NATIONAL INSTITUTE OF PUBLIC HEALTH AND ENVIRONMENTAL
PROTECTION (RIVM) BILTHOVEN
THE NETHERLANDS

Report nr. 659101003

EVALUATION OF QUANTITATIVE METHODS FOR THE DETERMINATION OF THE ACCEPTABLE DAILY INTAKE

M.J. Zeilmaker, W. Slob, E.H.J.M. Jansen and
F.X.R. van Leeuwen

maart 1995

This investigation was performed within project 659101 on behalf and for the account of the Directorate General for Environmental Protection, the Directorate for Chemicals, Safety and Radiation Protection.

MAILING LIST

01-05 Directoraat-Generaal Milieubeheer, Directie Stoffen, Veiligheid en Straling
06 Plv. Directeur-Generaal Milieubeheer, Dr ir B.C.J. Zoeteman
07 Prof. dr C.J. van Leeuwen, DGM/SVS
08 Dr K.R. Krijgsheld, DGM/LE
09 Dr H. Roelfzema, Directie Voeding en Veiligheid van Produkten
10 Dr ir P.C. Bragt, Hoofdinspectie Gezondheidsbescherming
11 T. Staarink, Hoofdinspectie Gezondheidsbescherming
12 Dr ir G. Kleter, Hoofdinspectie Gezondheidsbescherming
13 Directie RIVM
14 Hoofd Bureau Voorlichting en Public Relations
15 Depot Nederlandse Publikaties en Nederlandse Bibliografie
16 Dr ir G. de Mik, Directeur Sector 6, Stoffen en Risico's
17 Dr W.H. Könemann, Hoofd Laboratorium voor Toxicologie
18 Dr J.E.T. Moen, Waarnemend Hoofd Adviescentrum voor Toxicologie
19-49 Leden Dwarsverband Dosis Effect Modellering
50 Dr P.W. Wester, PAT
51 Drs. A.G.A.C. Knaap, ACT
52 Ing. P.J.C.M. Jansen, ACT
53 Dr G.J.A. Speijers, TOX
54 Dr A. van der Giessen, Hoofd Centrum voor Wiskundige Methoden
55 Dr J. Hoekstra, CWM
56-59 Auteurs
60-61 Bibliotheek RIVM
62 Bureau Projecten- en Rapportenregistratie
63-85 Reserve exemplaren

LIST OF CONTENTS pag.

MAILING LIST ii
LIST OF CONTENTS iii
SUMMARY iv
SAMENVATTING v

1. INTRODUCTION 1

2. TRADITIONAL WAY OF DETERMINING THE ACCEPTABLE DAILY INTAKE 1

2.1 The definition of the No Observed Adverse Effect Level 1
2.2 The Safety Factor Method 2

3. VALIDITY OF THE SAFETY FACTORS 4

3.1 The intraspecies extrapolation factor 4
3.2 The interspecies extrapolation factor 4
3.3 The subchronic => chronic extrapolation factor 4
3.4 The LOAEL => NOAEL extrapolation factor 5

4. LIMITATIONS OF THE CURRENT APPROACH TO ASSESS THE ACCEPTABLE DAILY INTAKE 5

5. QUANTIFYING DOSE-RESPONSE RELATIONSHIPS 8

5.1 Crump's Benchmark Dose 8
5.2 Gaylor's Linear Extrapolation Method 8
5.3 Hoekstra's Bounded-Effect Dose 10
5.4 Dose-Severity Diagrams 10

6. CONCLUSIONS AND DISCUSSION 12

6.1 Comparison of alternative methods 12
6.2 The Non Adverse Effect Level 12
6.3 Benchmark Dose versus NOAEL 13
6.4 Hybrid approach: estimation of threshold 13
6.5 Consequences for experimental design 14
6.6 Validation of safety factors 14

7. REFERENCES 15

8. APPENDICES 17

SUMMARY

In cases where human toxicity data are not available the Acceptable Daily Intake (ADI) is determined by division of the No Observed Adverse Effect Level (NOAEL) for animal toxicity by a safety factor. In this calculation the NOAEL is used as a point estimate for the threshold of animal toxicity.

The way in which the NOAEL is used in calculating the ADI has several limitations. Of these limitations the weakness of the NOAEL as an estimator of the threshold for animal toxicity and the inability of the current approach to assess toxicity beyond the NOAEL form important issues. This report reviews three quantitative methods, e.g. Crump's Benchmark Dose, Gaylor's Extrapolation Method and Hoekstra's Bounded-Effect Dose, on their merits in improving the adequacy of determining the ADI from animal toxicity data.

Contrary to the NOAEL approach, the above mentioned methods take the experimental uncertainty into account in estimating the threshold for animal toxicity and make use of the entire dose-response relationship relation instead of only one of its data points in assessing chemical toxicity. The applicability of these methods, however, greatly depends on the quality of the data set to be analysed. In this context answering the following questions are of particular importance :

1. Do the data allow reasonable characterisation of the dose-response relationship ?
2. Do the data allow the estimation of acceptable toxicity levels *per se* or is extrapolation outside the experimentally observed dose range necessary ?
3. Can *a priori* safe vs. non-safe effect levels be defined ?

Reviewing the literature the following conclusions were drawn. In cases where the data allow the characterisation of a dose-response relationship, Crump's Benchmark Dose is favourable. However, toxicity data generated for regulatory purposes may not suffice for assessing a dose-response curve. Then Hoekstra's "model free" Bounded Effect Dose might be an alternative. When acceptable animal toxicity levels lie outside the experimental dose range Gaylor's Extrapolation Method is more appropriate.

Applying a benchmark dose approach may have consequences for the experimental design of toxicity studies. Replications per dose group are no longer necessary and the series of doses may equal the total number of animals used, the unreplicated design having less chance in allocating doses in the most interesting part of the dose response relationship, i.e. its changing part. The benchmark dose approach may thus lead to a reduction of the number of animals used in toxicity testing.

In all three methods a maximum safe effect level needs to be postulated. Current toxicological knowledge however does not allow for an equivocal threshold between adverse and non-adverse effect levels for most toxic endpoints. As the definition of adverse vs. non-adverse effect levels is a conditio sine qua non for the application of the methods reviewed in this report we therefore suggest to examine the possibilities to define standard levels for various endpoints, to be used for future assessment of ADI's.

SAMENVATTING

In gevallen waar humane toxiciteitsgegevens niet beschikbaar zijn wordt de "Acceptable Daily Intake" (ADI) berekend door deling van de "No Observed Adverse Effect Level" (NOAEL) in proefdieren door een veiligheidsfactor. In deze berekening wordt de NOAEL gebruikt als een puntschatting voor de drempelwaarde voor chemische toxiciteit.

De manier waarop de NOAEL gebruikt wordt voor het berekenen van de ADI heeft echter enkele beperkingen. Enkele van deze beperkingen zijn de zwakte van de NOAEL als schatter voor de toxiciteitsdrempel en de onmogelijkheid om de toxiciteit van blootstellingen groter dan de ADI te kunnen beoordelen. Dit rapport evalueert drie kwantitatieve methoden, te weten Crump's "Benchmark" Dosis, Gaylor's Extrapolatie Methode en Hoekstra's "Bounded-Effect" Dosis, op hun meerwaarde ten opzichte van de NOAEL methode ter bepaling van de ADI

In tegenstelling tot de NOAEL methode betrekken de genoemde methoden de gemeten experimentele variatie bij het bepalen van de toxiciteitsdrempel. Verder betrekken deze methoden alle experimentele waarnemingen bij het vaststellen van de toxiciteitsdrempel. Bij de traditionele manier methode om de ADI vast te stellen worden daarentegen slechts de waarnemingen bij één van de toegediende doses, nl. de NOAEL, gebruikt.

