135142

Fd Chem. Toxic. Vol. 25, No. 3, pp. 257–267, 1987
Printed in Great Britain.

0278-6915/87 $3.00 + 0.00

QUANTITATIVE CANCER RISK ASSESSMENTS FOR 2,3,7,8-TETRACHLORODIBENZO-*p*-DIOXIN (TCDD)

R. L. SIELKEN, JR*
Department of Statistics, Texas A&M University, College Station, TX 77843, USA

(*Received* 3 *April* 1986)

Summary—State-of-the-art quantitative risk assessment techniques, including consideration of new time-to-response data, have been applied to chronic animal bioassay data on the dietary intake of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). The non-linear shapes of the dose-response relationships for the hepatocellular carcinogenic responses have been estimated, and a review of the quantitative impacts of several of the choices involved in the quantitative risk assessment considers, particularly, the definition of the carcinogenic responses of concern, the experimental data set, the pathology evaluation, a biologically effective dose scale versus the administered dose, methods of making the fitted model responsive to the data at the lower experimental doses, consistency in dose-response shapes for different data sets, fitted model values versus bounds, the utilization of time-to-response information incorporating the lateness of the carcinogenic responses, and the method of characterizing the maximum acceptable dose. The estimated virtually safe dose for an increase of 0.000001 (one in a million) in the probability of hepatocellular neoplastic nodule and/or carcinoma in a female rat is approximately 0.1 ng/kg body weight/day in the diet. The estimated mean free dose, corresponding to a reduction in the expected amount of time without hepatocellular neoplastic nodule and/or carcinoma proportional to 1 wk in 70 yr, is in the range of 1–5 ng/kg body weight/day in the diet of a female rat. No species-to-species extrapolations nor human exposure assessments have been made. However, these estimated risks correspond to dietary intakes that are at least 150 times greater than the 0.0006365 ng/kg body weight/day intake described by the Centers for Disease Control as a reasonable level to begin consideration of action to limit human exposure.

Introduction

Quantitative risk assessments can be an important part of an overall risk assessment and risk management decision. However, the developer of a quantitative risk assessment faces many uncertainties that force him to make several choices, assumptions and judgements. Hence, several different quantitative risk assessments can be developed depending upon the decisions taken. These alternative quantitative risk assessments may not all have the same implications. Nor are they necessarily equally reflective of the available scientific information. Before a particular quantitative risk assessment is allowed to influence, or to continue to influence, risk management decisions, a thorough analysis of the quantitative impacts of the choices made in the particular assessment should be presented, and the scientific merit of these choices should be carefully reviewed.

The new scientific information available on 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and the new quantitative risk assessment models and techniques continually being developed, make an updated quantitative risk assessment for TCDD appropriate.

*Address for correspondence: Sielken, Inc., Suite 410, 3833 Texas Avenue, Bryan, TX 77802, USA.

Abbreviations: MFD = mean free dose; NTP = National Toxicology Program; TCDD = 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; VSD = virtually safe dose.

Impact of alternatives in TCDD risk assessment

Definition of the response of concern

Quantitative dose-response models depend heavily on which events are defined as being 'responses'.

Table 1 reproduces the Kimbrough *et al.* (1984) summary of tumour incidences found to differ significantly in the treated and control Sprague–Dawley rats in the Kociba *et al.* (1978) diet study on TCDD. It is interesting that there are as many significant decreases in tumour incidence as there are significant increases. Table 2 reproduces the Kimbrough *et al.* (1984) summary of the incidence of dose-related tumours observed in the NTP gavage studies on TCDD in Osborne–Mendel rats and B6C3F$_1$ mice (National Toxicology Program, 1982).

Any of the tumours or tumour-like lesions listed in Table 1 or 2 could be defined as the 'response of concern'. In this discussion, the focus will be on the liver lesions. These were the only lesions suggesting a possible dose-response relationship in both of the rat strains and in the mice. In addition, the virtually safe doses (VSDs) suggested by Kimbrough *et al.* (1984) were smaller for liver lesions than for any other response.

Experimental data set

The variability in the risk characterization over different experimental data sets is illustrated in Table 3. For each of several liver data sets, the fitted model

Table 1. Total and individual tumours differing significantly in incidence in TCDD-treated and in untreated control male and female Sprague–Dawley rats

Tumour or tumour-like lesion	Incidence in groups of rats fed diets providing TCDD doses (μg/kg body weight/day) of: 0 M	0 F	0.001 M	0.001 F	0.01 M	0.01 F	0.1 M	0.1 F
No. of rats examined...	85	86	50	50	50	50	50	49
Hepatocellular neoplastic nodules	6	8	0	3	3	18*	2	23*
Hepatocellular carcinoma	2	1	0	0	3	2	1	11*
Hepatocellular neoplastic nodule or carcinoma	—	9	—	3	—	18	—	34
Stratified squamous-cell carcinoma of hard palate or nasal turbinates	0	0	0	0	0	1	7*	4*
Keratinizing squamous-cell carcinoma of lung	0	0	0	0	0	0	1	7*
Benign tumour of uterus	—	28	—	12	—	11	—	7†
Benign neoplasm of mammary gland	—	73	—	35	—	36	—	24†
Mammary carcinoma	—	8	—	4	—	4	0	0†
Pituitary adenoma	26	43	6	18	11	13	13	12†
Subcutaneous carcinomas	10	1	1†	1	5	0	6	0
Acinar adenoma of pancreas	14	0	7	1	5	0	2†	1
Adenoma of adrenal cortex	0	9	0	6	2	2	5*	5
Phaeochromocytoma of adrenal gland	28	7	6	3	10	1	4†	2

TCDD = 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin

Data adapted from Kimbrough *et al.* (1984) and Kociba *et al.* (1978). Values marked with superscripts are greater (*) or lower (†) than the control value to a statistically significant degree ($P < 0.05$) by the Fischer Exact Probability Test.

value for the VSD is listed for the multistage model. For the moment only one particular risk characterization is being considered—the VSD corresponding to an increase of 0.000001 (one in a million) in the probability of the specified response over the background probability of the same response. Similarly, only one particular quantal dose-response model is being considered—the three-stage multistage model. There are many other ways of characterizing the risk associated with a particular dose-response relationship and many other models.

