NATIONAL INSTITUTE OF PUBLIC HEALTH AND THE ENVIRONMENT (RIVM)
BILTHOVEN
NETHERLANDS ORGANIZATION FOR APPLIED SCIENTIFIC RESEARCH (TNO)
ZEIST

RIVM Report no. 620110007, TNO Report no. V97.880

Assessment factors for human health risk assessment: a discussion paper

T.G. Vermeire, H. Stevenson, M.N. Pieters,
M. Rennen, W. Slob, B.C. Hakkert

March 1998

This investigation has been performed by order of the Ministry of Housing, Spatial Planning and the Environment, Directorate-General for Environmental Protection, Direction of Chemicals, Safety and Radiation Protection (DGM-SVS) and the Ministry of Social Affairs and Employment (SZW).

National Institute of Public Health and the Environment (RIVM), P.O. Box 1, 3720 BA Bilthoven, the Netherlands
Tel.: 3130 - 2749111, fax: 3130 - 2742971
TNO Nutrition and Food Research Institute, Department of Occupational Toxicology, P.O. Box 360, 3700 AJ Zeist, the Netherlands
Tel.: 3130 - 6944144, fax: 3130 - 6944926

MAILING LIST/VERZENDLIJST

1	Directoraat-Generaal Milieubeheer, Directeur Stoffen, Veiligheid en Straling, t.a.v. dr. C. Plug
2 - 6	Ministerie van Sociale Zaken en Werkgelegenheid, A en O, Onderzoek, t.a.v. drs. A.A. Vijlbrief en dr.ir. A. Hollander
7	Plv. Directeur-Generaal Milieubeheer, dr. B.C.J. Zoeteman
8	Dr. D.W.G. Jung, DGM/SVS
9	Dr. Ir. P.C. Bragt, VWS/HIGB
10	Dr. H. Roelfzema, VWS/GZB
11 - 17	EC-OECD-Commissies d.t.v. Dr.C.J.van Leeuwen
18	Dr. W.F. ten Berge, DSM, Heerlen
19	Prof.dr. J.S.M. Boleij, CTB, Wageningen
20	Dr. F.M. Carpanini, ECETOC, Brussel, België
21	Prof.dr. P.T. Henderson, RUL, Maastricht
22	Dr. J.I. Delic, HSE, Bootle, UK
23	Dr. M.L. Dourson, TERA, Cincinnati, USA
24	Dr. H.E. Falke, CTB, Wageningen
25	Dr. J. Gift, EPA-NCEA, USA
26	Dr. S. Hanai, JCIA, Tokyo, Japan
27	Dr. B. Hansen, ECB, Ispra, Italië
28	Dr. R.F. Hertel, BGVV, Berlijn, Duitsland
29	Dr. J. Herman, IPCS, Genève, Zwitserland
30	Dr. A. Jarabek, EPA-NCEA, USA
31	Dr. F. Kalberlah, FoBiG, Freiburg, Duitsland
32	Prof.dr. J. Koeman, LUW, Wageningen
33	Prof.dr. R. Kroes, RITOX, Utrecht
34	Dr. C. Lally, P&G, Strombeek-Bever, België
35	Dr. F.X.R. van Leeuwen, WHO, Bilthoven
36	Dr. E. Löser, Bayer, Duitsland
37	Dr. B. Meek, Health Canada, Ottawa, Canada
38	Dr. M. Mercier, IPCS, Genève, Zwitserland
39	Dr. G. Moore, KEMI, Sweden
40	Dr. W.F. Passchier, Gezondheidsraad, Den Haag
41	Dr. M. Reynier, INRS, Parijs, Frankrijk
42	Prof.dr. W. Seinen, RITOX, Utrecht
43	Dr. Y. Sikezawa, NIHS, Tokyo, Japan
44	Dr. C. Sonich-Mullin, IPCS, Genève, Zwitserland
45	Dr. R. Taalman, SHELL, London, UK
46	Dr. R. Tomerlin, NoviGen Sciences, Washington D.C., USA
47	Dr. W.F. Tordoir, SHELL, Den Haag
48	Dr.H.G. Verschuuren, Dow Europe, Horgen, Zwitserland
49	Dr. E. van Vliet, Gezondheidsraad, Den Haag
50	Dr. A.A.E Wibowo, Coronel Laboratorium, Amsterdam

51 Dr. A. Wilson, HSE, Bootle, UK
52 Directie RIVM
53 RIVM: Sectordirecteur Stoffen en Risico's
54 RIVM: Sectordirecteur Volksgezondheidsonderzoek
55 RIVM: Hoofd Laboratorium voor Blootstellingsonderzoek
56 RIVM: Hoofd Centrum voor Stoffen en Risicobeoordeling
57 RIVM: Hoofd Laboratorium voor Effectenonderzoek
58 RIVM: Hoofd Afdeling Voorlichting en Public Relations
59 - 69 RIVM: Beoordelingsgroepen H en H/M d.t.v. Drs.A.G.A.C. Knaap
70 - 71 RIVM: Centrum voor Stoffen en Risicobeoordeling
72 - 81 RIVM: intern
82 Directie TNO
83 TNO: Hoofd Divisie Toxicologie
84 TNO: Hoofd Afdeling Arbeidstoxicologie
85 - 91 TNO: Afdeling Arbeidstoxicologie
92 -101 TNO: intern
102-106 Auteurs
107 Rapportenregistratie RIVM
108-109 Bibliotheek RIVM
110-111 Bibliotheek TNO
111-150 Reserve exemplaren
151 Depot van Nederlandse publicaties en Nederlandse bibliografie

TABLE OF CONTENTS

ABSTRACT

The general goal of this discussion paper is to contribute towards further harmonisation of the human health risk assessment. It discusses the development of a formal, harmonised set of default assessment factors. The status quo with regard to assessment factors is reviewed. Options are presented for a set of default values or probabilistic distributions for assessment factors based on the state of the art. Methods of combining default values or probabilistic distributions of assessment factors are described. The benchmark dose concept is proposed for better characterisation of the true human no-effect level in a probabilistic manner. It is shown how the probabilistic benchmark dose distribution can be combined with distributions of assessment factors to arrive at the distribution of a Human Limit Value.

SUMMARY

The general goal of this discussion paper is to contribute towards further harmonisation of the human health risk assessment. Although much of the contents of this report is applicable to the human health risk assessment of chemical substances in general, it concentrates on the assessment of new and existing substances within the scope of European Union legislation. More specifically, it intends to be a contribution towards the development of a formal, harmonised set of assessment factors to be applied within the scope of the EU risk assessment for new and existing substances.

MOS versus NOAEL
In the European Union, Directive 92/32/EC and EC Council Regulation (EC) 793/93 require the risk assessment of new and existing substances in accordance with a detailed package of Technical Guidance Documents (TGD). In the human risk assessment an attempt is made to identify the hazards of the substances and to relate them to exposure. For those substances for which a threshold for toxicity is assumed to exist, a No-Observed-Adverse-Effect Level (NOAEL) has to be derived or, if this is not possible, a Lowest-Observed-Adverse-Effect Level (LOAEL). The risk is characterised by comparing estimated (or measured) concentrations in air or on skin or total daily intakes to the results of the effects assessment. This Margin Of Safety (MOS) is determined separately for each population potentially exposed and for each effect.

Another widely used approach is the explicit derivation of a Human Limit Value (HLV[1]) from experimental or epidemiological toxicity data by dividing the NOAEL or LOAEL by an overall assessment factor. This overall assessment factor is a multiple of several factors, accounting among others for inter- and intraspecies variations, differences in exposure time scales, the nature of the adverse effects and the adequacy of the database. In this report the Margin of Safety approach of the current TGD is compared to the assessment factor approach. It is noted that the current TGD does not provide any quantitative guidance on the size of the Margin of Safety and concluded that decision criteria for human risk characterisation need to be made explicit. This could be achieved by establishing a formal, harmonised set of default assessment factors accompanied by elaborate guidance. The default set should only be applied in the absence of data which permit a more substance-specific, scientific choice. The default set should allow for differentiation with regard to exposure scenarios, including groups at risk, and the toxicological database. Harmonisation should not hamper further developments, i.e. should not be seen as standardisation for everything and for ever.

Assessment factors
This report gives an overview of published extrapolation methods based on the assessment factor approach. It discusses the status quo with regard to the type of factors to be identified, the range of values assigned as well as the presence or absence of a scientific basis for these values. It is shown that all methods use experts for judgement of the underlying toxicological

[1] HLV: general term covering various limit values such as the ADI, RfD, PNAEL and HBORV (see section 2.2.)

database and the severity of the effects. Few approaches are based on scientific data, but most methods basically rely on the arbitrary 100-fold factor used to derive the Acceptable Daily Intake (ADI). It is recommended to investigate the probabilistic nature of assessment factors and to try to derive their distribution. An attempt is made to estimate the distribution of several assessment factors from historical data, i.e. NOAEL-ratios. This analysis shows that the alleged worst case character of the traditional default assessment factors is doubtful: the 95-percentiles of the proposed distributions for the interspecies factor and the semi-chronic to chronic factor are considerably higher than 10 and the limited data on intraspecies variation also indicate that a default factor of 10 may not be sufficient. More work is needed to better characterise the distributions of assessment factors. Probabilistic multiplication of these distributions is preferred above simple multiplication to avoid extreme conservatism without indication how conservative it may be.

The benchmark dose concept
The NOAEL selected from the toxicological database may be a poor substitute for the unknown, true NAEL. New developments are presented with regard to the estimation of a NAEL. The already widely discussed Benchmark Dose concept can be extended to obtain an uncertainty distribution of the Critical Effect Dose (CED). This CED-distribution can be combined with estimated uncertainty distributions for assessment factors. In this way the full distribution of the HLV will be derived and not only a point estimate, whereas information on dose-response relations is taken account of. The method potentially can reveal the full range of variability and uncertainty in both the HLV and the exposure estimate as well as elicit expert judgment in a transparent way. It thus allows the risk manager to make a quantitative cost-benefit analysis. However, consensus on the definition of the Critical Effect Size is needed and this requires further research. The method also requires a certain statistical experience to fit mathematical dose-response models and in many cases the data available will not allow modelling, since current protocols are not intended to generate dose-response curves.