De toepasbaarheid van de genoemde methoden als alternatief voor de NOAEL methode wordt voor een groot deel bepaald door de kwaliteit van de te analyseren waarnemingen. In het bijzonder zijn daarbij de volgende punten van belang :

1. In hoeverre laten de waarnemingen een kwantitatieve dosis-effect analyse toe ?
2. Kan de drempelwaarde voor toxiciteit uit de waarneminegn zelf afgeleid worden of is hiervoor extrapolatie buiten het waarnemingsgebied nodig ?
3. Kunnen *a priori* veilige cq. niet-veilige effect niveau's aangegeven worden ?

Wat betreft de toepasbaarheid van genoemde methoden kunnen de volgende conclusies getrokken worden. In gevallen waar de waarnemingen een kwantitatieve dosis-effect analyse toelaten heeft Crump's "Benchmark" Dosis methode de voorkeur boven de NOAEL methode. Toxiciteitsgegevens die in het kader van regelgeving beschikbaar zijn lenen zich in het algemeen echter niet voor een gedegen dosis-effect analyse. In dergelijke gevallen zou Hoekstra's "Bounded Effect" Dosis als alternatief voor de NOAEL methode gebruikt kunnen worden. In gevallen waar extrapolatie buiten het waarnemingsgebied noodzakelijk is voor het vastellen van toxiciteitsdrempels cq. veilige blootstellingsniveau's verdient Gaylor's Extrapolatie Methode de voorkeur.

Het toepassen van de "Benchmark" Dosis methode kan belangrijke consequenties hebben voor het experimentele protocol dat toegepast wordt bij het vastellen van veilige blootstellingsnivau's. Toepassing van deze methode maakt het gebruik van herhaalde waarnemingen per toegediende dosis overbodig. Deze eigenschap maakt een efficiënter gebruik van proefdieren bij het vaststellen van toxiciteitsdrempels mogelijk. Verwacht mag worden dat deze methode daarom tot een aanzienlijke besparing van het proefdiergebruik in het toxicologisch onderzoek kan leiden.

Voor de toepassing van zowel Crump's "Benchmark" Dosis, Hoekstra's "Bounded Effect" Dosis en Gaylor's Extrapolatie Methode is kennis van veilige vs. niet-veilige effect niveau's noodzakelijk. Deze kennis kan veelal niet uit het huidige toxiciteitsondezoek verkregen worden. Aan het definiëren van veilige vs. niet-veilige effect niveau's voor toxiciteitsparameters dient daarom de hoogste prioriteit gegeven te worden.

1. INTRODUCTION

Central to the setting of safe oral human exposure levels (Acceptable Daily Intake (ADI)) is the determination of the exposure level which does not induce adverse health effects after chronic exposure.[1,2] Usually this exposure level cannot be determined by direct experimentation in humans. If suitable epidemiological studies on dose-effect relationships are not available safe human exposure levels therefore are extrapolated from dose-response studies in experimental animals.

In extrapolating oral toxicity data of non-carcinogens and non-genotoxic, carcinogenic chemicals from animals to man it is assumed that such compounds exhibit a threshold in inducing toxicity, implying that an exposure level can be defined below which no adverse health effects occur. Consequently, safe exposure can be defined as an exposure level below which adverse health effects are unlikely to occur (Johannsen, 1990; Ecobichon, 1992).

In current practice the ADI is determined by the safety factor method. In this method the ADI is calculated as the ratio of the No Observed Adverse Effect Level (NOAEL) as observed for animal toxicity and a safety factor. In this calculation the NOAEL is used as an estimate for the toxicity threshold. A safety factor is applied to compensate inaccuracies in inter- and intraspecies differences in chemical sensitivity. The safety factor method however is prone to several uncertainties and limitations. One of these limitations concerns the inaccuracy of the use of the NOAEL as an estimate for the toxicity threshold and the inability of the safety factor method to characterise the complete dose-response relationship of chemical toxicity, a feature severely impairing its application in actual risk assessment procedures. In the last decade several methods have been proposed for the improvement of the safety factor method on these points. This report reviews the current procedure in determining the ADI (safety factor method), its experimental validation as well as the methods suggested as its improvements in determining the toxicity threshold and in incorporating the complete dose-response characteristics of chemical toxicity in the derivation of the ADI.

2. TRADITIONAL WAY OF DETERMINING THE ACCEPTABLE DAILY INTAKE

2.1 The No Observed Adverse Effect Level

As mentioned in the introduction a safe human exposure level is extrapolated from a NOAEL observed in laboratory animals. Usually the NOAEL is derived by pairwise comparison of the mean/median (adverse) toxic effect levels in exposed vs. control

---

[1]Adverse effects are considered as effects "leading to functional impairment and/or the induction of pathological lesions which may affect the performance of the whole organism, or which reduce an organism's ability to respond adequately to additional challenge" (EPA, 1980)

[2]In EPA terminology the definition of the ADI equals the definition of the Chronic Reference Dose (RfD), the RfD being defined as "a quantitative estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk, i.e. without suffering significant adverse health effects, during a lifetime (EPA, 1993).

animals. The NOAEL is then defined as the highest experimental dose which does not induce a statistically significant change in the toxic response (see Fig 1.).

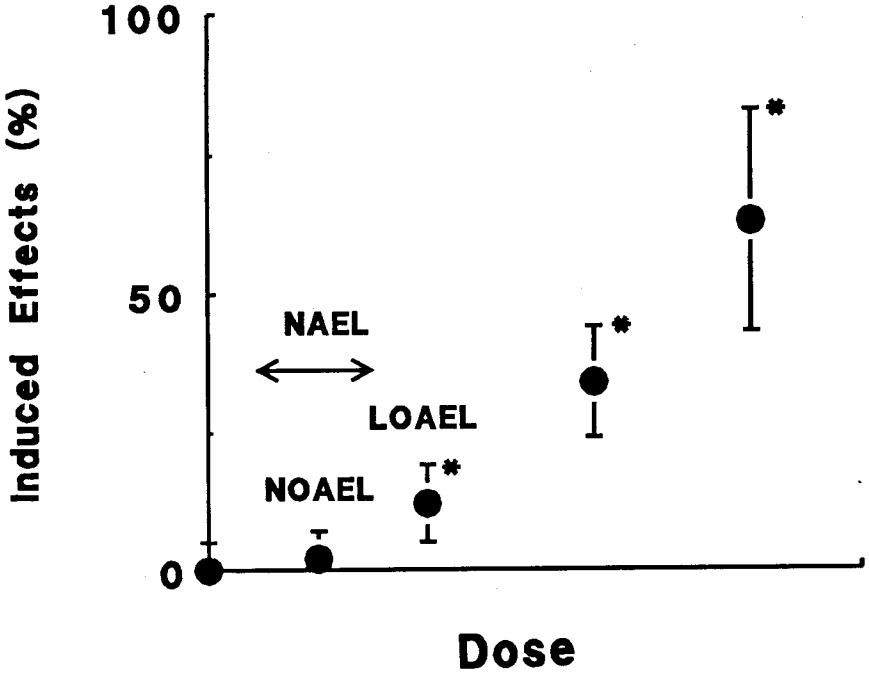

Fig. 1 The No Observed Adverse Effect Level in a hypothetical dose-response relationship existing of five doses

NOAEL : No Observed Adverse Effect Level
LOAEL : Lowest Observed Adverse Effect Level
NAEL : No Adverse Affect Level (threshold)

* : Differing statistically significantly from controls

2.2 The Safety Factor Method

The NOAEL as determined in chronic animal experiments serves as the basis for deriving the ADI. In this extrapolation the ADI is calculated by dividing the $NOAEL_{animal}$ through a safety factor :

$$ADI = \frac{NOAEL_{animal}}{SF_1 * SF_2 * SF_3 * SF_4} \quad (1)$$

with :