In Table 3 there is a considerable difference in the

Table 2. Incidences of dose-related tumours in Osborne–Mendel rats and B6C3F$_1$ mice treated with TCDD

Response	Sex	Rats: Dose (μg/kg body weight/wk)	Rats: Tumour incidence	Mice: Dose (μg/kg body) weight/wk)	Mice: Tumour incidence
Thyroid: follicular-cell adenoma	M	0.00	1/69		
		0.01	5/48		
		0.05	6/50		
		0.50	10/50 ($P = 0.001$)		
	F	0.00	3/73	0.00	0/69
		0.01	2/45	0.04	3/50
		0.05	1/49	0.20	1/47
		0.50	6/47	2.00	5/46 ($P = 0.009$)
Lymphoma and leukaemia	F			0.00	18/74
				0.04	12/50
				0.20	13/48
				2.00	21/47
Hepatocellular nodule, adenoma or carcinoma	M	0.00	0/74	0.00	15/73
		0.01	0/50	0.01	12/49
		0.05	0/50	0.05	13/49
		0.50	4/50	0.50	27/50
	F	0.00	5/75	0.00	3/73
		0.01	1/49	0.04	6/50
		0.05	3/50	0.20	6/48
		0.50	14/49 ($P = 0.001$)	2.00	11/47
Hepatocellular carcinoma	M	0.00	0/74	0.00	8/73
		0.01	0/50	0.01	9/49
		0.05	0/50	0.05	8/49
		0.50	1/50	0.50	17/50 ($P = 0.002$)
	F	0.00	0/75	0.00	1/73
		0.01	0/49	0.04	2/50
		0.05	0/50	0.20	2/48
		0.50	2/49	2.00	6/47 ($P = 0.014$)

TCDD = 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin

Adapted from Kimbrough *et al.* (1984) and NTP (1982).

Table 3. The quantitative impact of the definition of the response of concern and the experimental data set on the fitted multistage model value for the VSD of TCDD in the diet

Reference	Species	Sex	Fitted multistage value of VSD (μg/kg body weight/day): Hepatocellular adenoma, neoplastic nodule, or carcinoma	Fitted multistage value of VSD (μg/kg body weight/day): Hepatocellular carcinoma	Ratio of VSDs for different responses*
Kociba *et al.* (1978)	Sprague–Dawley rat	M	NDR	NDR	NDR
		F	7.7×10^{-8}	4.0×10^{-7}	5.2
NTP (1982)	Osborne–Mendel rat	M	2.6×10^{-3}	1.6×10^{-3}	0.6
		F	1.3×10^{-4}	1.3×10^{-3}	10.0
NTP (1982)	$B6C3F_1$ mouse	M	1.3×10^{-7}	2.6×10^{-7}	2.0
		F	1.4×10^{-6}	2.5×10^{-6}	1.8
	Ratio: largest VSD/smallest VSD...		33,766	6154	

VSD = Virtually safe dose TCDD = 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin
NDR = No apparent dose-related increase in the response

*The ratio in this column is the fitted model value for the VSD when the response is defined as hepatocellular carcinoma divided by the fitted model value for the VSD when the response is defined as hepatocellular adenoma, neoplastic nodule or carcinoma.

fitted model value for the VSD for either specified response. When the specified response is hepatocellular adenoma, neoplastic nodule or carcinoma, the largest VSD is approximately 34,000 times the smallest VSD, and when the specified response is hepatocellular carcinoma, the largest VSD is approximately 6000 times the smallest VSD. Furthermore, there is usually at least a tenfold difference between the sexes for the same species and strain. In rats the females appear to be at greater risk than the males while the revese is apparently the case in mice.

Since the female Sprague–Dawley rats appear generally to be the most sensitive species/strain/sex combination, the remaining discussion will focus primarily on this data set. This does not mean, however, that the dose-response relationship in female rats is necessarily predictive of any human dose-response relationship.

It is also important to note that the VSD is very dependent upon the definition of the 'response of concern'. For the female Sprague–Dawley rats the VSD is over five times larger for the more severe response (hepatocellular carcinoma) than for the less severe response (hepatocellular neoplastic nodule or carcinoma). Since most regulatory decisions should be concerned with the health effects of an exposure and since the health effects of a neoplastic nodule may be considerably less than that of a carcinoma, the choice of the response of concern should be very carefully evaluated.

Which pathologist's evaluations to use?

Frequently, different pathologists evaluate lesions differently. This is particularly true of the less severe carcinogenic responses. The original pathology done by Kociba *et al.* (1978) was reviewed by R. A. Squire (Environmental Protection Agency, 1980). Figure 1 and Table 4 compare the original pathology for hepatocellular neoplastic nodule or carcinoma with Squire's evaluation. Although the magnitudes of the

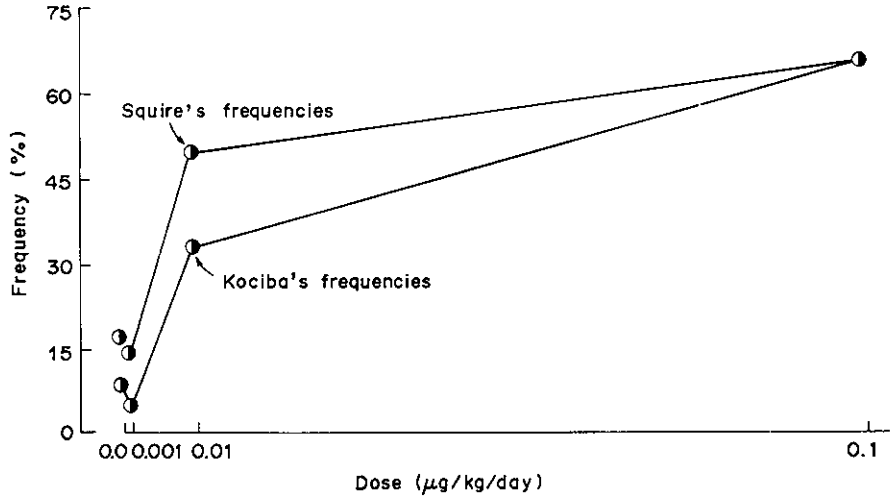

Fig. 1. The observed frequencies of hepatocellular neoplastic nodule or carcinoma in female Sprague–Dawley rats originally reported by Kociba *et al.* (1978) show a dose-response relationship very similar to that reported by Squire in his re-evaluation of the slides from the Kociba *et al.* study.