The various methods discussed are applied to an example substance to show the technical procedures involved.

The way forward
It should be considered to replace the traditional application of default assessment factors of 10 by more database-derived defaults and distributions of assessment factors. Probabilistic multiplication of distributions of these factors is then preferred above simple multiplication to avoid extreme conservatism. This can be considered as a first step for the implementation of a complete probabilistic method of risk assessment but more work is needed to fully override the traditional method. Through this stepwise implementation further research in this area will benefit from the experience gained. At the same time both risk managers and the public will have to adapt themselves to the new type of judgements they face: rather than relying on one point estimate for the HLV, a decision has to be made on the acceptable degree of confidence in the estimation of the HLV as well as in the estimation of exposure.

SAMENVATTING

De algemene doelstelling van dit discussierapport is bij te dragen aan verdere harmonisatie van de humaan-toxicologische risicobeoordeling. Hoewel een groot deel van dit rapport van toepassing is op de risicobeoordeling van stoffen in allerlei kaders, concentreert het zich op de beoordeling van nieuwe en bestaande stoffen in EU-kader.

MOS versus NOAEL
In de Europese Unie vereisen Richtlijn 92/32/EC en EC Verordening (EC) 793/93 de beoordeling van de risico's van nieuwe en bestaande stoffen overeenkomstig de richtlijnen in de "Technical Guidance Documents" (TGD). In de humaan-toxicologische beoordeling moeten de gevaren van de stof gerelateerd worden aan blootstellingsniveaus. Voor stoffen die geacht worden een toxicologische drempelwaarde te hebben moet een "No-Observed-Adverse-Effect Level" (NOAEL) of, indien dit niet mogelijk is, een "Lowest-Observed-Adverse-Effect Level (LOAEL) afgeleid worden. Het risico wordt dan gekarakteriseerd door a. de geschatte (of gemeten) blootstellingsconcentraties in de lucht of op de huid of b. totale dagelijkse opnames te vergelijken met de resultaten van de effectbeoordeling. Deze "Margin of Safety" (MOS) wordt voor elke potentieel blootgestelde populatie en voor elk effect vastgesteld.

Een ander wijd en zijd gebruikte benadering is de expliciete afleiding van een grenswaarde ("Human Limit Value", HLV[1]) op basis van experimentele of epidemiologische toxiciteitsgegevens waarbij de NOAEL of LOAEL wordt gedeeld door een samengestelde "assessment" factor (AF). Deze factor is opgebouwd uit verschillende assessment factoren die onder meer de onzekerheden kwantificeren ten gevolge van de variatie tussen de soorten, de variatie binnen de humane populatie, de verschillen in tijdschalen van blootstelling, de aard van de nadelige effecten en de kwaliteit van de onderliggende data. In dit rapport wordt de MOS-benadering vergeleken met de assessment factor benadering. Opgemerkt wordt dat de huidige TGD geen kwantitatieve richtlijnen geeft voor de grootte van de MOS en dat beslissingscriteria voor de humane risicokarakterisering expliciet gemaakt dienen te worden. Dit zou kunnen worden bereikt door het vaststellen van een formele, geharmoniseerde set default assessment factoren vergezeld van een uitgebreide leidraad. De default set zou alleen moeten worden toegepast als er geen gegevens voorhanden zijn die een meer stofspecifieke, wetenschappelijke keuze mogelijk maken. De default set zou rekening moeten kunnen houden met de mogelijke verschillen in blootstellingsscenario's, met inbegrip van risicogroepen, en in de onderliggende toxicologische gegevens. Harmonisatie moet echter geen rem zijn op verdere ontwikkelingen en niet gezien worden als een permanente standaardisatie.

Assessment factoren
Dit rapport geeft een overzicht van de gepubliceerde extrapolatiemethoden die gebaseerd zijn op de assessment factor benadering. De bestaande situatie wordt besproken voor wat betreft de te onderscheiden typen factoren, de toegekende waarden en de aan- of afwezigheid van een

[1] HLV: algemene term die verschillende soorten limietwaarden zoals ADI, RfD, PNAEL en HBORV dekt (zie 2.2)

wetenschappelijke verantwoording. Het blijkt dat alle methoden deskundigen inzetten ter beoordeling van de onderliggende toxicologische data en de ernst van de effecten. Enkele benaderingen baseren zich op wetenschappelijke gegevens, maar over het algemeen wordt afgegaan op de willekeurig gekozen factor 100 die wordt toegepast in de afleiding van de Acceptable Daily Intake (ADI). Aanbevolen wordt het probabilistische karakter van assessment factoren nader te onderzoeken en om te trachten de verdelingen daarvan te achterhalen. Uitgaande van historische gegevens, NOAEL-ratio's, wordt een poging gedaan om de verdeling van verschillende assessment factoren te schatten. Deze analyse toont aan dat het vermeende "worst case" karakter van de traditionele assessment factoren betwijfeld kan worden: de 95-percentielen van de voorgestelde verdelingen voor de factoren die corrigeren voor verschillen tussen de soorten and tussen de semi-chronische en chronische tijdschaal zijn aanzienlijk hoger dan 10. Ook de beperkte gegevens over de variatie binnen de humane populatie tonen aan dat een factor 10 niet voldoende groot is. Meer onderzoek naar de verdelingen van de assessment factors is noodzakelijk. Om extreem conservatisme te vermijden wordt de voorkeur gegeven aan de probabilistische vermenigvuldiging van deze verdelingen.

Het "benchmark dose" concept

De uit de toxicologische data geselecteerde NOAEL kan een slechte schatter zijn voor de onbekende, werkelijke NAEL. Er worden nieuwe methoden gepresenteerd om deze NAEL te schatten. Het reeds breed besproken "benchmark dose" concept kan zodanig uitgebreid worden dat een probabilistische verdeling van de kritische effectdosis ("Critical Effect Dose", CED) wordt verkregen. Deze CED-verdeling kan worden gecombineerd met de geschatte onzekerheidsverdelingen van assessment factoren. Op deze wijze wordt de volledige probabilistische verdeling van de HLV afgeleid en niet alleen een puntschatting, terwijl bovendien rekening wordt gehouden met hetgeen bekend is over dosis-response relaties. The methode kan in potentie de volledige variabiliteit en onzekerheid in zowel de HLV als de blootstellingsschatting in beeld kwantificeren en de meningen van deskundigen op een transparante wijze peilen. Dit stelt de beleidsmaker in staat om een kwantitatieve kosten-baten analyse te maken. Consensus over de definitie van de CES is echter vereist en dit vraagt om nader onderzoek. Een zekere statistische ervaring is ook nodig om mathematische dosis-response modellen te kunnen hanteren. In veel gevallen zullen er onvoldoende gegevens beschikbaar zijn voor deze modellering, aangezien de huidige testprotocollen niet zijn opgesteld om dosis-response curven te genereren. Om inzicht te geven in de technische procedures worden de verschillende methoden die in dit rapport besproken worden op een modelstof toegepast.

Hoe verder?

Overwogen dient te worden om de traditionele toepassing van default assessment factoren van 10 te vervangen door defaults die meer zijn gebaseerd op de beschikbare data en door Om extreem conservatisme te vermijden wordt de voorkeur gegeven aan de probabilistische vermenigvuldiging van deze verdelingen verdelingen van assessment factoren. Dit kan als een eerste stap beschouwd worden in de implementatie van een volledig probabilistische risicobeoordelingsmethode. Meer onderzoek is echter nodig om de traditionele methode

volledig te kunnen vervangen. Door een dergelijke stapsgewijze benadering kan nader onderzoek profiteren van de opgedane ervaringen. Tegelijkertijd kunnen zowel beleidsmakers als de gemeenschap gewend raken aan dit nieuwe type risicobeoordelingen: moest eerst worden vertrouwd op een puntschatting van de HLV, dan moet een beslissing genomen worden op basis van een aanvaarbaar geacht betrouwbaarheidsniveau van zowel de HLV als de blootstellingsschatting.

LIST OF ABBREVIATIONS

ADI	Acceptable Daily Intake
AF	Assessment Factor
AOEL	Acceptable Operator Exposure Level
BMDL	Benchmark Dose Level
CED	Critical Effect Dose
CES	Critical Effect Size
ECETOC	European Centre for Ecotoxicology and Toxicology of Chemicals
EPA	Environmental Protecton Agency (USA)
EUSES	European Union System for the Evaluation of Substances
FAO	Food and Agricultural Organisation
FDA	Food and Drug Administration (USA)
GM	Geometric Mean
GSD	Geometric Standard Deviation
HLV	Human Limit Value
HBORV	Health Based Occupational Reference Value
HC	Health Council (of The Netherlands)
IPCS	International Programme on Chemical Safety
JECFA	Joint Expert Committee on Food Additives
JMPR	Joint Meeting of Experts on Pesticide Residues
LLN	Lewis/Lynch/Nikiforov model
MOS	Margin Of Safety
LOAEL	Lowest-Observed-Adverse-Effect Level
MF	Modifying Factor
NAEL	No-Adverse-Effect level
NOAEL	No-Observed-Adverse-Effect Level
OECD	Organisation for Economic Co-operation and Development
PNAEL	Predicted No-Adverse-Effect level
P_x	xth percentile
SF	Safety Factor
RfD	Reference Dose
RIVM	National Institute of Public Health and the Environment
TDI	Tolerable Daily Intake
TGD	Technical Guidance Documents
TNO	Netherlands Organisation for Applied Scientific Research
UF	Uncertainty Factor
WHO	World Health Organisation

SUMMARY OF EXPRESSIONS

Geometric mean of the lognormal distribution

$$GM = \exp(\frac{1}{n}\sum_{i=1}^{n} \ln X_i))$$

Sample variance of log-entities

$$s^2_{\ln X} = \frac{1}{n-1}\sum_{i=1}^{n}(\ln X_i - \ln GM)^2$$

Geometric standard deviation

$$GSD = \exp(s_{\ln X})$$

95th percentile $P_{0.95}$

$$P_{0.95} = GM \cdot GSD^{z_{0.95}}$$

n: number of observations
X_i: lognormally distributed ith observation (e.g. NOAEL)
s_{lnX} : sample standard deviation of lognormally distributed X
$z_{0.95}$: 95th percentile of the standard normal distribution

1. INTRODUCTION

1.1 TGD approach towards human risk characterisation

The general goal of this report is to contribute towards further harmonisation of the human health risk assessment. Although much of the contents of this report is applicable to the human health risk assessment of chemical substances in general, it concentrates on the assessment of new and existing substances within the scope of European Union legislation. In the European Union, Directive 92/32/EC (EC, 1992) and EC Council Regulation (EC) 793/93 (EC, 1993a) require the risk assessment of new and existing substances, respectively. Principles for this risk assessment have been laid down (EC, 1993b; EC, 1994), supported by a detailed package of Technical Guidance Documents (TGD; EC, 1996a) and the software implementation EUSES (EC, 1996b). This risk assessment process proceeds along a causal chain following the substance from its origin to the place where it is available to organisms and may exert adverse effects. The exposures can be identified as acute, semi-chronic or chronic. Populations considered are consumers, workers and man exposed through the environment.