SF_1 safety factor encompassing interindividual differences in chemical sensitivity

SF_2 safety factor encompassing interspecies differences in chemical sensitivity
SF_3 safety factor encompassing the extrapolation of a less than chronic $NOAEL_{animal}$ to a chronic $NOAEL_{animal}$
SF_4 safety factor encompassing the extrapolation of a $LOAEL_{animal}$ to a $NOAEL_{animal}$

As the safety factor approach lends itself to gross generalisation guidelines for its standardisation have been proposed (see Table 1 for a characteristic example of such a proposal, Dourson and Stara, 1983; Barnes and Dourson, 1988; Johannsen, 1990; EPA 1993; van Leeuwen, personal communication). These guidelines all rest on the following assumptions :

1. the chemical sensitivity of the most sensitive human subpopulation is maximally 10 times as high as the average sensitivity of that population
2. average man is maximally 10 times as sensitive as the average animal to chemical insult
3. the chronic $NOAEL_{animal}$ is maximally 10 times lower than the subchronic $NOAEL_{animal}$
4. the chronic $NOAEL_{animal}$ is maximally 10 times lower than the chronic $LOAEL_{animal}$

As can be seen from Table 1 an uncertainty factor of 100 is applied to animal toxicity data providing a $NOAEL_{animal}$ in chronic toxicity studies. Table 1 also indicates that intra- and inter-species differences in chemical sensitivity form important uncertainties in the extrapolation of animal toxicity data to man. Each of these two uncertainties has two components, uncertainties on differences in toxicokinetics and uncertainties on differences in toxicodynamics. This led Renwick et al. (1990; 1991; 1993) to an analysis of intra- and interspecies differences in toxicokinetics and toxicodynamics. Based on a comparison of human and animal physiology and toxicokinetics in a (limited) number of case studies, mainly on food additives and pesticides, he concluded that both the intra- and interspecies extrapolation factors can be subdivided into a factor of 4 for differences in toxicokinetics and 2.5 for differences in toxicodynamics.

In addition to the safety factor it has been suggested to extend the procedure described above with a so called modifying factor (MF, default value 1), defined as "a subjective adjustment factor accounting for any additional scientific information that may affect the value of the ADI (Barnes and Dourson, 1988). This MF, which is numerically set between 1 and 10, covers uncertainties related to the (in)completeness of the data-base, a limited number of species tested, etc. (EPA, 1993). Including a modifying factor would then result in the following definition of the ADI :

$$ADI = \frac{NOAEL_{animal}}{\prod_{i=1}^{4} SF_i * MF} \qquad (2)$$

Table 1. Proposed guidelines for the use of safety factors	
Chronic $NOAEL_{animal}$ known (interspecies extrapolation)	SF_1=10, SF_2=10, SF_3=1,SF_4=1
Chronic $NOAEL_{animal}$ unknown Subchronic $NOAEL_{animal}$ known	SF_1=10, SF_2=10, SF_3=3-10, SF_4=1
Chronic $NOAEL_{animal}$ unknown Chronic $LOAEL_{animal}$ known	SF_1=10,SF_2=10, SF_3=1, SF_4=3-10

3. VALIDITY OF THE SAFETY FACTORS

In a limited number of cases the results of experimental studies have been used to test the validity of the safety factors. The results of these studies are reviewed in this section.

3.1. The intraspecies extrapolation factor

To evaluate intraspecies differences in chemical sensitivity Dourson and Stara determined the ratios between the LD_{50} and the $LD_{0.15}$[3] of 490 chemicals in experimental animals. This ratio can be interpreted as an intraspecies adjustment factor for differences in chemical sensitivity. The frequency distribution of the $LD_{50}/LD_{0.15}$ ratio, which is shown in Fig. 2, indicated that in 451 out of 490 cases studied (92%) the intraspecies adjustment factor was $\leq$ 10 (Dourson and Stara, 1983). The average intraspecies adjustment factor was found to be 2.4. Comparable studies in humans are not available. However, at the biochemical level it is known that intrahuman variation in chemical metabolism may exceed a factor of 10 (see Table 2), a finding being significant in cases where chemical toxicity can directly be related to an underlying biochemical lesion.

3.2. The interspecies extrapolation factor

The only study aimed at validating the interspecies extrapolation factor was performed by Hayes (1967). These investigators compared the acute and chronic toxicity of pesticides in man and rat and found a 1.9- to 100-fold difference in acute toxicity (geometric mean : 11) and a 0.58- to 9.4-fold difference in chronic toxicity (geometric mean : 2.9).

3.3. The Subchronic => Chronic extrapolation factor

When chronic toxicity studies are not available a NOAEL may be estimated from subchronic studies. In this extrapolation a factor of 3-10 is applied for the calculation of

[3]the $LD_{0.15}$ approximately represents the dose deviating three standard deviations from the LD_{50} in a probit analysis

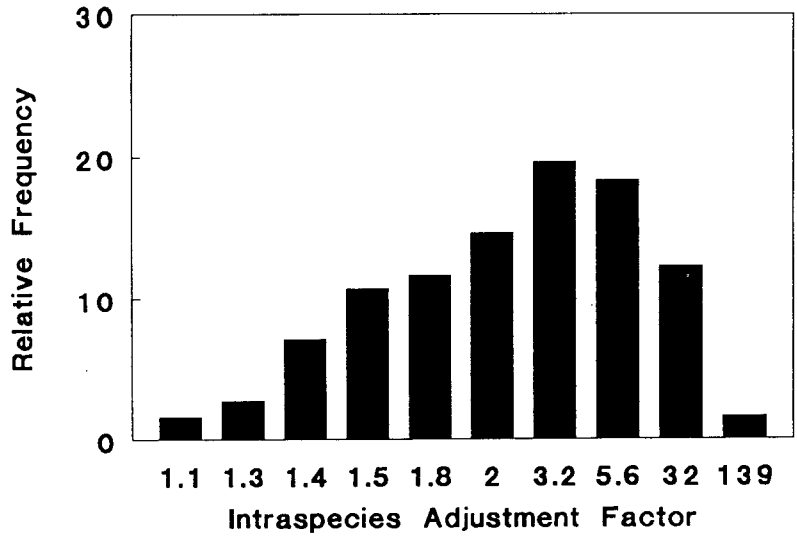

Fig. 2 Frequency distribution of the ratio between the LD_{50} and the $LD_{0.15}$ (intraspecies adjustment factor) as determined for 490 chemicals in experimental animals (Weil et al., 1972, as cited in Dourson and Stara, 1983)

the $NOAEL_{chronic}$ from the $NOAEL_{subchronic}$. To test the validity of this procedure Weil and McCollister determined the frequency distribution of the ratio of these two parameters for a total of 52 chemicals (Weil and McCollister, as quoted in Dourson and Stara, 1983). From this distribution, which is shown in Fig. 3, it can be concluded that in 50 out of 52 cases the ratio of these parameters was ≤ 10. Similar conclusions could be drawn from a study performed by McNamara (1976, all of 41 chemicals studied by these authors showed a ratio $NOAEL_{subchronic}/NOAEL_{chronic} \leq 3$) and Beck (1993, see Fig. 3).

3.4. The LOAEL => NOAEL extrapolation factor

When chronic toxicity studies do not provide a NOAEL its value is approximated by the LOAEL value. As with the subchronic => chronic extrapolation factor an extrapolation factor of 3-10 is applied in this approximation. A comparison of 52 $LOAEL_{chronic}/NOAEL_{chronic}$ ratios revealed a ratio ≤ 5 in 50 out of 52 cases (see Fig. 4).