Table 4. Alternative pathologists' evaluations of the Kociba *et al.* (1978) female Sprague–Dawley rat data on hepatocellular neoplastic nodule or carcinoma and the corresponding fitted multistage model value for the VSD

TCDD dose (μg/kg body weight/day)	Pathologist	Incidence
0.0	Kociba	9/86
	Squire	16/86
0.001	Kociba	3/50
	Squire	8/50
0.01	Kociba	18/50
	Squire	27/50
0.1	Kociba	34/49
	Squire	33/47

TCDD = 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin
Fitted model values for the virtually safe dose (VSD) derived from the Kociba and Squire evaluations were 7.7×10^{-8} and 8.6×10^{-8} μg/kg body weight/day, respectively.

response frequencies are higher in Squire's evaluation, the two corresponding dose-response relationships in Fig. 1 are nearly parallel. Also the fitted multistage model value for the VSD is nearly the same using either evaluation. The original pathology will serve as the primary data set in this paper.

Modelling on an administered dose scale or a biologically effective dose scale

The modelling of the dose-response relationship should use a dose scale that is as biologically relevant as possible. The quantal response models (including the multistage model) and even the time-to-response models are based on simplistic conceptions of the occurrence of a carcinogenic response and do not even attempt to be detailed biological or mechanistic models of cancer causation. Hence many scientists view the fitting of these simplistic models more as 'curve fitting' than as really 'modelling' the carcinogenic process. This view of the data fitting and the associated high-to-low-dose extrapolation as primarily curve fitting is reinforced if the dose level is expressed on an administered dose scale (e.g. the concentration in the air or diet) as opposed to a biologically effective dose scale. The biologically effective dose scale should reflect the available information on the amount of cancer-related biological activity associated with the administered dose. Thus, the biologically effective dose should reflect the physiological and pharmacokinetic components of the chemical's delivery (or the delivery of its active metabolite) to the target site, the proportion of the amount delivered that can enter into the cancer causation, and the effects of repair systems and immune systems.

Kociba *et al.* (1978) reported that terminal liver samples contained mean TCDD concentrations of 0.540, 5.10 and 24.0 ppb in the groups administered 0.001, 0.01 and 0.1 μg/kg body weight/day, respectively. Assuming that these concentrations were stable for most of the 2-year study, they provide a dose scale that is more biologically relevant for modelling the dose-response relationships concerning hepatocellular responses than is the administered dose scale.

Accepting or rejecting a model's fit to the experimental data

Figure 2 shows the fits of the three-stage multistage model to the frequencies of hepatocellular neoplastic nodule or carcinoma in female rats (Kociba *et al.* 1978) using both the administered dose scale and the liver concentration scale. The fits are both really very poor. Both fitted models are too high at the lowest non-zero dose and too low at the intermediate non-zero dose.

The chi-square goodness-of-fit test is not powerful enough to reject the fits of the three-stage multistage model to these experimental data on hepatocellular neoplastic nodule or carcinoma in a female rat. If:

$d_i = 0,\ 0.001,\ 0.01,\ 0.1$ for $i = 1,\ 2,\ 3,\ 4$ respectively,

Observed_i = number of female rats with hepatocellular neoplastic nodule or carcinoma at dose d_i,

n_i = number of rats at risk at dose d_i

and

$\hat{P}(d_i)$ = maximum likelihood estimate using the three-stage multistage model of the probability of the specified response at dose d_i,

then the chi-square goodness-of-fit statistic is:

$$\sum_{i=1}^{4} [\text{Observed}_i - n_i \hat{P}(d_i)]^2 / \{n_i \hat{P}(d_i)[1.0 - \hat{P}(d_i)]\}$$

The observed values of this statistic are 8.7 using the administered dose scale and 3.0 using the liver concentration scale. Both of these values are less than the 95th percentile of a chi-square distribution with four degrees of freedom ($\chi^2_{4,0.95} = 9.49$).

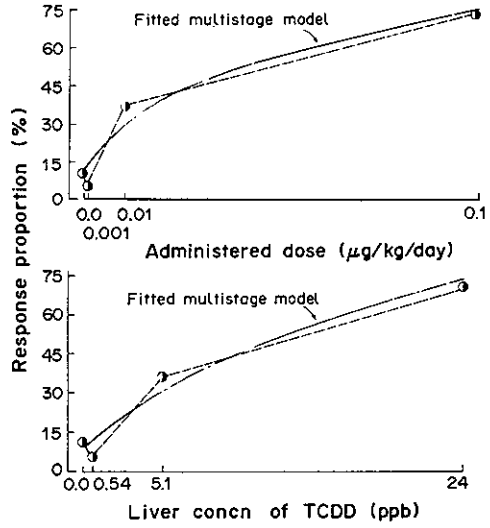

Fig. 2. When the multistage model is fitted to the experimental data on hepatocellular neoplastic nodule or carcinoma in female rats in the Kociba *et al.* (1978) study, the fitted models do *not* reflect the observed behaviour at the lower experimental doses, regardless of whether the dose is expressed as the administered dose (a) or as the liver concentration (b).

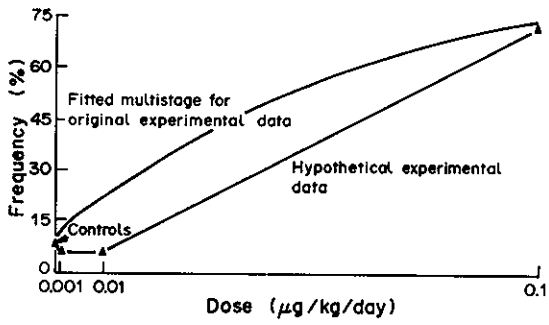

Fig. 3. The chi-square goodness-of-fit test is not powerful enough to reject the fitted three-stage multistage model for the frequency of hepatocellular neoplastic nodule or carcinoma in female rats in the Kociba *et al.* (1978) study, even if the experimental data were the hypothetical experimental data depicted.

Unfortunately, proposed dose-response models are often blindly accepted whenever the chi-square goodness-of-fit test does not reject them. Although the test will not reject a true model more often than is specified by the significance level, it will frequently not reject a false model either. The test's weakness stems primarily from its failure to detect the patterns in the discrepancies between the fitted model's predictions and the observed experimental data. In other words, the test is weak because it does not consider the *shape* of the fitted dose-response relationship. For example, the fitted multistage model shown in Fig. 2 for the administered dose scale would not have been rejected by the chi-square goodness-of-fit test even if the experimental data had been as shown in Fig. 3. Thus, a fitted model can be well above the low-dose experimental data points and the chi-squared goodness-of-fit test may still not be powerful enough to reject the fitted model.

Highest experimental dose level

False impressions can arise from fitting quantal response models to very high administered doses. In the Kociba *et al.* (1978) study, the highest experimental dose level (0.1 μg/kg body weight/day) is 100 times bigger than the lowest non-zero experimental dose level (0.001 μg/kg/day). Even the second highest experimental dose level (0.01 μg/kg/day) is ten times the lowest non-zero experimental dose level.