The following summarises important aspects of the human effects assessment and risk characterisation from the current TGD. In the human risk assessment an attempt is made to identify the hazards of the substances and to relate them to exposure. For those substances for which a threshold for toxicity is assumed to exist, a No-Observed-Adverse-Effect Level (NOAEL), has to be derived or, if this is not possible, a Lowest-Observed-Adverse-Effect Level (LOAEL). The NOAEL is the highest concentration or amount of a substance, found by experiment or observation, which causes no detectable adverse alteration of morphology, functional capacity, growth, development, or life span of the target organisms under defined conditions of exposure (WHO, 1979). Unless a N(L)OAEL value is available from human data, the N(L)OAEL values are those derived from animal studies. In the risk assessment for man, the risk is characterised by comparing estimated (or measured) concentrations in air or on skin or total daily intakes to the results of the effects assessment. This analysis is made separately for each population potentially exposed and for each effect.

Where the exposure estimate is higher than or equal to the N(L)OAEL, the TGD qualifies the substance as "of concern"[1]. If the exposure estimate is less than the N(L)OAEL, the risk assessor is asked to decide whether the magnitude by which the N(L)OAEL exceeds the estimated exposure, the "Margin of Safety", is of concern. The TGD recommends the following parameters to be considered in assessing the Margin of Safety:

- the uncertainty arising, among other factors, from the variability in the experimental data and intra- and interspecies variation;
- the nature and severity of the effect;

[1] If it is concluded that a substance is "of concern", it is considered likely that adverse effects can be expected in human populations from known or reasonably forseeable use. In that case further data have to be requested or risk reduction recommendations made.

- the human population to which the quantitative and/or qualitative information on exposure applies;
- the differences in exposure (route, duration, frequency and pattern);
- the dose-response relationship observed;
- the overall confidence in the database.

The TGD further states that expert judgement is required to weigh these individual parameters on a case-by-case basis. Transparency is required. The TGD refers to several relevant publications which may support this expert judgement.

1.2 Methods using assessment factors

Another widely used approach is the explicit derivation of a Human Limit Value (HLV)[1] for man from experimental or epidemiological toxicity data by dividing the NOAEL or LOAEL by an assessment factor. This overall assessment factor is a multiple of several factors (called safety factors, uncertainty factors, extrapolation factors, adjustment factors, conversion factors; see definitions below), accounting for inter- and intraspecies variations, differences in exposure time scales, the nature of the adverse effects, the adequacy of the database etc. This approach is described and discussed extensively in the scientific literature. Pertinent reviews are produced by Dourson (1996), the (Dutch) Health Council (1985), McColl (1989), IPCS (1994), ECETOC (1995) and Stevenson et al. (1995a).

The consistent use of terminology is a prerequisite for harmonisation discussions. Reviewing the available literature on this subject, the differences in terminology with regard to these factors is striking. Therefore, before turning to a comparison of the two approaches in human effects assessment and risk characterisation as introduced above, it is essential to present the definitions as they will be used throughout this report, acknowledging that the use of another set of definitions may be equally justified:

Assessment factor	*general term to cover all factors designated as safety factor, uncertainty factor, extrapolation factor, adjustment factor, conversion factor, etc. and the composite thereof.*
Extrapolation factor	*database-derived factor, used in the extrapolation from experimental or epidemiological toxicity data to a health based recommended exposure level for man, which takes into account uncertainty due to inter- and intraspecies variability, variability in exposure duration, variability in nature and severity of effects, including the dose-response, and variability in the adequacy of the database.*

Other terminology will not be used in this report or explicitly explained, unless the term is a quote (chapter 2).

[1] HLV: general term covering various limit values such as the ADI, RfD, PNAEL and HBROV (see section 2.2.)

Within the scope of the EU legislation on risk assessment for new and existing substances two approaches for the application of assessment factors in human health risk assessment have been proposed: one by a Task Force of the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC, 1995) and one by TNO Nutrition and Food Research Institute (Stevenson et al., 1995a; Hakkert et al., 1996). TNO Nutrition and Food Research Institute also compared both approaches (Stevenson et al., 1995b). A harmonised WHO-scheme for the derivation of guidance values for health-based exposure limits (ADI) by Task Groups of the International Programme on Chemical Safety has been adopted (IPCS, 1994).

1.3 A comparison

In this section the Margin of Safety approach (MOS-approach) of the current TGD is compared to the assessment factor approach (AF-approach). Note that the current TGD does not provide any quantitative guidance on the size of the Margin of Safety. This subject clearly was not yet ready for EU-wide harmonisation at the time of the development of the TGD given the time constraints and the gaps in opinions to bridge. The AF-approach can be used either to derive a limit value or to indicate the minimum size of the Margin of Safety.

Aspects to be covered in the comparison of the current MOS-approach and the AF-approach are: variability, transparency, consistency and acceptability.

Variability
The current MOS-approach is inherently more variable than the AF-approach. The risk characterisation on the basis of a MOS is heavily dependent on expert judgement, whereas assessment factors need to be applied on the basis of more or less fixed criteria. The more elaborate these criteria are, the less degrees of freedom are left for expert judgement. Expert judgement, however, always plays a role in the application of assessment factors, but in a more formalised way. The outcome of expert judgement is highly uncertain as was illustrated by Dourson and Lu (1995) who compared two sets of 65 risk assessments developed by WHO and US-EPA. ADI and RfD values were within a 3-fold range for 38 sets, a 3 to 30-fold range for 20 sets, a 30 to 300-fold range for 6 sets and in one case differed 700-fold.

Transparency
Transparency in this context indicates how clear the choices made in the human risk characterisation stage are to all stakeholders. The high variability of the current MOS-approach should force experts to explain elaborately their way of thinking in each risk characterisation. Very often risk assessors fail to do so and therefore transparency is not easily achieved. Although the application of assessment factors also needs to be made transparent, the burden on the expert is less since certain criteria are already explicit.

Consistency
Consistency in this context denotes the logical agreement of risk characterisation decisions made for different substances in similar situations. Unless the application of the MOS-

approach is accompanied by careful documentation of decisions and criteria applied by experts, it will be very difficult to maintain consistency. "Institutional memory" is a great asset here, but no guarantee to full consistency over longer periods of time. Clearly, the demand for careful documentation also applies to the AF-approach, but in this case many decisions to be made and criteria are already explicit and only need to be referred to. Problems to maintain consistency are increased when more expert groups are involved sharing the burden of the work such as within the scope of the EU risk assessment for new and existing substances.

Acceptability

- *Acceptability to the risk assessor*

One could argue that a highly variable risk characterisation system such as the current MOS-approach may be more acceptable than an inherently more rigid AF-approach because it allows each party involved to keep its own standards high and therefore keeps everyone satisfied. However, from the above it is also clear that in a co-operative risk assessment scheme such as that of the EU for new and existing substances, it will be virtually impossible to maintain consistency across all substances and to be transparent at the same time.

- *Acceptability to risk managers*

In this situation risk managers are confronted with outcomes of risk assessments which are difficult to classify and to base risk reduction strategies upon. Would the risk manager have the luxury of a second opinion, chances are high she or he would be confronted with differing risk statements for the same exposure scenario. The lack of guidance in the TGD on the acceptable size of the Margin of Safety leads to lengthy discussions for each priority substance with regard to the acceptability of risks to man. As a consequence, the risk manager is left in doubt.

It is concluded that decision criteria for human risk characterisation need to be made explicit. This could be achieved by establishing a formal, harmonised set of default assessment factors accompanied by elaborate guidance. This was also one of the conclusions of a Workshop held under the auspices of the European Chemicals Bureau for an exchange of experience within the scope of risk assessment for new substances under Directive 92/32/EEC (Vollmer et al., 1996).

Several prerequisites can be formulated for a harmonised default set of assessment factors:

- The default set should only be applied in the absence of data which permit a more substance-specific, scientific choice. The routine use of data driven assessment factors will increase the confidence in risk assessments and encourage mechanistic research (Stevenson et al., 1995a; Dourson, 1996);
- In view of the possible differences in exposure scenarios, including groups at risk, and in the toxicological database the default set should allow for differentiation with regard to these differences;

- Harmonisation should not hamper further developments, i.e. should not be seen as standardisation for ever.

It is too easy to argue that the problem of quantifying human health risks will be solved completely by such a harmonised set of assessment factors. Unless the criteria are very rigid - which is not desirable and probably not possible in view of the uncertainties involved - there will always remain opportunities for scientific debate on the choices to make. Furthermore, it is still extremely difficult to indicate to the risk manager what the risk of a threshold substance will be in terms of the number of people at risk, their geographical distribution and the nature of the effect(s), once the minimum Margin of Safety is not reached or a no-effect level exceeded. The risk manager needs the information on exposure and the adjusted dose-response curve in order to set priorities. The latter discussion, however, is outside the scope of this report.

1.4 Goals

This report intends to be a contribution towards the development of a formal, harmonised set of assessment factors to be applied within the scope of the EU risk assessment for new and existing substances. This means that both acute, subchronic (covering both subacute and semi-chronic) and chronic exposure and occupational as well as non-occupational exposure need to be addressed. The report will first discuss the status quo with regard to the type of factors to be identified, the range of values assigned as well as the presence or absence of a scientific basis for these values (Chapter 2). Chapter 3 is a discussion on the options possible with regard to a set of default assessment factors based on the state of the art. For each factor conclusions will be presented on the scientific basis, the most likely distribution and the assumptions made. Methods for combining assessment factors are presented as well. In Chapter 4 the benchmark dose concept is discussed as well as methods of combining probability density functions for the benchmark dose and the assessment factors to arrive at the probability density function of the HLV. The various methods presented will be illustrated using an example substance in Chapter 5. Chapter 6 will summarise the major findings of this report and address recommendations.