4 LIMITATIONS OF THE CURRENT APPROACH TO ASSESS THE ADI

Although the method to determine the ADI, as discussed above, has been used for more than thirty years several limitations impair the application of this method to establish accurate safe human exposure levels. Shortly these limitations relate to the following points :

1. The estimation of the toxicity threshold by the NOAEL heavily depends on the experimental protocol chosen (number and interval of experimental doses). Further-

Table 2.

Representative examples of inter-human variability in enzymatic activity

Enzymatic activity	Variation range	Reference
Acetylation of aromatic amines	3.7 - 13 fold	Glowinski, 1978
Debrisoque oxidation	75-85% of the population within a 10-fold range	Idle and Smith, 1979
Epoxide hydroxylation	80-90 % of the population within a 10-fold range	Glatt and Oesch, 1984
Catalase	95% of the population falling within a 0.77-1.44 fold range	Aebi and Wyss, 1978
Superoxide dismutase	2-4 fold	Legge, 1977

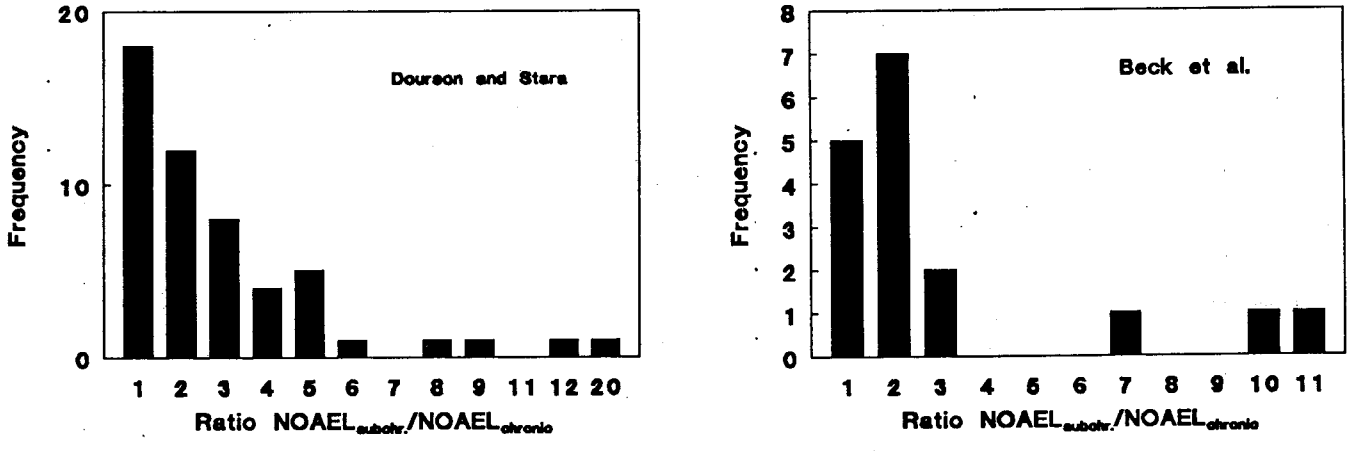

Fig. 3 Frequency distributions of $NOAEL_{subchronic}/NOAEL_{chronic}$ ratios in experimental animals as reported by Dourson and Stara (1983, N=52) and Beck et al. (1993, N=18).

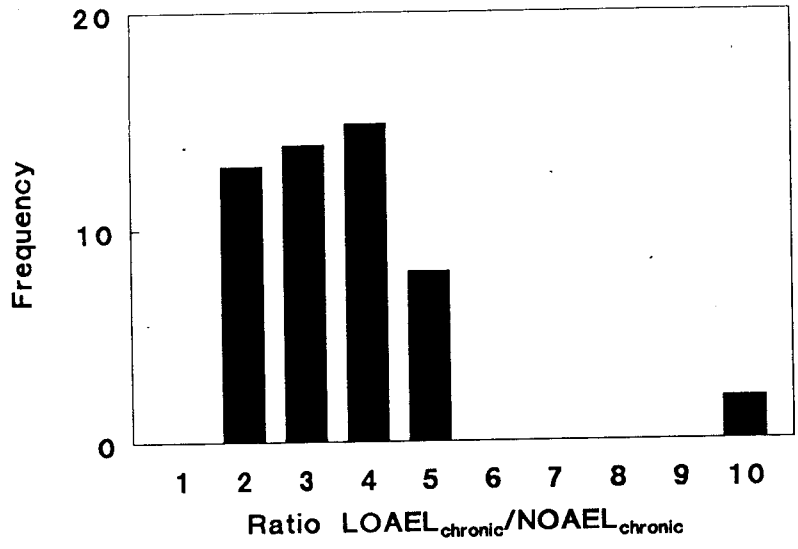

Fig. 4 Frequency distribution of the $LOAEL_{chronic}/NOAEL_{chronic}$ ratio in experimental animals as reported by Dourson and Stara (1983, N=52)

more, poor experimental designs (small sample size, large measurement errors) lead to a higher NOAEL (Hoekstra, 1993).

2. At the NOAEL no statistically significant effect between the treated group and its control is found. However, this does not mean that toxicity at this dose level can be ruled out. An example of this principle is given by Hoekstra (1993). This author cites an experiment investigating the effects of patulin, a mycotoxin, on the renal creatinine clearance in female rats. The results of this experiment showed a NOAEL of 30 mg/l. The estimated ratio of the creatinine clearance at this dose level relative to the control group was found to be 0.93% with the 95% confidence interval ranging from 0.73 to 1.18. Therefore an effect of up to 27% reduction in renal creatinine clearance cannot be ruled out the NOAEL derived from these data. If a reduction in this order of magnitude is considered toxicologically relevant the use of the NOAEL as a basis for deriving the ADI is unwarranted.
3. The determination of the NOAEL may be indicative for the threshold of toxicity, the induction of toxic effects however results from a gradual process obeying distinct dose-response characteristics. Consequently the uncertainty factor method is not suitable to estimate toxic risk at dose levels exceeding the NOAEL.
4. The safety factor method aims at preventing toxic effects after chronic, relatively low, exposure. The ADI derived with this method cannot be extrapolated across time and dose. In this respect it should be noted that :

   "Because in most cases, data are extrapolated from life-time animal studies, the ADI relates to life-time use and provides a margin of safety large enough not to be particularly concerned about short-term use at exposure levels exceeding the ADI providing the average intake over longer periods does not exceed it" (WHO, 1987, as quoted in Renwick and Walker, 1993)

To reduce several of the limitations mentioned it has been proposed to incorporate quantitative methods characterising the dose-response relationship of chemical toxicity into the current approach of deriving safe human exposure levels. These methods all make use of the complete dose-response relationship of chemical toxicity and the experimental error. The basic concepts underlying each of these methods and their merits in improving the accuracy with which the human reference dose is calculated from animal toxicity data are presented and evaluated in the following sections.

5 QUANTIFYING DOSE-RESPONSE RELATIONSHIPS

5.1 Crump's Benchmark Dose

Crump (1984) suggests to use the dose corresponding to a known or postulated toxic effect threshold (for example the increase of the cadmium concentration in the kidneys associated with the onset of organ failure) as the threshold dose for toxicity (benchmark dose). As defined by Crump a benchmark dose then is : "the lower statistical confidence limit for the dose corresponding to a specified, (safe), increase in level of health effect over the background level". The benchmark dose is estimated as follows (see Fig. 5). First a curve is fitted through the data. Next this curve is used to determine the dose level corresponding to the accepted threshold effect level. Finally the experimental error is taken into account to calculate the confidence interval of this dose level. The lower limit of this interval is then taken as the benchmark dose.

