT at parts per million (ppm) concentration (mg/kg) and the number exhibiting tumors was recorded. Table 10 compares dose values corresponding to a 10^{-8} risk generated by Q-Risk and those from the Food Safety Council, 1980. These values were read from the extrapolated graphs of these data (Figures 43 and 44). All data points were analyzed except the zero dose-response data point. Recall that Q-Risk handles only incremental responses

TABLE 8

PARAMETER COMPARISONS FOR EXAMPLE 1

PROBIT MODEL

PARAMETERS	Q-RISK RESULTS	LITERATURE RESULTS	Relative Percent Difference (RPD)
A (Y-intercept)	-1.010	-1.010	0
B (slope)	0.606	0.609	0.49
r (correlation co-efficient)	0.995	0.995	0

LOG-LOGISTIC MODEL

PARAMETERS	Q-RISK RESULTS	LITERATURE RESULTS	Relative Percent Difference (RPD)
A (Y-intercept)	-1.700	-1.700	0
B (slope)	1.170	1.170	0
r (correlation co-efficient)	0.992	0.992	0

WEIBULL MODEL

PARAMETERS	Q-RISK RESULTS	LITERATURE RESULTS	Relative Percent Difference (RPD)
A (Y-intercept)	0.167	0.167	0
B (slope)	0.480	0.480	0
r (correlation co-efficient)	0.990	0.990	0

TABLE 9

DDT DOSE RESPONSE DATA FOR EXAMPLE 2

DOSE (ppm)	% RESPONSE (Tumors)
2	0.0381
10	0.0887
50	0.1250
250	0.6667

Source: (Food and Safety Council, 1980)

TABLE 10

DOSE COMPARISON DATA FOR EXAMPLE 2 AT A 10^{-8} RISK

MODEL	Q-RISK RESULTS	LITERATURE RESULTS	Relative Percent Difference (RPD)
One-Hit	6E-05	6E-05	0
Multi-stage	< 1E-20	5E-05	200
Weibull	7.E-09	2.0E-03	200

Source: (Food Safety Council, 1980)

TABLE 11

POLYNOMIAL COEFFICIENT COMPARISONS FOR EXAMPLE 2

PROBIT MODEL

COEFFICIENTS	Q-RISK RESULTS	LITERATURE RESULTS	Relative Percent Difference (RPD)
α_0	2.346E-02	4.483E-02	63
α_1	7.422E-03	2.038E-03	114
α_2	-1.989E-05 OR 0[1]	0	0
α_3	-2.122E-06 OR 0[1]	0	0
α_4	8.605E-09	1.39E-09	144

Notes:

1 - If the coefficient value is negative it is considered to be zero.

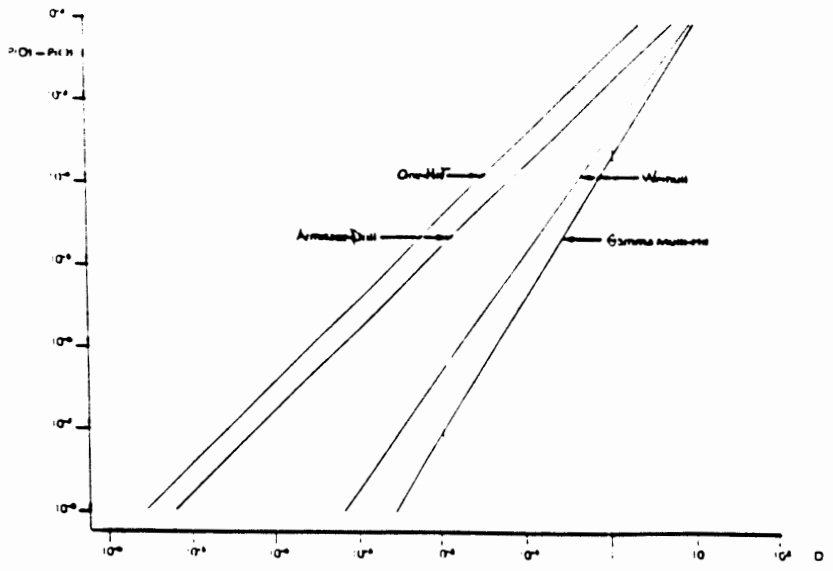
Source: (Guess and Crump, 1976)