The report will extensively refer to relevant earlier studies done in this area both in- and outside the EU-framework. It is intended as a contribution towards further discussion within the EU-framework in the first place but also explicitly taking into account the status quo and developments elsewhere. One harmonisation effort, currently underway, at the international level is the IPCS project "Harmonisation of approaches to the assessment of risk from exposure to chemicals" (Sonich-Mullin, 1995). This project considers qualitative and quantitative risk assessment methods as well as methods used for determining endpoint-specific effects. The project, among others, considers the issue how uncertainty and variability are taken into account in risk assessment. The harmonised use of terminology is a subproject undertaken in collaboration with the OECD.

2. REVIEW OF APPLIED ASSESSMENT FACTORS IN HUMAN HEALTH RISK ASSESSMENT

2.1 Introduction

This chapter presents an overview of published extrapolation methods based on the assessment factor approach for establishing Human Limit Values (HLVs). This overview is based on the report of Stevenson et al. (1995a). It is limited to a safety factor approach, thus to effects with threshold characteristics and is not meant to be exhaustive. Several terms are used for the factors introduced for the translation of NOAELs from experiments as described in the previous chapter. In this chapter, the terminology of the regarding authors is used when reference is made to the methods described. The term "assessment factor" is used when no reference is made to a specific term or method. The assessment factor can cover both extrapolation and uncertainty.

This chapter is not intended to give a complete overview of all procedures described. Some alternative extrapolation methods (e.g. the Benchmark approach) will be described in Chapter 3.

2.2 Extrapolation methodology

Acceptable Daily Intake (ADI)

Historically the so-called safety factor approach (SF-approach) was introduced in the United States in the mid-1950s in response to the legislative guideline needs in the area of food additives (Food and Drug Administration (FDA)). This approach proposed that a safe level of food additives or contaminants can be derived from a chronic NOAEL (in mg/kg of diet) from animal studies divided by a 100-fold safety factor (Lehman & Fitzhugh, 1954; ECETOC, 1995). In a slightly modified form this proposal was adopted by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) and by the Joint Meeting of Experts on Pesticides Residues (JMPR) of the WHO/FAO in 1961: the safe level was called the Acceptable Daily Intake (ADI) and was expressed in mg/kg body weight per day. The procedure involved collecting all relevant data, ascertaining the completeness of the available dataset, determining the NOAEL using the most sensitive indicator of toxicity, and applying an appropriate safety factor to derive the ADI for humans. The ADI approach is now widely used as well as the comparable Tolerable Daily Intake (TDI) approach for contaminants.

The ADI is defined as "the daily intake of a chemical which, during the entire lifetime, appears to be without appreciable risk on the basis of all known facts at the time".

The rationale for the 100-fold safety factor is reviewed by Dourson and Stara (1983). Initially, Lehman and Fitzhugh (1954) reasoned that the safety factor 100 accounts for several areas of uncertainty:

- intra (human) species variability;
- inter (animal to human) species variability;

- allowance for sensitive human populations due to illness as compared to healthy experimental animals;
- possible synergistic action of the many intentional and unintentional food additives or contaminants.

Bigwood (1973) and Lu (1979) justified the 100-fold factor on the basis of:
- differences in body size of the laboratory animals versus that of human;
- differences in food requirements varying with age, sex, muscular expenditure, and environmental conditions within species;
- differences in water balance of exchange between the body and its environment among species;
- differences in susceptibility to the toxic effect of a given contaminant among species;

Vettorazzi (1980) justified the use of the 100-fold factor by:
- differences in susceptibility between animals and humans;
- variations in sensitivities in the human population;
- the fact that the number of animals tested is small compared to the size of the human population that may be exposed;
- the difficulty in estimating the human intake;
- the possibility of synergistic action among chemicals within the human diet.

It is apparent that the factor 100 was arbitrary and is retrospectively justified in several ways. Others have attempted to interpret this factor as the product of two uncertainty factors with default values of 10, one for intra- and one for interspecies variability. But, by either interpretation, the purpose of the safety factor is to allow for uncertainties in knowledge of the toxic response of a small number of rather homogeneous laboratory animals in establishing safe doses for a heterogeneous human population (Stevenson et al., 1995a). The safety factor should not be considered immutable. When setting the ADI, various test data and judgmental factors should be considered and are needed to be taken into account, e.g., adequacy of data base, nature of the effects, age-related effects, metabolic and pharmacokinetic data, and available human data.

The overall safety factor ranges from 10 to greater than 1000, and the most commonly used factor is 100. The FDA recommends an additional factor 10 when estimating an ADI from short-term toxicity data.

Comments

Usually an arbitrary chosen safety factor of 100 is used for establishing ADIs. Retrospectively, some attempts have been made to support this factor. The procedures of the JECFA and JMPR do not force to a clear motivation for deviation from the factor 100. However, in some individual cases an expert explanation is given for the use of different factors.

Health Council of the Netherlands

The Health Council of the Netherlands (Health Council, 1985) presented an approach for the

establishment of HLVs for the general population based on the same safety factor 100. Two steps are considered: a real extrapolation step to compensate for differences in body size between the test species and human, and the application of safety factors to compensate for observation errors, possible species-specific differences in biological availability, and susceptibility between test species and human.

Two ways of extrapolation were compared:

1. On basis of body weight: This method assumes that body weight is a good relative measure of factors determining the concentration in blood.
2. On basis of caloric demands equivalent to (body weight)$^{0.75}$: Because caloric demands (basal metabolism) varies with (body weight)$^{0.75}$ interspecies adjustment factors can be calculated.

If extrapolation on basis of caloric demands is chosen, the Health Council proposed a safety factor of 30 in combination with adjustment factors. They justified this safety factor as follows:

- The safety factor for interspecies variation should be smaller, because the variation introduced by differences in body weight have been accounted for.
- The safety factor can also be rationalised otherwise. The variations and errors (interspecies, intraspecies, observation errors) are independent log normal variations. The intra- and interspecies variation are estimated to be 10: the factor for observation errors is estimated to be 3. The calculation of the safety factor is as follows:

$$\log(totalvariation) = \sqrt{(\log 10)^2 + (\log 10)^2 + (\log 3)^2} \approx 1.4925 \approx \log 31.1$$

The Health Council emphasised the role of experts in order to judge the quality of the data base, to formulate the toxicological starting-points, and to establish the safety factors. The absence of relevant data should be taken into account.

At present, the safety factor approach is still in use. However, the Health Council Committee believes that limits should be derived using a method which makes systematic use of data on the relationship between exposure and response (Health Council, 1996).

Comments

This method is not commonly used in The Netherlands. An advantage of this method is the theoretical distinction between extrapolation and uncertainty. However, in practice, such a distinction is not always possible. Scaling on caloric demands is considered to be preferable above scaling on body weight, both on theoretical grounds and because of the similarity with setting respiratory HLVs from inhalation studies. This method is still based on the arbitrary factor of 100 used in the derivation of the ADI. The rationale for the formula is not given but lognormal uncertainty measures can indeed be summed logarithmically (Slob, 1994).

EPA: Reference dose (RfD)

In 1988, also the US-EPA adopted the ADI approach in its regulatory measures against environmental pollution, however, with a number of modifications. Instead of the terms ADI and safety factor (SF), the terms "reference dose" (RfD) and uncertainty factor (UF) are used, respectively. The RfD is derived from the NOAEL by consistent application of generally one order-of-magnitude UFs that reflect the various types of data set used to estimate RfDs. UFs generally consist of:

- a 10-fold factor for valid human data;
- a 10-fold factor for data from long term animal studies;
- a 10-fold factor from less than chronic studies in animals;
- a 10-fold factor from a LOAEL in the absence of NOAEL.

In addition, a modifying factor (MF) ranging from less than 1 to up to 10 is applied when the data base includes for example a very large number of animals per dose level (MF less than 1), or lacking in biological studies (MF 1 to 10).

Method of Calabrese and Gilbert

The modifications in uncertainty as used by the EPA, i.e. a factor 10 for interspecies and 10 for intraspecies variations were also proposed by Calabrese and Gilbert (1993). They suggested modifications of UFs by the lack of total independence of these factors. The interspecies UF is generally recognized as providing an extrapolation from the average animal to an average individual assuming that humans may be 10-fold more sensitive. The intraspecies UF assumes that most human responses fall within approximately a 10-fold range. Calabrese and Gilbert stated that, given this assumption, the application of a 10-fold intraspecies UF should begin with the average human and extend to cover the higher risk segments of the population. Consequently, an UF of 5 would be expected to protect most humans. However, an UF 10 is indicated when the HLV is based on an occupational epidemiological study since this type of study does not consider the most sensitive humans.

The factor used when a semi-chronic study is used as starting point, incorporates an age-dependent factor which is comparable in some respects to the age-dependent factor in the intraspecies uncertainty factor. The age component in the intraspecies uncertainty factor concerns the age differential response over the entire life-time span while the age differential of the less-than-lifetime uncertainty factor concerns only the age-related differences from the end of the study to the end of the normal life span. High susceptibility is not exclusive for young animals. In certain circumstances susceptibility may be greater in adulthood than in the young and may further increase in elderly animals. Assuming that age differences account for 50% of the intraspecies variation, they stated that this factor could be reasonably apportioned as 60% for prior to weaning and 40% for after weaning. If a 24-month rodent exposure accounts for 40% of age effects, then it would be reasonable to reduce the age component of the intraspecies variation by the proportion described to age. If the intraspecies factor is 5 as recommended, then this reduces the factor to 4.

Table 1 provides the uncertainty factors recommended by Calabrese and Gilbert in the light of the above considerations.