The trade-offs inherent in curve fitting may lead to questionable fits in the low-dose region of primary concern. The fitted models in Fig. 2 both reflect trade-offs in the sense that the fitted models are shaped differently from the experimental data and the fitted model response rates relative to the observed rates are too large at the lowest non-zero experimental dose level and too small at the intermediate non-zero level. Furthermore, the need to accommodate the highest experimental dose forces the fitted models to be very non-responsive to whatever is observed at the lowest non-zero dose level. For example, as illustrated in Fig. 4, the presence or absence of the experimental data at the lowest experimental dose has very little effect on the overall shape of the fitted models and has only a relatively small effect on the fitted model values for the VSD.

The non-parametric Kaplan–Meier estimates of the probability of a female rat developing a hepatocellular neoplastic nodule or carcinoma by the end of 25 months (the duration of the Kobica *et al.* (1978) study) are shown in Fig. 5. These estimates are computed separately for each dose level and take into account the observation times of each rat. The values of these estimates for the two highest dose levels are 0.81 at 0.01 μg/kg/day and 1.00 at 0.1 μg/kg/day. Thus, by the end of the experimental period the number of rats that could develop a hepatocellular response is completely saturated at 0.1 μg/kg/day and nearly saturated at 0.01 μg/kg/day. It is impossible for the multistage model to portray both this type of saturation phenomenon and the observed non-linearity at the lower experimental dose levels (0, 0.001 and 0.01 μg/kg/day). The conflict caused by the model's inability to reflect both behaviours is dealt with in the fitting process by essentially ignoring the lower-dose behaviour and focusing on depicting the relative flatness at the higher doses.

The flattening out or levelling off of the observed dose-response relationship at high doses causes the fitted models to modify their lower-dose shape dramatically in order to have a relatively flat higher-dose

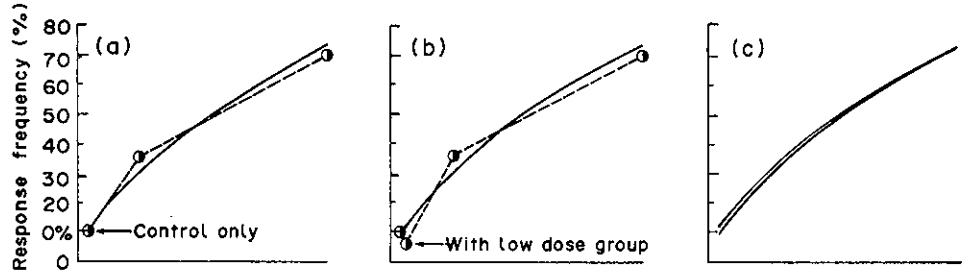

Fig. 4. Both the fit of the multistage model to the Kociba *et al.* (1978) experimental data on hepatocellular neoplastic nodule or carcinoma in female rats and the corresponding estimate of the virtually safe dose (VSD) are dominated by the higher experimental dose levels and are essentially unaffected by the lower dose level, as shown by (a) from which the experimental data at the lower dose level are *excluded* (VSD = 4.1×10^{-8} μg/kg body weight/day), (b) in which domination of the high dose data continues even when the experimental data at the lower dose level are *included* (VSD = 8.1×10^{-8} μg/kg/day) and (c) which demonstrates the *non-responsiveness* to the inclusion of the experimental data at the lower dose level (VSD 8.1×10^{-8}/VSD $4.1 \times 10^{-8} \triangleq 2$).

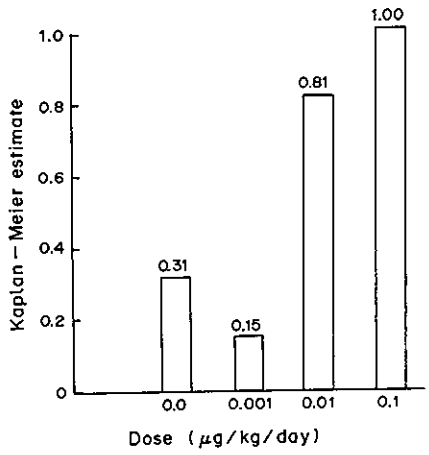

Fig. 5. Kaplan–Meier estimates of the probability of a female rat developing a hepatocellular neoplastic nodule or carcinoma in the Kociba *et al.* (1978) study suggest that the dose-response relationship resembles one in which a saturation-like phenomenon occurs at the highest experimental dose level.

shape. This change in the shape of the fitted models at the lower doses is illustrated in Fig. 6, where the fitted multistage model on the administered dose scale is relatively linear if 0.1 μg/kg/day is included in the fitting process and is upward curving in the lower-dose region if 0.1 μg/kg/day is excluded. These two fits have very different shapes in the lower-dose region, particularly between 0 and 0.001 μg/kg/day. Hence the use of the data for the very high dose here leads to fitted models with an unsupported false impression of linearity in the lower-dose region. This pitfall will be lessened by fitting the models to the observed proportions excluding the highest dose level.

A figure similar to Fig. 6 results if the definition of a response is changed from hepatocellular neoplastic nodule or carcinoma to hepatocellular carcinoma. Hence the fitted models for this latter response will also be obtained using the observed proportions from only the three lower dose levels.

For the three lower experimental dose levels, the administered dose levels (0, 0.001 and 0.01 μg/kg/day) and the liver concentration levels (0, 0.54 and 5.1 ppb) are almost directly proportional (linearly related). Any two dose scales that are directly proportional yield equivalent fitted models and estimated VSDs. Thus, for the remainder of this paper only the μg/kg body weight/day scale will be used.

When the very highest experimental dose level is not allowed to dominate the shape of the fitted multistage model for hepatocellular neoplastic nodule or carcinoma in the Kociba *et al.* (1978) study, the non-linearity in the fit of the multistage model to the three lower experimental dose levels is much more consistent with the observed non-linearity in the fitted multistage model for the same carcinogenic response in the NTP (1982) study (Fig. 7). Almost contradictory shapes arise (Fig. 8) if the high-dose data in the former study is allowed to dominate the fit.

Estimating or bounding the VSD

The dose level corresponding to an increase of 0.000001 in the fitted model's value for probability of a specified response is the fitted model's estimate of the VSD. This VSD estimate is the estimate that is most consistent with the presumed model family. The corresponding estimated VSD for an increase of 0.000001 in the probability of hepatocellular neoplastic nodule or carcinoma in a female rat is 0.14 ng/kg body weight/day in the diet using the Kociba *et al.* (1978) data and 0.13 ng/kg body weight/day in the diet using the NTP (1982) data. The corresponding estimated VSDs for male rats are higher, as are those for hepatocellular carcinoma alone.