above background. The multi-stage and Weibull low-dose extrapolation lines differed significantly from those generated by the Food Safety Council (1980). There is clearly a significant difference (a large RPD) in the Weibull and multi-stage models. Figure 43 and Figure 44 clearly show these differences. This appeared to be due to the Food Safety Council using a correction for background response of another variation to the Armitage-Doll model equation that was not made evident to the reader. All other model plots from Q-Risk seem to match those from the literature in example 2. Table 10, which compares the dose read from the extrapolated graphs corresponding to a risk of 10^{-8} for the Weibull and one-hit models, shows almost an exact match for the one-hit model values but a vast difference for the Weibull and multi-stage models values. Therefore, Q-Risk produces a lower dose estimate (more conservative) when using the Weibull and multi-stage models, but is exactly similar with respect to the one-hit model.

Example 3. Guess and Crump (1976) developed a maximum likelihood estimation procedure to calculated the polynomial coefficients for the Armitage-Doll multi-stage model. Table 11 compares the coefficients calculated by Q-Risk to those from Guess and Crump (1976) for a fourth degree polynomial.

It is evident from Table 11 (large RPD), which compares the polynomial coefficients calculated for the multi-stage model by Q-Risk and by Guess and Crump, that some of the differences in the coefficients between Guess and Crump (1976) values and the value generated by Q-Risk were significant. However, due to the exrtemely small (10^{-9}) value of these coefficients these differences could be

FIGURE 44
FOOD AND SAFETY COUNCIL OF DDT DATA



Source: (Food and Safety Council, 1980)

to the exrtemely small (10^{-9}) value of these coefficients these differences could be due to rounding or significant figure differences between the two methods of estimation. The data used for the parameter estimation in Q-Risk may vary with that used by Guess and Crump (1976). For example Guess and Crump (1976) could have included a zero dose data point that produced a response when example 2 did not (i.e. background correction). The data used by Guess and Crump were taken from a mouse DDT study. The interpretation of dose-response data from the literature could have been significantly different than that of Guess and Crump. That is, extracting the dose-response data required interpreting instructions from the author as to which data to use (i.e. female or male). That is, numerous tables of dose and response data were given. Percentage of mice with tumors had to be calculated for both the female and male mouse data. Guess and Crump (1976) did not state which sex they used in their study. Also, when calculating the response was not apparent which numbers were to be used (i.e. the number of mice exhibiting).

By graphing the low-dose extrapolation region, the user can select a risk level that corresponds to a VSD by just selecting a point from the graph. A safety factor should be used in estimating a VSD. The EPA uses a 95 percent confidence level and the FDA uses a 99 percent confidence level to estimate risk. At present, the model does not complete these calculations.

The selection of some near-zero lifetime risk, either 10^{-8} (proposed by Mantel and Bryan) or 10^{-6} (proposed by the FDA) is a decision made by the user

for determination of VSD (Food Safety Council, 1980). The Food Safety Council (1980) suggests that the decision should be left up to the regulatory authorities.

CHAPTER V

CONCLUSIONS

Q-Risk was designed to aid the user in the process of performing risk assessments for carcinogens or toxic chemicals that pose a health risk to the human population. The code specifically addresses the dose-response portion of the Risk Assessment process by applying five commonly employed models to extrapolate from the high-dose, short duration testing typically completed in toxicity testing to the low dose, long term patterns thought typical of chronic disease propagation. The goal was to combine the power of Quickbasic, a modern, graphics-based complex computer programming language, with the mathematics of the various dose-response models. This provided the user with a program that requires little computer knowledge to operate. "Help screens" were added to aid the user in decision making. Although not all mathematical models that exist are made available to the user, the ones most frequently used in the scientific community for low-dose extrapolation were incorporated. The models included in the program were:

- Probit
- Log-Logistic
- Weibull
- One-Hit
- Multi-Stage

The multi-hit model was described in Q-Risk but was not included in the low-dose extrapolation performed by Q-Risk. This was done because of the similarity of the multi-hit model with the Weibull model (Brown, 1984).