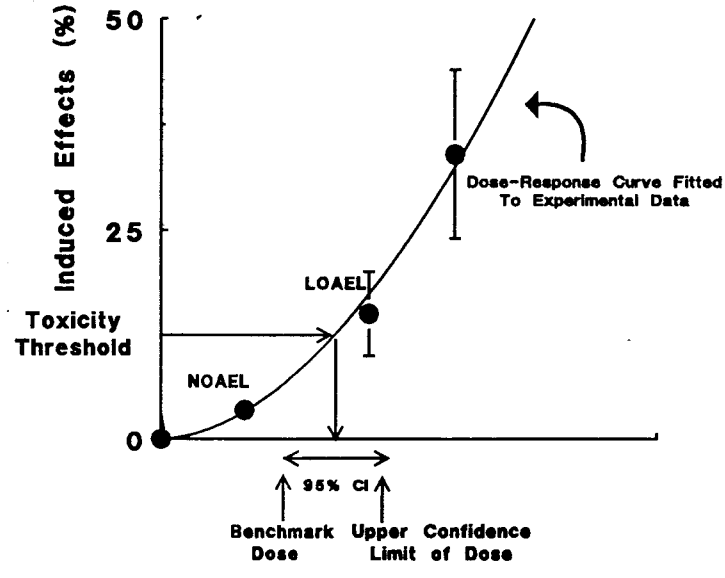

Fig. 5 A graphical presentation of Crump's benchmark dose

5.2 Gaylor's Linear Extrapolation Method

A typical toxicological animal study is not able to discern effect levels smaller than its

level of detection, say a 10% increase of effect over controls. In cases where acceptable effect levels are much lower (e.g. genotoxic or teratological endpoints) the method of Crump becomes unreliable. In particular different regression models, equally fitting the data, result in quite different Benchmark Doses. This problem is related to the low-dose extrapolation used for (genotoxic) carcinogens.

To approach this problem Kimmel and Gaylor (1988) proposed a two-step procedure. First a dose-response curve is fitted to the data and the upper confidence limit of effect, as a function of dose, is calculated. The 10% effect level is intersected with the confidence limit of effect. The associated dose (LED_{10}) is comparable to the Benchmark Dose of Crump. The second step is to linearly extrapolate the latter point to dose zero :

$$Safe\ Dose\ Level_{animal} = \frac{LED_{10}}{F} \qquad (4)$$

with F being the ratio between the 10% effect level and the (arbitrarily set) level considered as acceptable for the endpoint under investigation. The second step, based on the assumption of a linear-dose response relation at low doses, is considered as conservative.

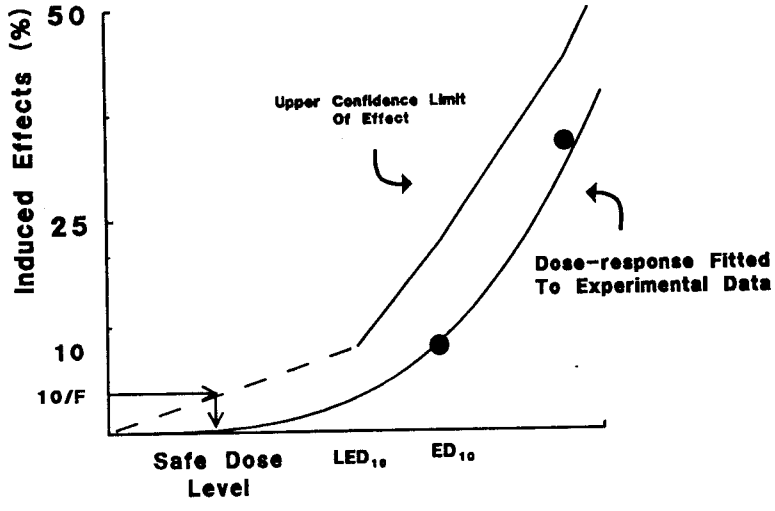

Fig. 6 A graphical presentation of Gaylor's calculation of safe exposure levels

ED_{10} : experimental dose inducing an excess risk of 10%
LED_{10} : dose associated with an UCL_{effect} of 10%
F : factor used to calculate an acceptable, safe, dose level for animal toxicity from an unacceptably high dose level, the value of F, being greater than 1, depending on the severity of the effect to be extrapolated

5.3 Hoekstra's Bounded-Effect Dose

As with Gaylor's extrapolation method Hoekstra's Bounded-Effect Dose method approximates low dose toxicity by a linear function. This approach is conservative given the assumption that the true dose-response relationship is convex below a certain effect level. This effect level, referred to as the bounded-effect level, therefore should be sufficiently low to have confidence in the convexity of the dose-response relation. In investigating 72 ecotoxicological dose-response relationships Hoekstra found in 94% cases that doses corresponding with a 25% change of the mean/median effect in treated groups relative to the control group met the convexity assumption. Hence a 25% change of the toxicological response relative to its control value is suggested as the bounded-effect. The bounded-effect dose then is "The highest dose for which the confidence limits for the toxic effect does not exceed 25%". Once the bounded-effect dose is determined doses corresponding to safe toxicity levels are calculated by interpolating these effect levels on the linear function running from the upper confidence limit of effect at the bounded-effect to zero (see Fig. 7, note that no regression function is fitted to these data).

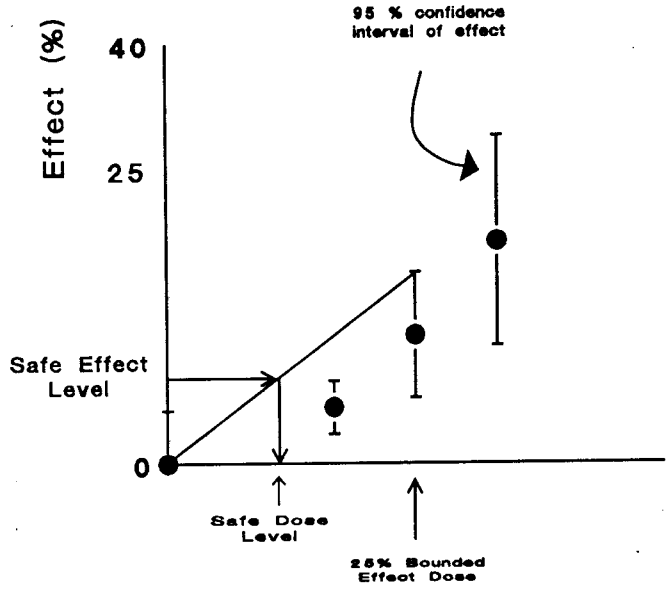

Fig. 7 A graphical presentation of Hoekstra's determination of safe effect levels on the basis of the 25% Bounded Effect Dose

5.4 Dose-Severity Diagrams

The methods used for the calculation of the safe human exposure level presented so far all rely on the evaluation of the most sensitive toxic endpoint in the most sensitive test species. In this calculation the severity of the effect, other then being defined as adverse, only plays a role in qualitative terms. To assess the human reference dose on the basis of the combined pattern of all experimental toxic effect parameters in all species tested and their mutual rating of severity, US EPA has developed a system combining the rating of the severity of effect with the rating of the dose level (potency) of the chemical under investigated (Stara et al., 1987, as cited in McColl, 1990). In this system toxic effect

parameters are scaled according to their severity, Rating Values for severity of effect (RV_e) ranging from 1 (example : enzyme induction) to 10 (example : death)(see Appendix 2). Next to RV_e scaling corresponding Rating Values for dose (RV_d) are calculated from their Minimum Effective Dose (MED), the dose in milligrams per day required to produce a detectable response at a specified level. For practical reasons MED values are then transformed to values from 1 to 10 on a logarithmic scale as follows :

$$RV_d = -1.5 * logMED + 5.5 \qquad (6)$$

Once paires of RV_e and RV_d values are determined a two-dimensional scattergram of both parameters can be plotted (see Fig. 8). For a given RV_e level the point furthest to the right represents the occurrence of an effect in the most sensitive organ or species. Consequently, the line drawn to the right of all the RV_e vs. RV_d data points defines the minimal dose required to produce an effect level in the most sensitive organ or species at various levels of severity. The slope of this line can be extrapolated downward to intersect the X-axis immediately to the left of several NOAEL's for different organs or species. This intercept is then taken as the "overall" NOAEL for the chemical, taking into account effects occurring at all levels of severity and for multiple biological endpoints in different organs and species (McColl, 1990). The "overall" NOAEL may then be taken to calculate the human reference dose by division by the safety factor of 100.