In addition to the model family's best estimate of the VSD, 95% upper and lower confidence limits (or other percentages) on that VSD can also be constructed. The purpose of these limits is to bound the true VSD and not to estimate the VSD. The difference

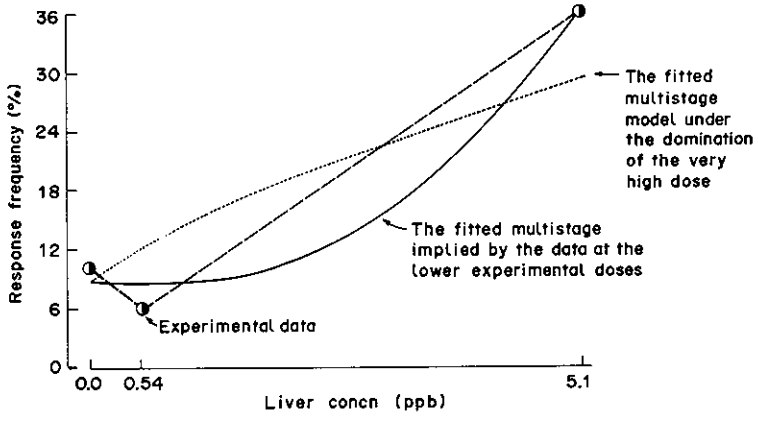

Fig. 6. The fit of the multistage model and the estimated virtually safe dose (VSD) for the Kociba *et al.* (1978) data on hepatocellular neoplastic nodule or carcinoma in female rats are substantially different when the domination of the very high dose is retained (VSD 0.00008 ng/kg body weight/day) and when it is removed (VSD 0.14 ng/kg/day).

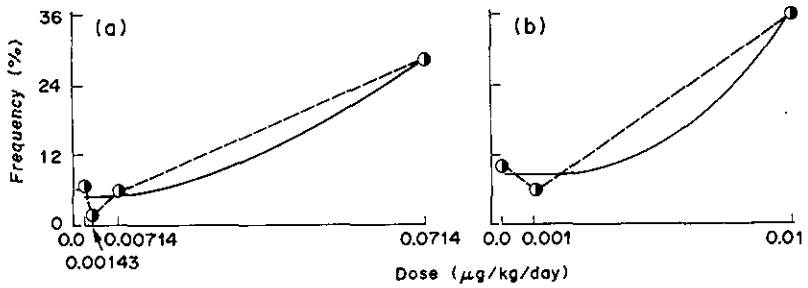

Fig. 7. The shape of the fitted multistage model for the experimental data on hepatocellular neoplastic nodule or carcinoma in female rats in the NTP (1982) study (a) supports the non-linearity in the shape of the fitted multistage model for the experimental data on hepatocellular neoplastic nodule or carcinoma in female rats at the lower doses in the Kociba *et al.* (1978) study (b).

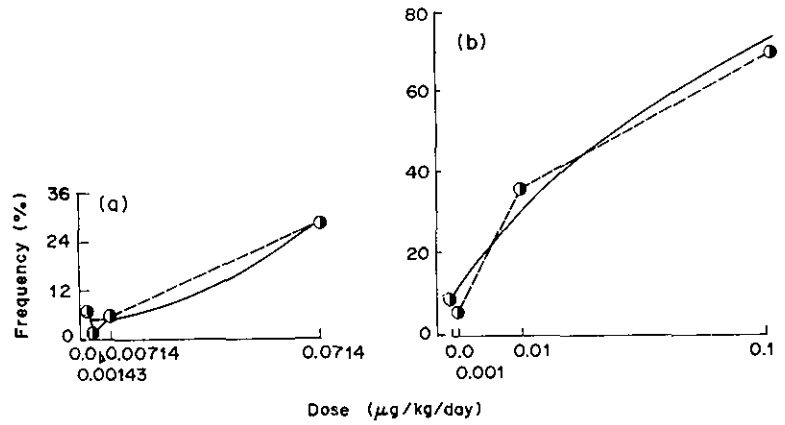

Fig. 8. The shape of the fitted multistage model for the experimental data on hepatocellular neoplastic nodule or carcinoma in female rats in the NTP (1982) study (Fig. a) does not support the linearity in the shape of the fitted multistage model for the experimental data on hepatocellular neoplastic nodule or carcinoma in female rats in the Kociba *et al.* (1978) study (Fig. b), if the data at the very high dose are allowed to dominate the fit.

between the fitted model values and the bounds on the VSD may be very large.

There is more than one way to construct a bound or confidence limit, and, even for a particular model family and a given set of experimental data, the constructed bounds can differ by at least a few orders of magnitude. Empirical studies of the usual confidence limit procedure for the multistage model have documented that its lower bounds on the VSD are frequently much smaller than the true value of the VSD when the underlying dose-response relationship is sub-linear (i.e. convex or upward curving as in Fig. 7).

Table 5 shows the differences between the fitted model values for the VSD and the multistage model's usual 95% lower confidence limits on the VSD. (These limits are not calculated using the fitted multistage model values but arise from the so-called 'linearized multistage model', which is something of a misnomer since a 'model' usually refers to a best fit as opposed to a collection of worst case bounds. Crump (1981) contains a detailed discussion of this

Table 5. The difference between the fitted three-stage multistage model value for a VSD and the usual 95% lower confidence limit on the VSD in μg/kg body weight/day

Response	Highest dose level used in modelling	Fitted model value for VSD (A)	95% Lower confidence limit for VSD (B)	Ratio A/B
Hepatocellular neoplastic nodule or carcinoma	0.01	1.4×10^{-4}	2.2×10^{-8}	6400
Hepatocellular carcinoma	0.01	3.1×10^{-4}	1.2×10^{-7}	2600