Table 1: Recommended modifications in current uncertainty factors on the concept of independence and interdependence of uncertainty factors (Calabrese & Gilbert, 1993)

Extrapolation step	Uncertainty factor
Animal to human	10
Interindividual	
less-than-lifetime animal study	5
animal study with normal experimental lifetime	4
occupational epidemiological study	10
environmental epidemiological study (normal lifespan)	5
LOAEL instead of NOAEL	10
Less-than-lifetime	10

Comments:
The notion of interdependence between the factors is considered a valuable one. However, it is recommended to examine the interdependence of the factors in more detail before applying the concept in risk assessment procedures. The assumption of total independence of assessment factors should be recognised as a "worst-case" approach.

Method of Renwick
The approach proposed by Renwick (1991, 1993a,b) is also based on the 100-fold factor used to derive ADIs. It attempts to give a scientific basis to the default values of 10 for the interspecies and inter-individual differences. Renwick has proposed the division of each of these UFs into subfactors to allow for separate evaluations of differences in toxicokinetics and toxicodynamics. The advantage to such a subdivision is that components of these UFs can be addressed where data are available (e.g., if data exist to show similar toxicokinetic handling of a given chemical between laboratory animals and humans, then only an interspecies extrapolation factor would be needed to account for differences in toxicodynamics).
Renwick examined the relative magnitude of toxicokinetic and toxicodynamic variations between and within species in detail. He found that toxicokinetic differences were generally greater than toxicodynamic differences resulting in the proposal that the 10-fold uncertainty factors (for inter- and intraspecies) should, by default, be subdivided into factors of 4 for kinetics and 2.5 for dynamics. Factors up to 10 should be applied to critical effects as teratogenicity and non-genotoxic carcinogenicity. In cases where a reversible lesion (e.g. hyperplasia) is believed to be a precursor for a severe irreversible change, the NOAEL for each lesion should be used to calculate an ADI, using appropriate factors. The lowest value should be chosen as ADI. The rationale for an extra factor is the potential seriousness of any unrecognised aspect that has not been taken into account in the safety evaluation. This factor is not data-derived but based on scientific judgement. A factor for the adequacy of the overall database has been introduced to consider aspects other than the pivotal study and the determination of the NOAEL. A value of 1 assumes an adequate database consistent with national or international guidelines.

Comments
Renwick also uses the arbitrary factor of 100 as basis. The main feature of this approach is that kinetic and dynamic aspects are distinguished in inter- and intraspecies differences. This

offers the possibility to incorporate mechanistic information on these aspects in the establishment of the factors as shown for pharmaceuticals by Naumann et al. (1997), provided sufficient data are available. It should be remarked that the proposed default values are derived from only limited data. Renwick also uses expert judgement for the derivation of HLVs, but with an attempt for transparency and clear motivations. The International Programme on Chemical Safety (IPCS, 1994) has adopted the principles set forth by Renwick (1991, 1993a,b), but has suggested that while the UF for interspecies extrapolation be subdivided unequally into 4-fold and 2.5-fold, the UF for intraspecies extrapolation should be split evenly (3.16-fold for both kinetics and dynamics)(ECETOC, 1995).

Approach described by Lewis/Lynch/Nikiforov

In 1990 Lewis and his colleagues (Lewis et al., 1990) undertook revision of the long established practices, with the goal of introducing flexibility such that both new information and expert judgement could be readily incorporated. The Lewis-Lynch-Nikiforov (LLN)-method, and its refinements, are extensions of established principles and procedures. LLN guides the data evaluator to adjust experimentally determined 'no-effect' (or 'minimum effect') levels from laboratory animal studies, while taking the following aspects into account:

- differences between laboratory animals and humans;
- differences between experimental conditions and actual or anticipated human exposures;
- the sensitivity of the exposed human populations;
- weight of evidence indicating an actual human health hazard;
- quality of the experimental information base;
- uncertainties in extrapolating from animals to humans;
- potency of the toxic agent.

If suitable human data are not available, the HLV is estimated from laboratory results, using the following algorithm:

$$HLV = \frac{NOAEL_{animal}[S]}{[I][R][Q_1][Q_2][Q_3][U][c]}$$

The terms are described in table 2.

An aggregate adjustment of about 250 is typical and is approaching the practical maximum. The theoretical maximum adjustment value is 100,000. By application of the factors Q_1-Q_3 and U this method intends to separate scientific judgements from policy/value judgements. Factors for data quality should reflect the completeness and suitability of the available information. According to Lewis et al. (1990) there are three distinguishing features of the LLN-approach:

- careful discrimination among the adjustments;
- discrimination between "best estimates" of the correct adjustments for [S], [I], and [R] and the overall uncertainty;
- securing scientific consensus on the adjustment values.

Table 2: Adjustment factors of the Lewis/Lynch/Nikiforov model (1990)

AF[a]	Description	Range of values	Most likely value	Default value
S	Scaling factor to account for known quantitative differences between species and between experimental conditions and those likely to be encountered by humans	> 0	NS[b]	1
I	Intraspecies variability	1-10	1-3[c]	10
R	Interspecies extrapolation	> 0-10	NS	10
Q_1	Degree of certainty that the critical effect observed in laboratory animals is relevant to humans	0.1-1		1
Q_2	Subchronic to chronic extrapolation	1-10	1-3	10
Q_3	LOAEL to NOAEL extrapolation	NS	2	10
U	Accounts for residual uncertainty in estimates of S, I, and R	1-10	NS	10
C	A non-scientific, judgmental "safety" factor	1-10	≤ 3	1

[a] AF, Adjustment Factor
[b] NS, Not Stated by authors
[c] Most likely value based on study of high quality

Comments

The LLN-method discriminates to a large extent and therefore, this method is valuable to give guidance for which factors one should account. However, some remarks can be made:

- In practice it will not be possible to distinguish all these factors;
- It is likely that the scaling factor S influences the adjustment for interspecies differences R. However, the authors give no guidance for the choice of R when a value different from 1 is introduced for S;
- The factors I and R are not consistently distinguished. R is said to account for a possible wider range of susceptibility among individuals. However, this belongs to the factor I;
- The value Q_1 seems superfluous, because for the selection of the NOAEL it should be considered whether the critical effect is relevant for humans;
- It is not clear how the value of residual uncertainty in the estimates of S, I, and R can be determined;
- Introduction of a non-scientific "safety" factor C is not in accordance with the establishment of an HLV.
- One should be aware that some factors may not be independent of each other.

Method used in advising the Dutch Competent Authority for occupational risk assessment

The method used by TNO in advising the Dutch Competent Authority regarding the risk assessment of workers for New and Existing Chemicals is developed, using literature, e.g. as mentioned in the EU-Technical Guidance Document (EC, 1996a), supplemented with information from studies of Stevenson et al. (1995a, b) and with a guidance document for setting Acceptable Operator Exposure Levels (AOEL) (Anonymous, 1995). The method is used for setting Health-Based-Occupational-Reference-Values (HBORVs). The HBORV is defined as the maximum amount of a substance to which a worker can be exposed without adverse health effects being expected. For the time being a starting point is that workers may

be exposed predominantly, but not exclusively, by two routes: dermally and by inhalation. HBORVs are assessed for both routes separately and for every effect (if possible) as defined in the TGD (Hakkert et al., 1996).

The hazard assessment serves as starting point for the derivation of an HBORV. To translate the selected NOAEL into the HBORV assessment factors, compensating for uncertainties inherent to extrapolation of experimental (animal) data to a given human situation and for uncertainties in the toxicological data base, have to be applied. The assessment factors must be derived considering the toxicity profile of the substance. If no conclusions can be drawn a default factor will be used. The default factors are presented in table 3.

Table 3: Assessment factors applied for the calculation of HBORVs (Hakkert et al., 1996)

Aspect	Assessment factor (default value)
Interspecies differences	
- mouse	7[a] x 3
- rat	4[a] x 3
- rabbit	2.4[a] x 3
- dog	1.4[a] x 3
Intraspecies differences	3[c]
Differences between experimental conditions and exposure pattern for workers:	
- chronic to chronic exposure	1
- subacute to semi-chronic exposure	10[b]
- semi-chronic to chronic exposure	10[b]
- other aspects	1
Type of critical effect	1
Dose-response curve	1
Confidence of the database	1
Route-to-route	No default: if no relevant data on toxicokinetics and metabolism are available, worst case assumptions with respect to absorption% have to be made.

[a] this is a calculated adjustment factor, allowing for differences in basal metabolic rate (proportional to the 0.75 power of body weight)
[b] the actual factor applied is often lower, and is derived from the toxicological profile of the test substance
[c] a factor of 3 is used for workers, a factor of 10 for the general population

The overall factor is established by multiplication of the separate factors, unless the data indicate another method to be used. It is stated, that one should be aware that in practice, it will be possible to distinguish all above mentioned factors, and that some factors are not independent of each other. Therefore, straightforward multiplication may lead to unreasonable high factors. Discussion and weighing of individual factors is essential to establish a reliable and justifiable overall assessment factor (Hakkert et al., 1996).

Comments

This approach discriminates factors to a large extent in order to distinguish between the single adjustments and to separate best estimates from uncertainty. Discrimination forces to a rational choice and to greater transparency, and invites to apply scientific consideration. However, multiplication can result in an unrealistically high overall factor. Besides, in

practice, it is not possible to distinguish all above mentioned factors. One should be aware that some factors are not independent of each other. The overall factor will be the result of a number of different considerations which have to be made transparent.

ECETOC approach

The approach recommended by the ECETOC (ECETOC, 1995) to derive a scientific estimate of a human no adverse effect level (which is referred to in their report as the Predicted No-Adverse-Effect Level (PNAEL)) distinguishes three stages. In each of the stages an estimate is made of the most likely value of the factor described. At the end of the process the factors are multiplied together and the resultant number is used to derive the human PNAEL. The three stages are:

1. Application of a scientifically derived 'adjustment factor' to the NOAEL/LOAEL of the critical effect established in the pivotal study (summarised in table 4). If the data base is inadequate then human PNAELs cannot be derived scientifically and the recommended scheme cannot be developed further.
2. Application of an 'uncertainty factor' to the PNAEL to take into account the degree of scientific uncertainty involved. The following degrees of confidence in the human PNAEL are suggested as a guide based on several different required conditions:
   - high degree of confidence: 1
   - medium degree of confidence: 1-2
   - low degree of confidence: larger uncertainty factor.
3. Application of a non-scientific 'safety factor' taking into account political, socio-economic or risk perception factors. Non-scientific safety factors are intended to account for :
   - political aspects;
   - socio-economic aspects (cost-benefit considerations);
   - risk perception factors (the nature of the effect may justify the use of an additional factor).