The Probit, Weibull, and Log-Logistic model parameters were estimated using a linear regression sequence as opposed to maximum likelihood estimation. This was done to simplify coding and was considered appropriate given previous work by others. Based on the results from chapter IV, there was no significant difference between the two methods of estimation (i.e., example 3). The parameters estimated for these models by Q-Risk matched those cited in literature exactly in at least one example. The parameters for the One-Hit and Multi-Stage models were also estimated using a Gauss-Jordan elimination sequence instead of a maximum likelihood estimation procedure. This was done because of the need to normalize the method of estimation within the mechanistic models. The number of k-stages, of disease initiation and propagation, which equals the degree of the polynomial, were limited to five in the program, to be consistent with previous observations relative to physical evidence that cancer does not occur in more than two stages (Whittemore and Keller, 1978).

Some of the problems in the past with environmental decision-making processes include the lack of user friendly computer programs for those who are not computer literate. This limited the ability of the scientist to make valid risk decisions. Q-Risk was an attempt to lower the barrier between the scientist with little computer knowledge and his/her data while also supplying the scientist with

an aid in the area of risk assessment of toxic or carcinogenic compounds that pose a human health hazard. The complexity and time necessary to manually compute the parameters for the dose-response models and plotting of the results are greatly shortened.

The use of Microsoft® Quickbasic to generate the code allowed the production of user friendly screens and, powerful graphics, while incorporating powerful mathematic functions. The program was structured so that the user could easily view results with output files generated in a form that could be viewed or printed under any appropriate text editor. The code was compiled as "stand-alone execute files" it does not require BASIC files to run the program) so that the user could run the program from any IBM®-DOS based computer.

This program allows even the least-computer-knowledgeable scientist to precisely assess the incremental risk above background response of toxic chemicals or carcinogens that may pose a human health risk.

The following conclusions can be made:

- A user-friendly, graphics-based computer code was developed to allow comparison between dose-response models
- A powerful mathematical tool was developed to aid the user in calculating unit risk above incremental cancers by performing low-dose extrapolation.
- A program was developed for users with little computer knowledge, and addressed the problem of toxicity assessment of human health risks.
- A code was developed that was user friendly and aids the user in environmental decision making.

REFERENCES

Berkson, Joseph (1944). Application of the Logistic Function to Bio-Assay. Journal of the American Statistical Association, 39, 357-365.

Bliss, CI (1934). The Method of Probits. Science, 79, 38-9.

Bliss, CI (1935). The Calculation of The Dosage-Mortality Curve. Annals Of Applied Biology, 22, 134-67.

Brown, Charles, C. (1984). High-to Low-Dose Extrapolation in Animals. American Chemical Society, Assessment And Management Of Chemical Risks, Chapter 5, 57-79.

Carlborg, F. W. (1981). Dose-Response Functions In Carcinogenesis And The Weibull Model. Food And Cosmetics Toxicology, 19, 255-63.

Crump, K. S., Hoel, D. G., Langley, C. H., and Peto, R. (1976). Fundamental Carcinogenic Processes and Their Implications for Low Dose Risk Assessment. Cancer Research, 36, 2973-79.

EPA, (1989). RISK ASSESSMENT GUIDANCE FOR SUPERFUND VOLUME I Human Health Evaluation Manual (Part A) Interim Final. U.S. EPA Office of Emergency and Remedial Response, EPA/540/1-89/002, Washington, D.C., Means, Bruce (Ed.), 1989.

EPA, (1986). SW-846 METHODS VOLUME I. U.S. EPA, 1986.