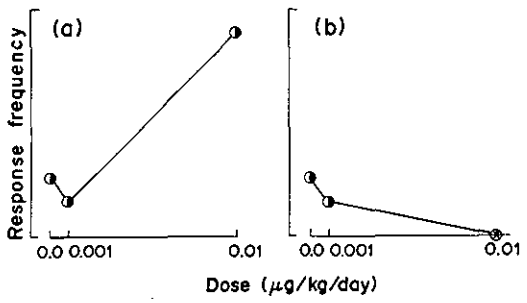

Fig. 9. The procedure used to construct a 95% lower confidence limit on the virtually safe dose is very *non-responsive* to the experimental data. The lower confidence limits here are very far below the lowest experimental dose even if the experimental data would have indicated a decreasing dose-response relationship. Thus for the experimental data on hepatocellular neoplastic nodule or carcinoma in the Kociba *et al.* (1978) study (Fig. a), the 95% lower confidence limit on VSD is 2.2×10^{-8} μg/kg body weight/day, which is 45,500 times smaller than the lowest non-zero experimental dose (0.001 μg/kg/day); the corresponding figure for a hypothetical continued decrease at 0.01 μg/kg/day (Fig. b) is 3.7×10^{-7} μg/kg/day (2700 times smaller than the 0.001 μg/kg/day dose).

bounding procedure.) The 95% lower confidence limit on the VSD is 6400 times smaller than the fitted model value for the VSD when the 'response' is defined as hepatocellular neoplastic nodule or carcinoma (or 2600 times smaller when the 'response' is hepatocellular carcinoma).

Empirical studies have shown that the multistage model's usual 95% lower confidence limits on the VSD are not very responsive to the experimental data, in the sense that very different experimental outcomes still lead to very similar lower bounds on the VSD. For example, as indicated in Fig. 9, if the VSD is bounded using the response frequencies for hepatocellular neoplastic nodule or carcinoma (namely 9/86, 3/50 and 18/50—i.e. 10.5, 6 and 36%—at 0.0, 0.001 and 0.01 μg/kg body weight/day, respectively), then the 95% lower confidence limit on the VSD is 2.2×10^{-8} μg/kg/day. If, by contrast, the response frequencies had been 9/86, 3/50 and 0/50 (i.e. 10.5, 6 and 0%—a decreasing response rate with increasing dose), then the 95% lower confidence limit would have been 3.7×10^{-7} μg/kg/day. Thus, two totally different experimental outcomes (one with an apparently increasing risk near 0.01 and one with an apparently decreasing risk over all experimental dose levels) lead to comparable lower bounds on the VSD. Both lower bounds are at least 2700 times smaller than the lowest non-zero experimental dose level (0.001 μg/kg/day). This non-responsiveness (which has misleadingly been referred to as 'stability') is not due to the sample sizes either. If the response rates had been 900/8600, 300/5000, 1800/5000, corresponding to a 100-fold increase in sample size, the 95% lower confidence limit would still have been only 1.2×10^{-6} μg/kg/day, which is approximately 830 times smaller than the lowest non-zero experimental dose level (0.001 μg/kg/day).

The same procedure which is used to generate the linearized multistage model (the upper bounds on the added probability of the specified response), from which the 95% lower confidence limits on the VSD are calculated, can also be used to generate lower bounds on the added probability of the specified response, from which 95% upper confidence limits on the VSD can be calculated. As is evident in Fig. 10, the corresponding upper and lower bounds on the added probability of the hepatocellular neoplastic nodule or carcinoma in a female rat computed using the Kociba *et al.* (1978) study data are very far apart and very different from the estimates implied by the best fit of the multistage model.

The true VSD can be considerably larger than its lower bound. If the true dose-response relationship is one of the curves in the model family, then the fitted model value is the most likely value for the VSD and the values near the fitted model value are much more likely to be the true value than are the values farther away.

Incorporating time-to-response information

A scientific assessment of the health effects of TCDD should include risk characterizations that reflect the time to the occurrence of the specified responses. Such time-to-response information is available from the individual animal records for the Kociba *et al.* (1978) study in the form of the presence or absence of the specified response at the time the experimental animal was observed (when it died, was killed or was found dead).

Plots of the Kaplan–Meier non-parametric estimates of the frequencies of hepatocellular carcinoma and hepatocellular neoplastic nodule or carcinoma at the time of death are shown in Fig. 11. These plots suggest that, if a rat does develop a carcinogenic response, it is likely to do so only after a considerable portion of an average rat lifetime has passed and that, on average, only a very small fraction of an average rat lifetime would be affected by the occurrence of a carcinogenic response. Furthermore, the more severe

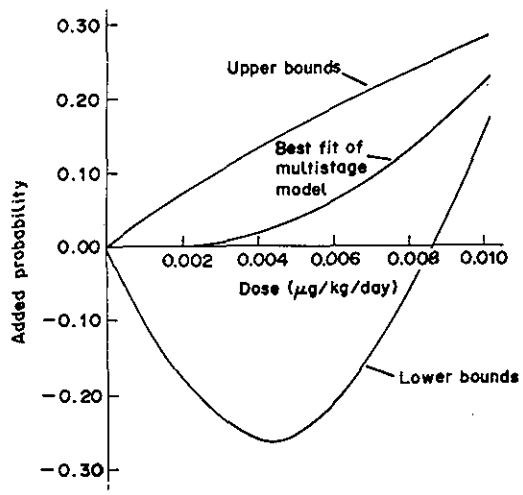

Fig. 10. The upper and lower bounds on the added probability of a hepatocellular neoplastic nodule or carcinoma developing in a female rat for the lower experimental doses in the Kociba *et al.* (1978) study are very far apart and quite different from the values implied by the best fit of the model.

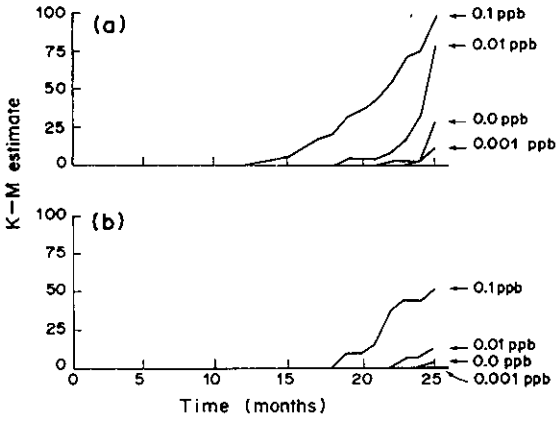

Fig. 11. Kaplan–Meier (K–M) non-parametric estimates of the frequency of hepatocellular neoplastic nodule or carcinoma (Fig. a) and of hepatocellular carcinoma (Fig. b) at the time of death in the Kociba *et al.* (1978) study show that the observed carcinogenic responses took a long time to occur, and the time to response increased markedly as the dose level decreased. The time to response also increased as the severity of the response increased.

the carcinogenic response, the later it is likely to occur.