   No additional factors should be used if conservative (default) values were used in the stages 1 and 2.

An important feature of this approach is the need to establish the route and duration of exposure to which the PNAEL refers before attempting to derive factors, since these may vary for different routes or exposure duration. For each element of the approach, ranges and default values for the numerical factors involved are recommended (ECETOC, 1995).

Comments

Like the previous method, this approach discriminates factors to a large extent in order to distinguish between the single adjustments and to separate best estimates from uncertainty. This method also gives guidance for setting occupational and non-occupational limit values. Discrimination forces to a rational choice and to greater clarity, and invites to apply scientific consideration. However, the ECETOC approach does not mention the establishment of the overall factor. Furthermore, ECETOC gives guidance to the extrapolation-step. Although they mention that all discriminated aspects introduce uncertainties, they don't give guidance how to account for this. Finally, it can be questioned whether a non-scientific factor should be

discussed in a scientific risk assessment.

Table 4: Adjustment factors (recommended default factors) for use in deriving human PNAELs from human or animal NOAELs/LOAELs (ECETOC, 1995)

Element	Factor (default value)	Additional information
1. Short-term repeated/ subchronic/ chronic extrapolation:		
- short-term to subchronic	3	
- subchronic to chronic	2-3	
2. LOAEL to NOAEL extrapolation	3	Extent and severity of effects may justify the use of another (higher/ lower) factor.
3. Route-to-route extrapolation	no default	Conversion factors must be calculated for each individual situation, making appropriate assumptions about body weight, minute volume, and percentage absorption.
4. Interspecies extrapolation		
- oral route	4	Based on caloric demands (4: default for the rat (Body weight 0.250 kg)).
- inhalation route	1	For substances with local effects on case-by case basis. Concerning toxicodynamic aspects: factor >1 only when human is considered to be more sensitive than the most sensitive species otherwise no factor.
5. Intraspecies extrapolation		
- general population	3	
- occupational population	2	

3. QUANTIFICATION OF ASSESSMENT FACTORS

3.1 Introduction

As shown in chapter 2, the approaches described in the previous chapters share many of the same underlying assumptions, judgements on critical effect, and choices of assessment factors or margins of safety. The approaches typically rely on existing human epidemiological and/or animal laboratory data. Scientists review all toxicity data, judge what constitutes an adverse effect, and determine the critical effect. Subsequently, the appropriate assessment (safety, extrapolation, or uncertainty) factors are applied to the NOAEL or LOAEL for the critical effect to account for the lack of data and inherent uncertainty in the extrapolations. Alternatively, in the case of the MOS approach, the magnitude by which the NOAEL or LOAEL exceeds the estimated exposure will be considered in view of several uncertainty parameters (e.g. interspecies differences, intraspecies differences). These parameters can be compared quantitatively. The description of typical assessment and modifying factors in the development of an HLV for the different approaches is summarised in Table 5.

The ideal method for the establishment of HLVs should fulfil a number of requirements:

- Factors for extrapolation should be based on scientific data.
- Extrapolation includes: short-term to long-term, interspecies, intraspecies, differences in exposure conditions between experimental or observational and the human situation for which the HLV is developed, and route-to-route;
- The system should give possibilities to differentiate for: severity of effects, dose-response curve, and data on kinetics or dynamics;
- Account should be made on the adequacy of the selected study and the completeness of the database;
- The choice of the factors should be motivated to assure a consistent application.
- The method should correct for worst case combination of assessment factors such as occurs with multiplication.

In Table 5 attention is also paid to these requirements.

Based on the evaluations made in the previous chapters it is clear that only a few approaches rely on scientific data. The factors applied will depend on the selected pivotal study and the critical effect. All methods use experts for judgement of the underlying toxicological database (completeness, relevance, and adequacy of the studies) and the severity of effect. The precision of the selected NOAEL is determined mainly by the sensitivity and relevance of the toxicological endpoint, the group size studied, and the increment between doses. Therefore, the selected NOAEL may be a poor estimate of the true but unknown NAEL.

Most approaches using assessment factors basically rely on the arbitrary 100-fold factor used to derive the ADIs. Example 1 in Chapter 5 illustrates this approach. Other approaches discussed in this and the next chapter will also be illustrated in chapter 5, using the same dataset. There are no scientific grounds for the factor 100. Some scientists interpret the

absence of widespread effects in the exposed human population as evidence of the adequacy of this factor. Several attempts were made to justify the subdivision in a factor 10 for interspecies and a factor 10 for intraspecies. Some apply modifying factors, on the basis of body weight, caloric demands, confidence in database, lack of independence of factors, variations in exposure circumstances, or differences between species in toxicokinetics and toxicodynamics. It is concluded that a scientific justification for the size of the factors used for intra- and interspecies differences is lacking. Several studies have been described concerning interspecies scaling, based on either a mechanistic approach or an empirical approach. The incorporation of this knowledge in a procedure for the establishment of HLVs deserves further study.

The LLN, TNO, and ECETOC-method, discriminate factors to a large extent in order to distinguish between the single adjustments and to separate best estimates from uncertainty. Discrimination enhances the clarity, forces a rational choice, and invites to substantiate the adjustment factors. However, multiplication can result in an unrealistically high overall factor. ECETOC does not mention the uncertainty and the establishment of the overall factor. In the TNO approach it is mentioned that weighing of the individual factors is essential because the overall factor is the result of a number of different considerations. The LLN-method introduces residual factors which do not belong to the establishment of HLVs or that cannot be quantified. Calabrese and Gilbert (1993) indicated that the applied assessment factors for a.o. inter- and intraspecies variation are not fully independent. Therefore, they proposed a modifying factor to incorporate the assumption of the lack of total independence of assessment factors in the establishment of the overall assessment factor.

The choice of the applied assessment factors is seldom motivated. Transparency is essential to assure a consistent application. Besides, if more detailed information on a specific situation would become available the assessment of an HLV may be refined. Therefore, the remaining uncertainties should be clarified. An approach which lends much more credibility to the use of the assessment factors is the investigation of the probabilistic nature of assessment factors (Dourson, 1996; see also chapter 4). The expression of the probability of the numerical value of each uncertainty factor can be based on the actual toxicity data on groups of chemicals for which HLVs have been developed . The most likely distribution for each assessment factor will be log-normal. For HLVs that have more than one area of uncertainty, the respective individual distributions can be multiplied using Monte Carlo techniques to develop an overall distribution reflecting total uncertainty which is then applied to the NOAEL or LOAEL of the pivotal study to develop a probabilistic HLV.

In the following, specific attention will be paid towards the scientific information on which assessment factors are based and the quantification of the overall assessment factor. The possibilities for the application of the probabilistic approach for the derivation of a more rationalised assessment factor will be further investigated in chapters 3 and 4.

Table 5: Evaluation of default assessment factors used or suggested for the establishment of HLVs

Assessment factors	ADI JECFA JMPR	HC	EPA (RfD)	Calabrese & Gilbert	Renwick IPCS	LLN	TNO	MOS	ECETOC
Intraspecies								- [C]	
non-occupational	10	10	10	4-10	+	+	10		10
- toxicokinetics					0-2.5[E]	1-10[F]			
- toxicodynamics					0-4.0[E]				
occupational	-	-	-	-	-	+	3		2
Interspecies	10	ND[I]	10	10	+	+	+		+
toxicokinetics					0-2.5	1-10[F]		ND	
toxicodynamics					0-4.0				
oral route							A[G] x 3		4
inhalation route							3	-	1
Duration of exposure	10 [H]	10	<10	10	-	+	+	ND	+
subacute/subchronic							1-10[F]		3
subchronic/chronic						1-10[F]	1-10[F]		2-3
other aspects							1		
LOAEL to NAEL	-	-	<10	[L]	-	2-10[F]	ND	ND	2-3
Route-to-route	-	-	-	[L]	-	-	ND	ND	ND
Type of critical effect	-	-	-	[L]	-	-	1	ND	-
Dose-response curve	-	-	-	[L]	-	-	1	ND	-
Confidence in database	-	-	1-100	[L]	1-10	+	1	ND	1: high 1-2:medium ND: low
Non-scientific factor	-	-	-	[L]	-	1[F] -10	-	ND	+
Modifying factor	-	-	0-10[K]	[L]	1-10	+	+	-	+
Motivation for choice of factors	-	-	-	[L]	+/-	+/-	+/-	?	+/-
Overall factor	mult.	other (see 2.2)	mult.	[L]	mult.	mult.	mult.	-	mult.

+	method accounts for
-	method does not account for
E)	IPCS recommended that the interindividual toxicokinetic and toxicodynamic default values should be 3.16 and 3.16.
F)	default factor
G)	this is a calculated adjustment factor, allowing for the differences in metabolic size (mouse: 7, rat: 4, rabbit: 2.4, dog: 1.4)
H)	the additional assessment factor for duration of exposure for establishment of the ADI has been recommended by FDA.
I)	scaling on body weight or caloric demands
K)	factor accounts for type of critical effect
L)	Calbrese and Gilbert do not describe a full method for the derivation of HLVs
mult.	multiplication of different factors to establish the overall assessment factor
ND	no default value, based on a case-by-case determination (expert judgement and scientific information)

3.2 Quantification of assessment factors: use of historical data

3.2.1 Introduction

In this section the quantification of assessment factors is addressed. As far as possible, this quantification will be based on historical data (i.e. NOAELs). The uncertainty in the value of the assessment factors obtained will be taken into account by describing their entire distribution (see also Chapter 4). Possible default values for the assessment factors will be derived from the higher end of these distributions for comparison with current worst case default values. The data of Kramer et al. (1996) confirm that assuming a lognormal distribution for each assessment factor is reasonable, based on the observation that ratios of NOAELs appear to be lognormally distributed (see also section 3.2.2).

3.2.2 Interspecies extrapolation

For extrapolation of data from animal studies to humans account should be taken of species-specific differences between animals and humans. These interspecies differences can be divided in differences in metabolic size and remaining species-specific differences. To account for differences in metabolic size three methods are used in practice: extrapolation based on body weight, surface area, and caloric demand. These methods can be described by an allometric equation: for that purpose body weight has to be raised to the power 1, 0.67, and 0.75, respectively.