Fenner-Crisp, Penelope A., Ph.D., D.A.B.T. (1986). Pesticides and Groundwater: A Health Concern for the Midwest. Proceedings of a conference held on Oct 16-17, 1986. Published by the Freshwater Foundation. 233-97.

Finney, D. J. (1971). Probit Analysis. Cambridge At The University Press, 3rd edition.

Fisher, J. C., Hollman, J. H. (1951). A Hypothesis For The Origin of Cancer Foci. Cancer; Diagnosis, Treatment, Research. 916-8.

Gaddum, J.H. (1933). Medical Research Council, Special Report, Series No. 183, 1933.

Gehring, P. J. (1977), Mechanisms Of Carcinogens: Dose Response. Journal of Environmental Pathology and Toxicaology, 1, 163-79.

Grossman, Elizabeth A., Maronik, John (1990). OSHA's Approach To Risk Assessment for Setting a Revised Occupational Exposure Standard for 1,3-Butadiene. Environmental Health Perspective, 86, 155-158.

Guess, H. A., Crump, K. S. (1977). Can We Use Animal Data To Estimate "Safe" Doses For Chemical Carcinogens?. Envrionmental Health Quantitative Methods, Proceedings of a Conference sponsered by SIAM, July 5-9, 1976. SIAM, Philadelphia, 1977. 13-27.

Guess, H. A., Crump, K. S. (1976). Low-Dose Extrapolation of Data from Animal Carcinogenicity Experiments-Analysis of New Statistical Technique. Mathematical Biosciences, 32, 15-36.

Guess, H. A., Crump, K. S. (1978). Best-Estimate Low-Dose Extrapolation of Carcinogenicity Data. Environmental Health Perspectives, 22, 149-152.

Hallenbeck, William H. (1988). Quantitative Risk Assessment for Environmental and Occupational Health, 2nd Edition, Chapter 4, 63-102.

Hanes, B., Wedel, T. (1985). A Selected Review of Risk Models: One-Hit, Multihit, Multistage, Probit, Weibull, and Pharmacokinetic. Journal Of The American College Of Toxicology, 4, Number 4, 271-278.

Hoel, D. G., Gaylor, D. W., Kirschstein, R. L., Saffiotti, Umberto, Schneiderman, M. A. (1975). Estimation Of Risks Of Irreversible, Delayed Toxicity. Journal of Toxicology and Environmental Health, 1, 133-151.

Mantel, N., Bryan, W. R. (1961). "Safety" Testing of Carcinogenic Agents. Journal Of The National Cancer Institute, 27, No. 2, 455-470.

Mantel, N., Bryan, W. R. (1975). An Improved Mantel-Bryan Procedure for "Safety" Testing of Carcinogens. Cancer Research, 35, 865-872.

Maxim, Daniel L., Harrington, L. (1984). A Review of the Food and Drug Administration Risk Analysis for Polychlorinated Biphenyls in Fish. Regulatory Toxicology and Pharmacology, 4, No. 2, 192-219.

Rai, K., Ryzin, J. V. (1979). Risk Assessment of Toxic Environmental Substances Using a Generalized Multi-Hit Dose Response Model. Energy and Health, Whitemore, Alice S. (Eds.). SIAM: Philadelphia. 99-117.

Rai, K., Ryzin, J. V. (1987). A Dose-Response Model Incorporating Nonlinear Kinetics. Biometrics, 43, 95-105.

Rall, David P. (1978). Thresholds?. Environmental Health Perspective, 22, 163-65

Rodricks, Joseph V., Tardiff, Robert G. (1984). Conceptual Basis for Risk Assessment. American Chemical Society, Assessment And Management Of Chemical Risks, Chapter 1, 3-12.

Schneiderman, Marvin A., Mantel, N., Brown, Charles C. (1975). From Mouse To Man--Or How To Get From The Laboratory To Park Avenue And 59TH Street. Annals of The New York Academy of Sciences, 246, 237-48.

Schoeny, R. S. (1991). Pesticide Residues and Food Safety: A Harvest of Viewpoints. ACS Symposium Series 446, 214-220.