The multistage model utilizes only the quantal response data and essentially ignores the time-to-response information even though the 'number of animals at risk' in the quantal response data is sometimes 'adjusted' for interim kills and early mortality. Unfortunately, quantal response data refers only to one point in time (approximately 25 months in the Kociba *et al.* (1978) study) and takes into account only the proportion of animals developing the specified response by that one time. The quantal response data does not reflect *when* a response occurs during a period.

The multistage model has the form:

$$P(d) = 1 - \exp\left[-\sum_{i=0}^{I} \alpha_i d^i\right]$$

where $P(d)$ is the probability at dose d that the specified response will occur by the end of the experiment. The $\alpha_0, \ldots, \alpha_I$ are unknown constants to the estimated from the experimental data.

The Hartley–Sielken time-to-response model (Hartley & Sielken, 1977) is one of the generalizations of the multistage model to include the information on the time to occurrence of a carcinogenic response. It has the form:

$$P(t;d) = 1 - \exp\left\{-\left[\sum_{i=0}^{I} \alpha_i d^i\right]\left[\sum_{j=1}^{J} \beta_j t^j\right]\right\} \ldots. \quad (1)$$

where $P(t;d)$ is the probability at dose d that the specified response will occur by time t. This model, which has a product from hazard rate, can itself be generalized to:

$$P(t;d) = 1 \exp\left[-\sum_{i=0}^{I}\sum_{j=1}^{J} \kappa_{ij} d^i t^j\right] \ldots. \quad (2)$$

which has a non-product form hazard rate (Society of Toxicology ED_{01} Task Force, 1981). A latency period parameter, say LP, can be incorporated into these models by replacing t by t-LP.

Both of these time-to-response models (equations 1 and 2) were fitted to the Kociba *et al.* (1978) time-to-response data with the response defined to either

(i) death from any cause,
(ii) the existence of a hepatocellular neoplastic nodule or carcinoma, or
(iii) the existence of a hepatocellular carcinoma.

Figure 12 shows the corresponding plots of the estimated mean response-free periods (i.e. the estimated amount of time that a female rat is expected to be without the specified response). For the 25-month observational period in the Kociba study, the estimated mean response-free periods hardly decrease in the probable area of public health interest.

The mean response-free period can be used to define alternative maximum acceptable doses that may be more relevant than the VSD, since the latter ignores the time of the response. The alternative considered here is the mean free dose (MFD) which is illustrated in Fig. 13 and is the dose level corresponding to a specified decrease in the mean response-free period. Table 6 indicates the estimated MFDs when the specified decrease in the female rat's mean response-free period for the 25-month experimental period is the same proportional decrease as 1 month, 1 week, 1 day and 1 hour in a 70-year period. These risk characterizations are summarized in Table 7. The estimated MFD corresponding to a 1-hour reduction in a 70-year period ranges between approximately 0.1 and 1 ng/kg body weight/day, depending upon the specified response, whether a product or non-product form of the hazard rate is used, whether

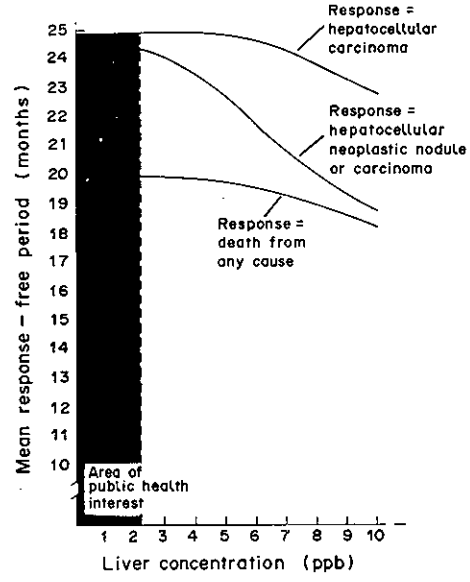

Fig. 12. Risk can be characterized in terms of the expected length of time without a specified carcinogenic response. For 2,3,7,8-tetrachlorodibenzo-*p*-dioxin, the mean response-free periods for female Sprague–Dawley rats during a 25-month experiment hardly decrease until the dose levels become relatively large.

Table 6. Fitted time-to-response model values for the virtually safe dose (VSD) and mean free dose (MFD) utilizing the individual female rat time-to-response information from the Kociba *et al.* (1978) study on 2,3,7,8-tetrachlorodibenzo-*p*-dioxin

Response	Dose scale	Form of the Hartley–Sielken time-to-response model		Estimated dose levels (ng/kg body weight/day)				
					MFD			
		Hazard rate	Latency period	VSD (0.000001)	Month	Wk	Day	Hr
Hepatocellular carcinoma	Liver concn	Non-product	Included	0.54	6.2	4.3	2.7	1.2
			Excluded	0.38	5.5	3.8	2.3	0.90
		Product	Included	0.59	6.5	4.5	2.8	1.3
			Excluded	0.26	5.0	3.1	1.6	0.57
	Administered	Non-product	Included	0.55	6.2	4.3	2.6	1.2
			Excluded	0.19	4.5	2.8	1.4	0.45
		Product	Included	0.59	6.7	4.7	2.9	1.3
			Excluded	0.29	5.3	3.3	1.8	0.63
Hepatocellular neoplastic nodule or carcinoma	Liver concn	Non-product	Included	0.12	2.5	1.6	0.82	0.29
			Excluded	0.37	3.1	2.1	1.3	0.59
		Product	Included	0.32	3.4	2.4	1.5	0.67
			Excluded	0.12	2.2	1.4	0.71	0.25
	Administered	Non-product	Included	0.11	2.4	1.5	0.76	0.27
			Excluded	0.12	2.1	1.3	0.66	0.23
		Product	Included	0.13	2.9	1.8	0.96	0.34
			Excluded	0.12	2.2	1.3	0.70	0.25
Death from any cause	Liver concn	Non-product	Excluded	0.15	4.5	2.8	1.4	0.25
		Product	Excluded	0.32	4.8	3.0	1.6	0.55
	Administered	Non-product	Excluded	0.07	3.2	1.6	0.59	0.12
		Product	Excluded	0.33	5.0	3.1	1.6	0.57

a latency period is included or excluded, and whether the administered dose or liver concentration scale of dose is used. The estimated MFD ranges between 0.6 and 3 ng/kg/day for a 1-day reduction in a 70-year period and between 1 and 5 ng/kg/day for a 1-week reduction in a 70-year period.

Conclusions

In both the Kociba *et al.* (1978) and NTP (1982) studies, there were no significant increases in carcinogenic response at the lower experimental doses. The only significant increases were at the very highest experimental dose, except in the case of hepatocellular neoplastic nodules in female rats; these were also significantly increased at the second highest dose level.