Although the scattergram resulting from this approach gives a comprehensive overview of the combined toxicity data collected for a given compound and of the effects to be expected when the ADI is exceeded, the quantitative analysis is ambiguous. First of all, the interval scale chosen for the levels of RV_e is arbitrary : it implies that the increase in adversity from, say, level 7 to 8 equals the increase from, say, level 2 to 3, which is unwarranted. Second, as objective methods are not available yet, the determination of the "apparent severity slope" completely relies on personal judgement. And third, the derived value of the "maximum NOAEL" does not seem to be much influenced by the data for RV_e values greater than 1, suggesting that the "maximum NOAEL" will not differ much from the hitherto determined "overall" NOAEL (= NOAEL determined for the most sensitive toxic endpoint in the most sensitive species).

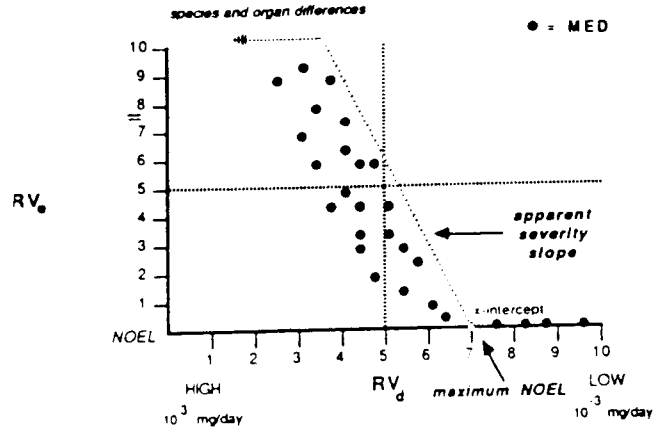

Fig. 8 The estimation of the "overall" NOAEL by Dose-Severity Diagrams (Stara et al.,(1987) as cited in McColl, 1990)

6. CONCLUSIONS AND DISCUSSION

6.1 Comparison of alternative methods.

The use of the NOAEL as a basis for deriving an ADI has been critized in the toxicological literature (see section 4). The alternatives as reviewed in section 5 largely meet these objections, in particular the statistical ones, i.e., the alternative methods all take the experimental uncertainties into account. We will now elaborate on the relative merits of these alternatives and their possible drawbacks.

The method proposed by Crump is straightforward from a statistical / data analytical point of view. Moreover, it may not be applicable in all situations. Firstly, the data must contain sufficient information to facilitate a reliable fitting of a dose-response curve. For example, if only two or three dose groups have been tested, fitting a dose-response curve will not be appropriate due to a high level of uncertainty. In such a situation widely differing curves may fit the data equally well, but result in largely varying benchmark doses. This effect may even be aggrevated when different toxic effects display marked heterogenicity at relative high dose-levels.

A second restriction is that the benchmark dose cannot be much lower than the lowest experimental dose, since extrapolation from fitted regression models is well-known to be unreliable.

The latter restriction is likely to prevail when the reference effect level still considered as harmless is small. The method of Gaylor et al (1980) circumvents this problem by first estimating the dose at which a moderately small effect level (say 10%) might still occur, using the 95% upper confidence level of the regression function. The ensuing extrapolation to lower effect levels, usually resulting in a dose far beyond the experimental range, is done linearly and thus considered as conservative.

As already mentioned, when the data do not reveal a clear dose-reponse relation, fitting a regression function may not be warranted. In that case the approach of Hoekstra may be adopted. This 'model-free' approach does not suffer from the fact that different regression models may give different benchmark doses, which will occur particularly when the number of dose groups is small. On the other hand the assumption that the underlying dose-response relationship is convex below that level of 25 % excess effect over controls was validated by quantal ecotoxicological data (mortality). Whether Hoekstra's method and its criterium for the bounded-effect dose are also valid for toxic endpoints other than mortality as determined in rodent toxicity studies has yet to be determined. Of course in applying Hoekstra's method a price has to be paid : of the three methods evaluated this method is the most conservative and will lead to the lowest value of the 'safe' dose.

In summary, the choice of method should depend on the quality of the data available (is the dose response relation reasonably determined?), and on the magnitude of the acceptable effect level (is the associated dose in- or outside the range of observation?).

6.2 The Non Adverse Effect Level.

Although the alternative methods discussed here do not suffer from the objections against the NOAEL, they are still not common practice. One of the reasons is that they introduce another basic difficulty: an effect level has to be defined that may be considered as negligible or acceptable. Thus, the qualitative problem of deciding what toxicological endpoint is adverse, has been extended with the quantitative, and even more difficult

question: "At which level is an effect (for that endpoint) adverse?" Current toxicological knowledge usually does not provide sufficient basis to quantitate such an effect-threshold indisputably or in an objective manner. The implied arbitrariness might be reason to reject the proposed alternatives, but that would be a fallacy. Both the benchmark dose (or analogues of the other methods) and the NOAEL are affected by arbitrariness. The difference is that the arbitrariness of the benchmark dose is explicit and under the direct control of the risk assessor, whereas the NOAEL, being determined via expert judgement, is not. The power of the alternative approaches is that at the derived benchmark-dose effects of the specified magnitude can be reasonably precluded. The weakness of the NOAEL is that effects of some unknown magnitude cannot be excluded.

Clearly, the question as to which effect-levels are non-adverse cannot be avoided. If it cannot be satisfactorily answered, we can do no other than postulate some value for the time being, for example effect levels usually found at the level of the NOAEL. On the long-term this issue can only be solved by developing toxicological theory on the induction of toxic effects in relation to their physiological control mechanisms. Only in this way biologically sound toxicity thresholds can be defined.

6.3. Benchmark Dose versus NOAEL

The NOAEL is a point estimate, ignoring any uncertainties whereas the Benchmark Dose, being a confidence limit, does take into account the implied uncertainties. This fact could suggest that the same data set would always give rise to a benchmark dose below the NOAEL derived from the same data set. In that case, the ADI based on the benchmark dose would always be lower than that based on the NOAEL. However, this will not hold in general: the benchmark can also turn out to be higher than the NOAEL in individual cases. The reason is that the NOAEL not only can have a lower, but also a higher value than the true No-Adverse Effect Level, because of an unfortunately chosen dose-interval. The Benchmark dose, on the other hand, will always (or rather with high probability) be lower than the true no-effect-level. Therefore, it is not clear beforehand whether an ADI derived from the benchmark dose will on the average be lower or higher than had it been derived from the NOAEL.

6.4. Hybrid approach: estimation of threshold

As mentioned above the setting of biologically sound toxicity thresholds and, hence, their corresponding threshold doses, is, other than by setting arbitrary boundaries for toxicity thresholds, not yet possible. We did not discuss an alternative, intermediate, approach for determining threshold doses : fitting a regression-model having a threshold-dose and determine a confidence interval for that threshold (being one of the model's parameters). A simple example of such a model for continuous data is the 'hockeystick model', assuming a linear response starting at the threshold-dose. However, it may be expected that other threshold models will result in different lower confidence limits for the threshold dose. Furthermore, a typical toxicological data set may be expected to result in a very large confidence interval for this threshold-dose, and the lower confidence limit might even equal zero. On the other hand, Kooijman and Bedaux (1995) have developed a threshold-model for quantal data that appears to result in acceptable confidence intervals for the threshold. This model and its use for deriving ADI's will be examined in the near future.