Scientific Committee, Food Safety Council (1980). Proposed System for Food Safety Assessment. Washington, D.C.. 137.

Sime, Rodney J. (1988). Physical Chemistry: Methods/Techniques/Experiments. Saunders Golden Sunburst Series. 72-130.

Simons, G. L. (1985). Expert Systems and Micros. NCC Publications, Oxford Road, Manchester M1 7ED, England.

Temple, Barker, & Sloane, Inc, Lexington, MA (1987). Principles of Risk Assessment. Environ Corp., Washington, D. C.

Wardlaw, Gordon M. (1985). Assessing the cancer risk from foods. Journal of the American Dietetic Association, 85, No. 9, 1122-7.

Whittemore, Alice, Keller, Joseph B. (1978). Quantitative Theories Of Carcinogenesis. SIAM Review, 20, No. 1, 1-30.

Wilson, Edwin B., Worcester, J. (1943). The Determination Of L.D. 50 And Its Sampling Error In Bio-Assay. Proceedings of the National Academy of Sciences of the United States of America, 29, 78-85.

APPENDICES

APPENDIX A

LINEAR TRANSFORMATIONS AND RPD CALCULATIONS

APPENDIX A

LINEAR TRANSFORMATIONS

GAUSS-JORDAN Elimination

The Gauss-Jordan elimination method involves eliminating all the variables except for one, then substituting it back into the equation and systematically solving for the other variables. Equation 38 represents the nth degree polynomial that was used in the Gauss-Jordan elimination estimation method (Sime, 1988).

$$y = \lambda_0 + \lambda_1 x + \lambda_2 x^2 + \lambda_3 x^3 + \ldots + \lambda_k x^k \qquad (38)$$

Where y equals the response, x the dose, and λ_k is the coefficient of the k_{th} stage of the cell. The following equations (39a-c) represent an example taken from Sime, 1988 of the Gauss-Jordan elimination method.

$$2x_1 + 3x_2 + 8x_3 = 84 \qquad (39a)$$

$$x_1 + 7x_2 - 3x_3 = 65 \qquad (39b)$$

$$5x_1 - 2x_2 + x_3 = 41 \qquad (39c)$$

By multiplying the second equation by -2, adding the product to the first equation, and replacing the second equation by the sum the following equations are produced.

$$2x_1 + 3x_2 + 8x_3 = 84$$

$$-11x_2 + 14x_3 = -46$$

$$5x_1 - 2x_2 + x_3 = 41$$

This eliminates x_1 from the second equation. By multiplying the third equation by -2.5, "... adding the quotient to the first equation, and replacing the third equation by the sum." (Sime, 1988).

$$2x_1 + 3x_2 + 8x_3 = 84$$

$$-11x_2 + 14x_3 = -46$$

$$3.8x_2 + 7.7x_3 = 67.6$$

Now by multiplying the third equation by 2.8947 (11/3.8), "adding the result to the second equation, and replacing the third equation by the sum." (Sime, 1988).

$$2x_1 + 3x_2 + 8x_3 = 84$$

$$-11x_2 + 14x_3 = -46$$

$$36x_3 = 149.6782$$

Next backward substitution starting with the value of x_3 in the second equation the values for the other two parameters can be calculated.

The parameters for equation 38 are solved for in the same manner as mentioned above. A first degree polynomial the derivatives produce two equations with two unknowns, λ_0 and λ_1. The derivatives give rise to p + 1 equations in p + 1 unknowns, namely, λ_0, λ_1, λ_2,. To evaluate these constants it is necessary to solve a system of p + 1 simultaneous linear equations.