Based on theoretical grounds, scaling on the basis of surface area or caloric demand can be considered more appropriate compared to extrapolation based on body weight (Gevel and Hakkert, 1997). Experimental work did not answer the question which of these two methods is the most correct. However, based on theoretical grounds the Health Council of the Netherlands (Health Council, 1985), TNO (Hakkert et al., 1996) and Kalberlah et al. (1997) consider the extrapolation based on caloric demands (the 0.75 power of body weight) as preferable above scaling on body weight.

Insert: Scaling according to caloric demand

Scaling according to caloric demand means that a comparable dose rate in milligram per kilogram body weight dose for the average person (70 kg) is equal to the rat (0.25 kg) dose rate divided by an interspecies factor which is equal to 70/0.25 to the power 0.25 (= 4).

In formulae:

Equivalent doses:

$mg_{rat}/mg_{human} = (kg_{rat}/kg_{human})^{0.75}$

Expressed as dose rates:

$mg_{rat}.kg_{rat}^{-1}/mg_{human}.kg_{human}^{-1} = (kg_{rat}/kg_{human})^{0.75} . (kg_{rat}/kg_{human})^{-1}$

Or:

$dose\ rate_{human} = dose\ rate_{rat}/(kg_{human}/kg_{rat})^{0.25} = dose\ rate_{rat}/4$

To express the dose in mg/kg body weight (to the power 1) assessment factors are calculated. The size of these factors are e.g. 7 for mice (25 g), 4 for rats (250 g), and 1.4 for dogs (15 kg),

etc. for the extrapolation from the test species to humans (see insert on page 30). For inhalation NOAELs for systemic effects no correction is made for differences in metabolic size, because extrapolation is already based on toxicological equivalence of a concentration of a substance in the air; animals and humans breath at a rate depending on their caloric requirements (Hakkert et al, 1996).

To account for the remaining interspecies uncertainties usually a default factor is used. In theory, the remaining uncertainty could be assessed by comparing NOAELs in test animals with estimates of human NOAELs. However, in practice, such an assessment must rely on data from studies derived experimentally for the same substance in different animal species because human data are lacking. The degree of remaining interspecies uncertainty may be obtained by examining the differences (ratios) of the NOAELs established for the same substance in different species. The actual uncertainty in extrapolating from animals to humans is likely to be at least as large as the uncertainty in extrapolating among mice, rats, and dogs.

For the purpose of assessing the remaining interspecies uncertainty, data from the TNO database (including Pesticide dossiers, Existing chemical dossiers, IPCS Environmental Health Criteria documents, JMPR evaluations, and public literature), including 184 substances tested in different species and via different routes, were analysed. NOAELs were selected from studies with mice, rats, and dogs exposed to the same test substance via the same route and with the same duration of exposure. In order to increase the comparability of the different animal experiments with respect to their duration of exposure, two categories of exposure duration were defined: subacute and (semi-)chronic. The definition of the categories is species-specific, partly depending on their maximum lifetime. For mouse, rat, and dog the categories are summarised in Table 6.

Table 6: Exposure duration categories for different species

Exposure duration	mouse (days)	rat (days)	dog (days)
Subacute	21-50	21-50	28-90
(Semi-)chronic	90-730	90-730	365-730

If within one category (same exposure duration, same test substance, and same species) more NOAELs were available, the lowest NOAEL has been used for the selection. NOAELs based on carcinogenicity have been left out of consideration. The oral NOAELs were adjusted to account for differences in metabolic size (as described above). In order to increase the comparability of the derived factors to the actual uncertainty (animal to human), the ratios were calculated by dividing the NOAELs derived in the smaller animal by the NOAEL derived in the larger animal. The following ratios were calculated: $NOAEL_{mice}/NOAEL_{rat}$, $NOAEL_{mice}/NOAEL_{dog}$, and $NOAEL_{rat}/NOAEL_{dog}$.

The ratios (both adjusted and unadjusted for metabolic size) were evaluated by examining their distributions. Table 7 presents the number of ratios (N), the geometric means (GM), the

geometric standard deviations (GSD), and the 90 and 95 percentiles of the distributions of the ratios. Percentiles are calculated from the GM and the GSD as shown on page 11.

With respect to dermal toxicity, insufficient relevant data were available. For respiratory toxicity data only the $NOAEL_{mice}$ - $NOAEL_{rat}$ ratios were analysed: with respect to the other ratios insufficient data were available to be statistically analysed.

These data suggest that the distribution of the ratios can be described well by a log-normal distribution. If the interspecies differences would depend only on the differences in metabolic size and if the method used were perfect (and if the NOAELs contained no errors), both the geometric mean and the geometric standard deviation of the distribution of the ratios would be

Table 7: Distribution parameters derived from the NOAEL ratios

Ratio	N	GM	GSD	P_{90}	P_{95}
$NOAEL_{rat}$ / $NOAEL_{dog}$ (oral, unadjusted)	63	1.3	5.1	10.4	18.8
$NOAEL_{rat}$ / $NOAEL_{dog}$ (oral, adjusted)	63	0.5	5.1	3.6	6.6
$NOAEL_{mouse}$ / $NOAEL_{rat}$ (oral, unadjusted)	67	4.2	5.7	39.3	73.9
$NOAEL_{mouse}$ / $NOAEL_{rat}$ (oral, adjusted)	67	2.4	5.7	22.5	42.2
$NOAEL_{mouse}$ / $NOAEL_{dog}$ (oral, unadjusted)	40	6.4	6.1	64.7	124.6
$NOAEL_{mouse}$ / $NOAEL_{dog}$ (oral, adjusted)	40	1.3	6.1	12.9	24.9
$NOAEL_{mouse}$ / $NOAEL_{rat}$ (respiratory)	21	3.1	7.8	43.6	91.8

N = number of ratios
GM = geometric mean
GSD = geometric standard deviation
P_{90} = 90th percentile
P_{95} = 95th percentile

unity. The geometric means of the ratios of adjusted NOAELs are closer to one than the means of the unadjusted NOAELs which supports the idea of accounting for the differences in metabolic size (scaling based on caloric demands). As an approximation of the remaining uncertainty in the extrapolation from animals to humans the mean of the distribution parameters will be used; a geometric mean of approximately 1 and a geometric standard deviation of 6. It should be noted that it is possible that the NOAELs were established based on different critical effects. Other differences not corrected for exist, such as in strain and substance. In reality, the variation of the distributions will therefore be smaller.

In summary: Based on theoretical grounds, and supported by the analysis given in Table 7, scaling on the basis of surface area or caloric demand to adjust oral NOAELs for metabolic size can be considered more appropriate compared to extrapolation based on body weight. The assessment factor accounting for the remaining uncertainty in the extrapolation from animals to humans may be characterised as approximately lognormally distributed with a geometric mean of about 1 and a geometric standard deviation of 6 (Figure 1). Based on this distribution, default values for the 90-, 95- and 99-percentiles can be calculated to be 10, 19 and 65, respectively.

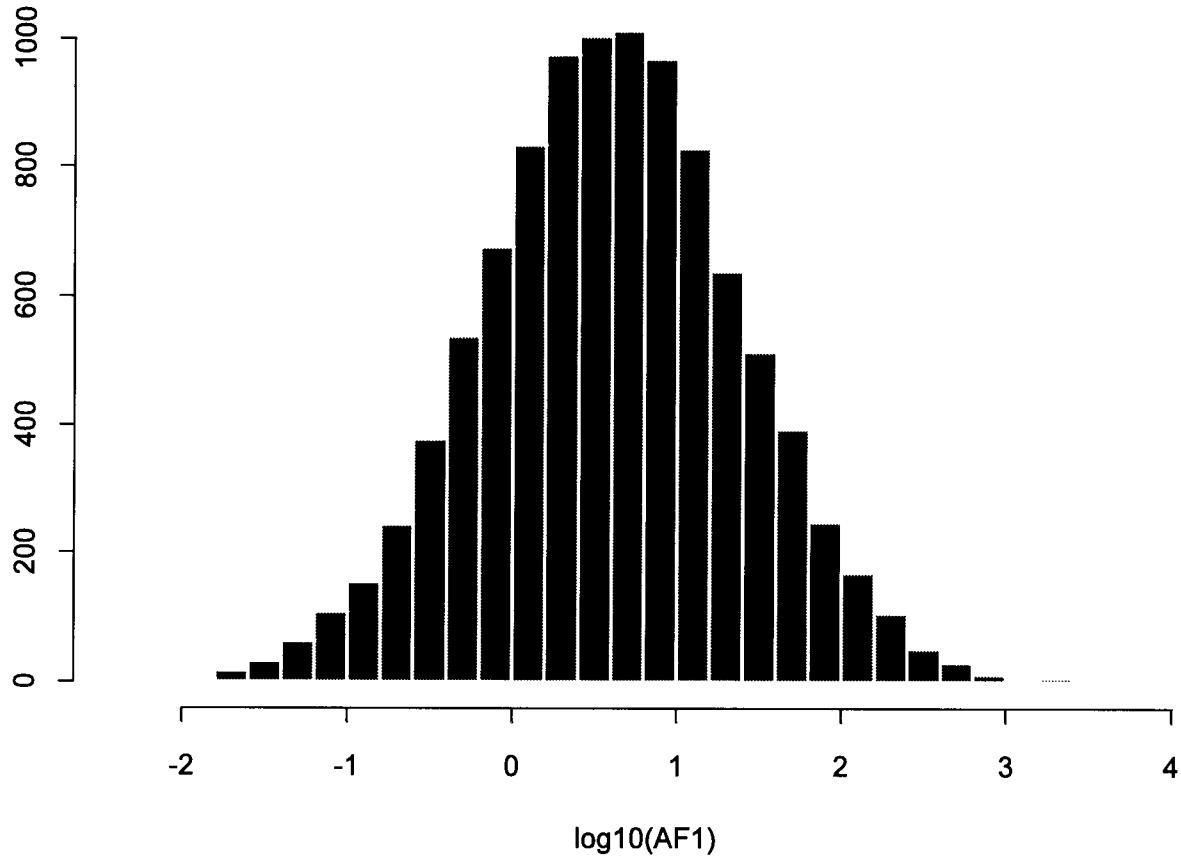

Fig.1: Distribution of the interspecies assessment factor, adjusted for metabolic size. In this distribution the often applied default factor of 12 (adjustment for metabolic size 4, remaining uncertainty 3) coincides with the 73th percentile

3.2.3 Intraspecies extrapolation

The response of humans to exposure of xenobiotic compounds may vary due to a number of biological factors, such as age, sex, genetic composition and nutritional status. To account for interindividual human variation a factor of ten for the extrapolation from the average to the sensitive human being is generally assumed to be appropriate for deriving HLVs. Though many publications indicate that human variability may grossly exceed a factor of 10, Calabrese (1995) reached the conclusion that in most cases a factor of ten would be sufficient to protect the majority (up to 80-95%) of the human population against adverse health effects. Exceptions may be due to e.g. increased susceptibility due to serious illness. Also genetic polymorphisms of metabolising phase I and phase II enzymes may cause a large variation in human responses (Daly et al., 1993).