6.5. Consequences for experimental design.

As yet it has not been discussed that the benchmark approach has certain consequences for the demands of experimental design. If an animal study is designed to derive a benchmark dose, there is no need to use replications per dose group. One could as well choose a series of doses equal to the total number of animals, and allocate each dose to a single animal. As a matter of fact, this design may be considered preferable to the usual design of replicated dose groups, because the unreplicated design would have less chance of lacking data in the most interesting part of the dose-response relation, i.e. that part in which the slope of the dose-response relationhip changes. Of course, this design implies more work regarding the performance of the experiment. On the other hand, this approach might even allow to diminish the total number of animals. We intend to further investigate this issue by a computer simulation study.

6.6. Validation of safety factors

Several studies addressing the validity of the safety factors were discussed. The relevance of these studies differs for the various safety factors. The study of Dourson and Stara (1983) appears the least relevant in this respect, as they quantify the variability in reponse among individual laboratory animals. Since in most animal studies one preferably uses strains showing little variation, their results do not support the intraspecies safety factor that is meant to cover the highly variable human population. To validate the intraspecies safety factor, studies as mentioned in Table 2 give more relevant information. The studies adressing the safety factor for subchronic -> chronic and for LOAEL -> NOAEL extrapolation do seem relevant. Concerning the interspecies extrapolation, the results of the study of Hayes et al. might be extended by examining data from pharmaceuticals where in many cases data from animal and human studies are available.

7. REFERENCES

1. H.E. Aebi and S.R. Wyss, (1978) Acatalasemia. In : The Metabolic Basis of Inherited Disease, J.B. Stanbury, J.B. Wyngaarden and D.S Frederickson (Eds), 4th ed., p 1792, McGraw-Hill, New York.
2. D.G. Barnes and M.L. Dourson, Regul. Toxicol. Pharmacol., 8, (1988), 471-486 : Reference Dose (RfD) : Description and Use in Health Risk Assessments.
3. B.D. Beck, R.B. Conolly, M.L. Dourson, D. Guth, D. Hattis, C. Kimmel and S.S.C. Lewis, Fund. Appl. Pharmacol., 20, (1993), 1-14 : Improvements quantitative noncancer risk assessment.
4. K.S. Crump, Fund. Appl. Toxicol., 4, (1984), 854-871 : A new method for determining allowable daily intakes.
5. W.R. Chappell and J. Mordenti, Adv. Drug Res., 20, (1991), 2-116 : Extrapolation of toxicological and pharmacological data from animals to humans.
6. M.L. Dourson and J.F. Stara, Reg. Toxicol. Pharmacol., 3, (1983), 224-238 : Regulatory history and experimental support of uncertainty (safety) factors.
7. D.J. Ecobichon, The Basis of Toxicity Testing, CRC Press, Boca Raton, 1992.
8. EPA (1980) U.S. Environmental Protection Agency. Guidelines and Methodology used in the preparation of health effects assessment chapters in the consent decree water quality criteria, Federal Register, 45, 79347-79357.
9. EPA (1993) U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS) Background Document 1A.
10. D.W. Gaylor and R.L. Kodell, J. Env. Pathol. Toxicol., 4, (1980), 305-312 : Linear interpolation algorithm for low dose assessment of toxic substances.
11. D.W. Gaylor, Toxicol. Ind. Health, 4, (1988), 453-459 : Applicability of cancer risk assessment techniques to other toxic effects.
12. H. Glatt and F. Oesch, (1984) Variations in epoxide hydrolase activities in human liver and blood. In : Genetic variability in response to chemical exposure : Banbury Report 16, G.S. Omenn and H.V. Gelboin (Eds.), p 189-201, Cold Spring Harbor Laboratory, New York.
13. I.R. Glowinski, H.E. Radtke and W.W. Weber, Mol. Pharmacol., 14, (1978), 940-949 : Genetic variation in N-acetylation of carcinogenic arylamines by human and rabbit liver.
14. W.J. Hayes, Proc. Roy. Soc. London, 167, (1007), (1967), 101-127 : Toxicity of pesticides to man : Risks from present levels.
15. J.A. Hoekstra, Statistics in Ecotoxicology, Thesis, Amsterdam, The Netherlands, 1993.
16. J.R. Idle and R.L. Smith, Drug Metabol. Rev., 9, (2), (1979), 301-317 : Polymorphisms of oxidation at carbon centers of drugs and their clinical significance.
17. F.R. Johannsen, Critical Rev. Toxicol., 20, (5), (1990), 341-367 : Risk Assessment of Carcinogenic and Noncarcinogenic Chemicals.
18. C.A. Kimmel and D.W. Gaylor, Risk Analysis, 8,(1), 1988, 15-20 : Issues in qualitative and quantitative risk analysis for developmental toxicology.
19. S.A.L.M. Kooijman and J.J.M. Bedaux, Statistical properties of no effect levels, submitted
20. M. Legge, M. Brian, C. Winterbourn and R. Carrell, Aust. Paediatr. J., 13, (1977), 25-28 : Red cell superoxide dismutase activity in the newborn.

21. R.S. McColl, Biological Safety Factors in Toxicological Risk Assessment, Environmental Health Directorate, Health Protection Branch, Health and Welfare Canada, Ottawa, Canada, 1990.
22. B.P. McNamara, (1976), Concepts in health evaluation of commercial and industrial chemicals. In : New Concepts in Safety Evaluation, M.A. Mehlman, R.E. Shapiro and H. Blumenthal, (Eds.), p 61-140, Hemisphere, Washington.
23. A.G. Renwick, Food. Add. Contam., 7, (4), (1990, 463-475 : Acceptable Daily Intake and the regulation of intense sweeteners.
24. A.G. Renwick, Food Add. Contam., 8, (2), (1991), 135-150 : Safety Factors and establishment of acceptable daily intakes.
25. A.G. Renwick and R. Walker, Reg. Toxicol. Pharmacol., 18, (1993), 463-480 : An analysis of the risk of exceeding the acceptable or tolerable daily intake.
26. C.S. Weil, Toxicol. Appl. Pharmacol., 21, (1972), 454-463 : Statistics vs Safety Factors and scientific judgment in the evaluation of safety for man.

Appendix 1. RENWICK'S ACCEPTABLE EXTENT OF EXPOSURE EXCEEDING THE ADI

1. Renwick's calculation of the acceptable extent of exposure exceeding the ADI : Critical Remarks

By definition an exposure situation at the level of the ADI should not lead to toxic effects. However, due to variation in individual exposure patterns exposure might, temporarily or permanently, be higher than the ADI. This led Renwick to the calculation of the maximum that the ADI may be exceeded without leading to significant toxicity. In this calculation the following procedure is followed. First a criterium for the upper bound of toxicity which can be considered as safe is set. Then the dose level corresponding to this safe effect level is determined. This dose level then is used for the calculation of the upper bound of the ADI, i.e. the maximum incursion of the ADI expected to be without significant toxicity.