In neither study does the observed incidence of hepatocellular neoplastic nodules and/or carcinomas appear to be linearly related to the dose level.

The Kaplan–Meier estimates of the probability of a female rat developing a hepatocellular neoplastic nodule or carcinoma in the Kociba study suggest that the dose-response relationship resembles one with a saturation-like phenomenon occurring at the highest experimental dose level. Since it is impossible for the fitted multistage model to portray both this type of phenomenon and the non-linearity observed at the

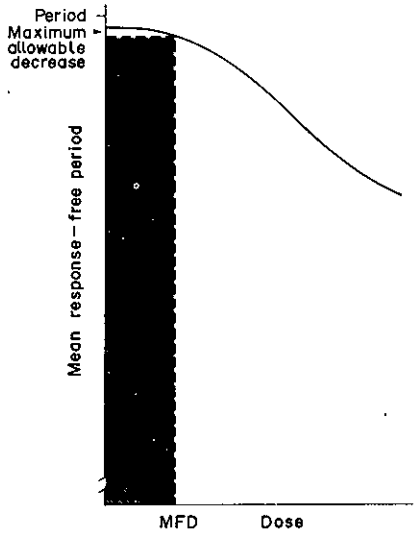

Fig. 13. A maximum allowable decrease in the mean response-free period can define a maximum allowable dose level (the mean free dose or MFD).

Table 7. Estimates of a maximum allowable dose of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in the diet for a female Sprague–Dawley rat

Dose estimated	Range (ng/kg body weight/day)
Virtually safe dose (VSD) corresponding to an increased probability of the specified response of 0.000001 (one-in-a-million)	0.07–0.60
Mean free dose (MFD) corresponding to a decrease in the mean response-free period proportional to	
1 month in a 70-yr period	2–7
1 wk in a 70-yr period	1–5
1 day in a 70-yr period	0.6–3
1 hr in a 70-yr period	0.1–1

The ranges of estimates were obtained:
(i) using the experimental data from the three lowest dose levels in the Kociba *et al.* (1978) study;
(ii) using Hartley–Sielken time-to-response models with either a product form or non-product form hazard rate and either including or excluding a latency period;
(iii) using either the administered or liver concentration dose scale;
(iv) defining the response of concern as either hepatocellular neoplastic nodule or carcinoma, hepatocellular carcinoma, or death.

lower experimental doses, and since including the very highest experimental dose level makes the fit of the multistage model very non-responsive to the situation observed at the lowest non-zero experimental dose; a better indication of the dose-response relationship in the region of the lower experimental dose levels should be obtained by fitting the multistage model only to the data at those levels.

When the very highest experimental dose level is not allowed to dominate the shape of the fitted multistage model for hepatocellular neoplastic nodules or carcinomas in the Kociba study, the non-linearity in the fit of the multistage model to the lower three experimental dose levels is much more consistent with the observed non-linearity in the fitted multistage model for the same carcinogenic response in the NTP study. The corresponding estimated VSD for an increase of 0.000001 (one in a million) in the probability of hepatocellular neoplastic nodule or carcinoma in a female rat is 0.14 ng/kg body weight/day in the diet using the Kociba data and 0.13 ng/kg body weight/day in the diet using the NTP data. The corresponding estimated VSDs for male rats are higher, as are those for hepatocellular carcinoma alone.

The procedure used in the linearized multistage model to construct 95% confidence limits on the VSD is very non-responsive to the experimental data, and these limits are very far away from the fitted model values.

Hepatocellular nodules or carcinomas were observed only very late in the Kociba *et al.* (1978) experiment, particularly at the three lower dose levels. When the time at which a hepatocellular neoplastic nodule or carcinoma might occur is incorporated into the dose-response modelling and risk characterization, the estimated VSD is in the approximate range of 0.1–0.6 ng/kg body weight/day in the diet of a female rat, and the estimated MFD corresponding to a reduction in the expected amount of time without the carcinogenic response proportional to 1 week in 70 years is in the range of 1–5 ng/kg body weight/day in the diet of a female rat.

Kimbrough *et al.* (1984) estimated the human daily TCDD intake corresponding to 1 ppb in residential soil as 636.5 fg/kg body weight/day for a person weighing 70 kg. They then combined this estimated intake with an essentially linear dose-response model to conclude that 1 ppb represented "... a reasonable level at which to begin consideration of action to limit human exposures ..." and "... a level of concern." However, these authors also stated that the "model was used on a hypothetical basis, and the cancer risk for TCDD should be reevaluated as the data base enlarges." The data base has been enlarged, and the dose-response model has been re-evaluated. The data now suggest that the 636.5 fg/kg body weight/day is more than 150 times smaller than the estimated VSD for an increase of one in a million in the probability of hepatocellular neoplastic nodule or carcinoma in a female rat and more than 1500 times smaller than the estimated MFD corresponding to a reduction in the expected amount of time without this carcinogenic response proportional to 1 week in 70 years.

REFERENCES

Crump K. S. (1981). An improved procedure for low-dose carcinogenic risk assessment from animal data. *J. envir. Path. Toxicol.* **5,** 339.

Environmental Protection Agency (1980). Risk Assessement on 2,4,5-Trichlorophenoxyacetic acid (2,4,5-T), 2,4,5-Trichlorophenoxypropionic acid (Silvex), and 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD). EPA 600/6-81-003, NTIS PB81-234-825.

Hartley H. O. & Sielken R. L., Jr (1977). Estimation of 'safe doses' in carcinogenic experiments. *Biometrics* **33,** 1.

Kimbrough R. D., Falk H. & Stehr P. (1984). Health implications of 2,3,7,8-tetrachlorodibenzodioxin (TCDD) contamination of residential soil. *J. Toxicol. envir. Hlth* **14,** 47.

Kociba R. J., Keyes D. G., Beyer J. E., Carreon R. M., Wade E. E., Dittenber D. A., Kalnins R. P., Frauson L. F., Park D. N., Barnard S. D., Hummel R. A. & Humiston C. G. (1978). Results of a two year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in rats. *Toxic. appl. Pharmac.* **46,** 279.

National Toxicity Program (1982). Carcinogenesis Bioassay of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (CAS no. 1746–01–6) in Osborne–Mendel Rats and B6C3F1 Mice (Gavage Study). NTP Technical Report Series, Issue 209, 195.

Society of Toxicology ED_{01} Task Force (1981). Re-examination of the ED_{01} study. *Fund. appl. Toxicol.* **1,** 26.