$$\lambda_0 n + \lambda_1 \Sigma x_i + \lambda_2 \Sigma x_i^2 + \ldots + \lambda_k \Sigma x_i^k + \ldots + \lambda_p \Sigma x_i^p - \Sigma y_i = 0 \quad (40)$$

$$\lambda_0 \Sigma x_i \; \lambda_1 \Sigma x_i^2 + \lambda_2 \Sigma x_i^3 + \ldots + \lambda_k \Sigma x_i^{k+1} + \ldots + \lambda_p \Sigma x_i^{p+1} - \Sigma x_i y_i = 0 \quad (41)$$

$$\lambda_0 \Sigma x_i^k y_i^k + \lambda_1 \Sigma x_i^{k+1} + \lambda_2 \Sigma x_i^{k+2} + \ldots + \lambda_k \Sigma x_i^{k+k}$$

$$+ \ldots + \lambda_p \Sigma x_i^{p+k} - \Sigma x_i^k y_i = 0 \qquad (42)$$

$$\lambda_0 \Sigma x_i^p + \lambda_1 \Sigma x_i^{p+1} + \lambda_2 \Sigma x_i^{p+2} + \ldots + \lambda_k \Sigma x_i^{p+k}$$

$$+ \ldots + \lambda_p \Sigma \lambda_i^{p+p} - \Sigma x_i^p y_i = 0 \qquad (43)$$

The solution to this system of simultaneous linear equations is the set of values of the coefficients λ_i. The augmented matrix is written

n	x_i	x_i^2	x_i^k	...	x_i^p ...	y_i
x_i	x_i^2	x_i^3	x_i^{k+1}	...	x_i^{p+1} ...	$x_i y_i$
x_i^k	x_i^{k+1}	x_i^{k+2}	x_i^{k+k}	...	x_i^{p+k} ...	$x_i^k y_i$
x_i^p	x_i^{p+1}	x_i^{p+2}	x_i^{p+k}	...	x_i^{p+p} ...	$x_i^p y_i$

The x_i (i.e. dose) and y_i (i.e. response) are the experimental points we wish to fit to a polynomial of degree p. "The number of x_i, y_i pairs equals n, so the summations are from i = 1 to n. The number of pairs must be greater than the degree of the polynomial and is often much greater." (Sime, 1988).

LOGISTIC Model

Equation 34 is transformed into a form of the linear equation y = mX + b.

$$P_e = 1/(1 + e^{-(a + b*\log D)}) \qquad (34)$$

This is done by first rearranging the equation algebraically and taking the

$$1/P_e = 1 + e^{-(a + b*\log D)}$$

$$-1-(1/P_e) = e^{-(a + b*\log D)}$$

natural logarithm of both sides. This results in the following equation.

$$- \ln [(1-P_e)/P_e] = a + b*\log D \quad (35)$$

WEIBULL Model

Equation 36 for the Weibull model is transformed in the same manner, except that the natural logarithm of both sides is taken twice to produce the double natural logarithm in equation 37.

RPD Calculations

Equation 43a represents the method by which RPDs are calcuated as described in EPA (1986). The x_1 and x_2 values represent the first and second values for which the RPD is being calculated for. For example the first value would be the Q-RISK value and the second the value from Guess and Crump (1976).

$$RPD = (x_1 - x_2)/((x_1 + x_2)/2) * 100 \quad (43a)$$

VITA

Bryce K. Smith

Candidate for the Degree of

Master of Science

Thesis: DEVELOPMENT OF ENVIRONMENTAL-SUPPORT SOFTWARE TO DETERMINE AND COMPARE DOSE-RESPONSE EXTROPLATIONS

Major Field: Environmental Engineering

Biographical:

Personal Data: Born in Bartlesville, Oklahoma, February 10, 1965, the son of DeLacey and Louise Smith. Married June 6, 1987. One child born October 5, 1990.

Education: Graduated from Bartlesville High School, Bartlesville, Oklahoma, in June 1983; received Bachelor of Science Degree in Zoology from the University of Oklahoma, Norman, Oklahoma in May, 1988; completed requirments for the Master of Science degreee at Oklahoma State University in December, 1994.

Professional Experience: Chemist, Vista Chemical Company, Austin, Texas, July 1988 to July 1991.
Chemist, Southwest Laboratory of Oklahoma, August, 1991, to June, 1994.
Environmental Engineer, PRC Environmental, Kansas City, Kansas, June 1994 to present.