Renwick considers all mean/median effect levels below the (for example) 95% upper confidence limit of effect at the NOAEL as safe (see Fig. 1). Interpolating this effect level on dose then gives the maximum dose corresponding to a non-significant, i.e. still safe, incursion of the dose-response curve at the NOAEL ($NOAEL_{upper}$). After correction for the inaccuracy with which the toxicity threshold (NAEL) is estimated by the NOAEL this dose level is used for the calculation of the maximum incursion of the ADI expected to be without significant toxicity. The calculation of this incursion, which is shown in § 2, ultimately results in the following equation relating the ADI to its safe upper boundary :

$$Upper\ Bound_{ADI} = ADI * \{\frac{NAEL}{NOAEL} + \alpha * \frac{0.5 * 95\%\ CI\ of\ effect\ at\ the\ NOAEL}{response\ slope\ at\ the\ NOAEL * NOAEL}\}$$

with :

$Upper\ Bound_{ADI}$	: the maximum incursion of the ADI expected to be without significant toxicity
ADI	: estimate of the toxicity threshold. Defined as the ratio of the NOAEL and the traditional uncertainty factor of 100
α	: response slope correction factor for difference in chemical kinetics at the NOAEL and the NAEL

The way the upper bound of the ADI is calculated by Renwick and Walker however is prone to several statistical and practical impediments. First, statistical entities, being prone to experimental error, will vary between experiments and hence are, as suggested, unsuited for the definition of (absolute) toxicity thresholds, i.e. safe vs. non-safe dose levels (different experiments leading to different values for toxicity thresholds). Second, the definition of the upper confidence limit of effect at the NOAEL as a safety criterium is based on the reasoning that all values lying within the (for example) 95% confidence interval of effect at the NOAEL cannot be excluded as mean/median effect levels at that dose level. Considering all such effect levels as safe and reasoning analogously then leads

to the definition of all effect levels lying within the confidence limits of effect found at the LOAEL as non-safe which, in cases where confidence intervals at the NOAEL and the LOAEL overlap, leads to the definition of certain effect level as both safe and non-safe. Third, within Renwick's definition of safe effect levels, i.e. the upper confidence limit of effect at the NOAEL, the dose corresponding to a non-significant incursion of the dose-response curve at the NOAEL ($NOAEL_{upper}$) represents a (mean) exposure level at which one-half of the exposed population is expected to show effect levels exceeding the defined safe effect level. Such exposure levels are hardly to be considered as "to give rise to a non-significant incursion of the dose-response relationship". Fourth, in calculating the upper bound of the ADI knowledge on the NAEL/NOAEL ratio is needed. However, as the absolute value of the NAEL is (and will be) unknown and its value can only be expressed in terms of statistical uncertainty no single, absolute value of the upper bound of the ADI can be calculated., as is suggested by the above mentioned equation.

Given these shortcomings we therefore consider the calculation of the maximal acceptable extent of the exposure exceeding the ADI as suggested by Renwick and Walker as invalid.

2. Definition of Renwick's maximal acceptable extent of exposure exceeding the ADI

Given the variation of response at the NOAEL and considering all mean/median response levels below the upper 95% confidence limit of response at the NOAEL as safe the maximum dose corresponding to a non-significant incursion of the dose-response curve at the NOAEL ($NOAEL_{upper}$) is found by interpolation of effect on dose (see Fig. 1). Assuming the dose-response relationship at the NOAEL approximately linear the value of the $NOAEL_{upper}$ is given by :

$$NOAEL_{upper} = NOAEL + \frac{0.5 * 95\%\ CI\ of\ effect\ at\ the\ NOAEL}{response\ slope\ at\ the\ NOAEL}$$

with :

$NOAEL_{upper}$	:	maximum dose corresponding to a non-significant, i.e. still safe, incursion of the dose-response curve at the NOAEL
NOAEL	:	the highest experimental dose at which there is no statistically significant increase/decrease in the frequency or the severity of an adverse toxic effect parameter in a treated group relative to its control group

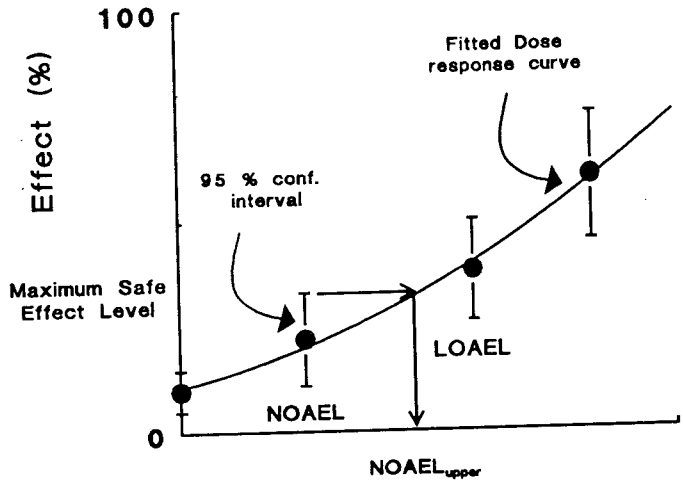

Fig. 1 A graphical presentation of the determination of the $NOAEL_{upper}$

Correcting for the inaccuracy with which the NAEL is estimated by the NOAEL and assuming the dose-response relationship at the NAEL and the NOAEL approximately linear with equal slopes, this equation is equivalent to :

$$NAEL_{upper} = NAEL + \frac{0.5*95\%\ CI\ of\ effect\ at\ the\ NOAEL}{response\ slope\ at\ the\ NOAEL}$$

with :

NAEL	:	true toxicity threshold
$NAEL_{upper}$	:	maximum dose corresponding to a non-significant incursion of the dose-response curve at the NAEL

Defining the true ADI as the ratio between the toxicity threshold (NAEL) and the traditional uncertainty factor of 100, the maximum human reference dose corresponding to a non-significant incursion of the dose-response curve at the true ADI is given by :

$$\frac{NAEL_{upper}}{UF} = \left\{\frac{NAEL + \frac{0.5*95\%\ CI\ of\ effect\ at\ the\ NOAEL}{response\ slope\ at\ the\ NOAEL}}{UF}\right\}$$

which, by definition, is equal to :

$$Upper\ Bound_{ADI} = ADI * \{ \frac{NAEL}{NOAEL} + \frac{0.5 * 95\%\ CI\ of\ effect\ at\ the\ NOAEL}{response\ slope\ at\ the\ NOAEL\ *\ NOAEL} \}$$

with :

Upper $Bound_{ADI}$	:	the maximum incursion of the ADI expected to be without significant toxicity
ADI	:	estimate of the toxicity threshold. Defined as the ratio between the NOAEL and the traditional uncertainty factor of 100.

Ignoring differences in toxicokinetics at the NOAEL and the NAEL the latter equation equals Renwick's equation for the calculation of the overall extent of exposure above the ADI that is still equivalent to "no significant toxic effect" (Renwick and Walker, 1993, pp. 473).

Appendix 2

EPA RATING VALUES FOR TOXIC EFFECT PARAMETERS (RV_e)

(After Hartung and Durkin, (1986) as cited in McColl, (1990))

Rating	Effect
1	Enzyme induction or other biochemical change with no pathological changes and no change in organ weight
2	Enzyme induction and subcellular proliferation or other changes in organelles, but no other apparent effects
3	Hyperplasia, hyperthrophy or atrophy but no changes in organ weights
4	Hyperplasia, hypertrophy or atrophy but changes in organ weights
5	Reversible cellular changes : cloudy swelling, hydropic change or fatty changes
6	Necrosis or metaplasia with no apparent behavioural, sensory or physiologic changes
7	Necrosis, atrophy, hypertrophy or metaplasia with a detectable decrement of organ functions. Any neuropathy with a measurable change in behavioural, sensory or physiologic activity
8	Necrosis, atrophy, hypertrophy or metaplasia with definite organ dysfunction. Any neuropathy with gross changes in behaviour, sensory or motor performance. Any increase in reproductive capacity. Any evidence of fetotoxicity.
9	Pronounced pathologic changes with severe organ dysfunction. Any neuropathy with loss of behavioural or motor control or loss of sensory ability. Reproductive dysfunction. Any teratogenic effect with maternal toxicity.
10	Death or pronounced life shortening. Any teratogenic effect without signs of maternal toxicity.