In only a few publications an attempt was made to investigate the human interindividual variation by data analysis. Hattis et al. (1987) investigated the variation in pharmacokinetic behaviour of 49 pharmaceuticals in healthy adults. Depending on the pharmacokinetic parameter studied (elimination half-life, area under the curve, peak concentration), a factor of ten would account for 2.5-9 standard deviations from the geometric mean. From this analysis it appears that a factor of ten will be sufficient for pharmacokinetic variation. Reanalysis of

the data of Hattis et al. showed that for the half-life time the variation between individuals was quite small. Defining the intraspecies factor as the ratio of the P_{50} and P_{05} resulted in a factor of 1.4 (Schaddelee, 1997). However, one should take into account that (i) variation also exist in pharmacodynamics and (ii) that only data of healthy volunteers were available so that the real variability in the human population is underestimated. Renwick (1993a,b) analysed interindividual differences of healthy volunteers and patients by comparing the maximum and mean values of pharmacokinetic parameters and the minimum and mean values of pharmacodynamic parameters. Based on this analysis he proposed to subdivide the factor of ten into a factor of four for pharmacokinetic differences and a factor of 2.5 for pharmacodynamic differences. Re-analysis of the Renwick data by using distributions instead of ratios max/mean and min/mean gave comparable results (Schaddelee, 1997).

The results of Renwick's analysis have been adopted by the IPCS (IPCS, 1994). However, rather than using default factors which comply with an arbitrary chosen default factor of ten, it would be better to use toxicity profile derived pharmacokinetic and pharmacodynamic factors. This would require further and better data analysis. Data availability, however, may be a problem here. By performing data analysis one should take experimental errors into account since these may be substantial. The latter has not been considered in Renwick's analysis.

Based on an analysis of the available human data, Kalberlah et al. (1997) propose an intraspecies factor of 25 for the general population, composed of a factor of 8 accounting for toxicokinetic variation and enzyme polymorphisms, and a factor of 3 accounting for toxicodynamic variation. For workers this factor is reduced and a total factor of 5 is considered to account for both inter and intraspecies variation (after adjustment for differences in metabolic size). They claim that only in a few cases the sensitivity of special groups at risk will exceed these ranges. As the authors admit, it can be noted that this proposal is based on an overall impression based on several substance-specific examples. The combined factor for workers accounting for both inter and intraspecies variation is not adequately explained.

In summary: currently no proposal for a database-derived distribution of the intraspecies factor can be made. Therefore for the time being it can be considered appropriate to remain consistent with the traditional default value of 10 and to assume that this value will protect the majority of the general human population. For workers one could remain consistent with the traditional default value of 3.

3.2.4 Subchronic to chronic extrapolation

When only subchronic (subacute and semi-chronic) data are available an extra assessment factor (usually ten) is currently used to extrapolate to chronic exposure. For the distribution of the extrapolation factor several studies comparing NOAELs from chronic and subchronic studies appear relevant (Weil and McCollister, 1963; McNamara, 1976; Rulis and Hattan, 1985; Kramer et al., 1995; Nessel et al., 1995; Kalberlah et al., 1997). Tables 8 and 9 summarise the oral data.

Table 8: Semi-chronic to chronic oral NOAEL-ratios[1]

N	GM	GSD	P_{90}	P_{95}	Semi-chronic exposure period	Chronic exposure period	Species	Reference
33	2.2	2.3	6.4	8.7	30-210 days	2 years	rats	Weil et al., 1963
41	1.0	1.7	2.0	2.5	not specified	not specified	rats, dogs[2]	McNamara, 1976
20	1.9	3.0	8.0	12.0	< 200 days	> 200 days	various	Rulis and Hattan, 1985
149	1.7	5.6	15.4	28.9	10-26 weeks	1-2 years	various	Kramer et al., 1995
23	2.0	1.8	4.2	5.1	90 days	2 years	rodents[3]	Nessel et al., 1995
9	2.4	1.3	3.4	3.7	90 days	1-2 years	mice	Kalberlah et al., 1997[4]
11	1.7	1.8	3.6	4.5	90 days	1-2 years	rats	Kalberlah et al., 1997[4]
20	2.0	2.4	6.1	8.4	90 days	1-2 years	mice + rats	Kalberlah et al., 1997[4]
21	1.7	1.7	3.3	4.1	90 days	2 years	mice	Kalberlah et al., 1997[5]
22	2.5	1.9	5.7	7.2	90 days	2 years	rats	Kalberlah et al., 1997[5]

[1] N = number of ratios, GM = geometric mean, , GSD = geometric standard deviation, P_{90} = 90-percentile, P_{95} = 95-percentile (percentiles calculated from the GM and GSD)
[2] 39 rat pairs, 2 dog pairs
[3] matched pairs
[4] Industry data from 13 agrochemicals
[5] Data from the US National Toxicology Program

Table 9: Subacute to chronic oral NOAEL-ratios[1]

N	GM	GSD	P_{90}	P_{95}	Subacute exposure period	Chronic exposure period	Species	Reference
57	6.5	3.5	32	51	3-6 weeks	1-2 years	various	Kramer et al., 1995
37	3.4	5.7	32	60	3-6 weeks	1-2 years	rats	Kramer et al., 1995[2]
20	3.1	1.9	7.0	8.9	14 days	2 years	mice	Kalberlah et al., 1997[3]
26	3.9	2.2	10.7	14.3	14 days	2 years	rats	Kalberlah et al., 1997[4]

[1] N = number of ratios, GM = geometric mean, , GSD = geometric standard deviation, P_{90} = 90-percentile, P_{95} = 95-percentile (percentiles calculated from the GM and GSD)
[2] Subset
[3] Industry data from 13 agrochemicals
[4] Data from the US National Toxicology Program

These studies assessed the ratios of observed NOAELs from subchronic versus chronic oral tests using historical data for a sample of various compounds. It is very likely that the databases used in these studies overlap each other significantly. In two studies the ratios have not been matched for the species concerned (Rulis and Hattan, 1985; Kramer et al., 1995). The NOAELs in these studies have not been normalised for any differences in basal metabolic rate between the testspecies used for the subchronic test and that used for the chronic test. No doubt there will also be differences in the interpretation of the tests available (if interpretation is done at all in the case secondary sources have been used).

It should be noted that subchronic toxicological studies usually have smaller sample sizes compared to chronic studies (typically twice as small). Therefore, it may be expected beforehand that NOAELs from subchronic studies will tend to be larger than NOAELs from chronic studies, even if the true dose-response relationships in both studies were identical. Thus, the geometric mean ratios for the NOAELs assessed in the mentioned studies most likely overestimate the median of the distribution of the $EF_{subchronic}$.

Which distribution can now be thought to approach reality best? The main dichotomy in the meta-studies performed is in the way the available database has been analysed: several authors have computed ratios regardless the species of both (the lowest) NOAELs, whereas others have done so for (the lowest) NOAELs of the same species only. In the latter case lower means may be calculated. However, as a result of the reduced number of ratios available, the estimate of the variation may be poor. In the daily practice of risk assessment an HLV can be derived from a subchronic test applying an extra assessment factor. In such cases the question is what, given the lowest subchronic NOAEL, the value of the chronic NOAEL would have been, if a chronic test had been performed that would be acceptable to derive an HLV. Ideally, this chronic test should have been carried out using the most appropriate animal model. It follows that the most relevant NOAEL-ratios are those based on the same exposure period and the same species (in order to exclude interspecies variation) and the most relevant distributions of NOAEL-ratios are those that include a sufficient number of matched pairs of NOAELs of various species. Unfortunately, the available distributions (Tables 8 and 9) are not "the most relevant" since these are based on rather variable exposure periods for the semi-chronic NOAELs, include interspecies variation (no matching for species), for which no correction was made, and in several cases use rather old data. Differences in endpoints were not considered. The distributions obtained from NOAELs of various species therefore will probably be overconservative, whereas the distributions obtained from NOAELs of one species and more strict criteria with regard to exposure period and overall study design will probably be too narrow.

In summary: it does not seem appropriate to rely on one particular study. The geometric means of the oral semi-chronic to chronic ratios were similar in all these studies, i.e. approximately 2, whereas the GSD ranges from 1.3 to 5.6 as a result of use of different species, variable size of the database, and criteria for data selection. Based on all data together a GSD of 4 is considered a reasonable approximation of the real standard deviation (Fig.2). This is lower than the standard deviation derived from the large database of Kramer et al. (1995). However, in the latter database, interspecies differences might have played a role. Based on this distribution, default values for the 90-, 95- and 99-percentiles can be calculated to be 12, 20 and 50, respectively.

The geometric means of the oral subacute to chronic ratios were significantly higher than of the semi-chronic to chronic ratios and included more variation. Based on the data available it seems reasonable to approximate their real distribution with a geometric mean of about 4 and a geometric standard deviation of 4. Based on this distribution, default values for the 90-, 95- and 99-percentiles can be calculated to be 24, 39 and 101, respectively.

Whether the distributions also apply to inhalatory and dermal subchronic-chronic ratios is questionable. It might be possible that the influence of exposure period on the toxicological effect depends on the route of exposure. Preliminary results from analysis of inhalation data indicate that this is not the case, i.e. different means and different standard deviations were found. Further analysis of results from dermal and inhalatory studies must give more insight in this uncertainty.

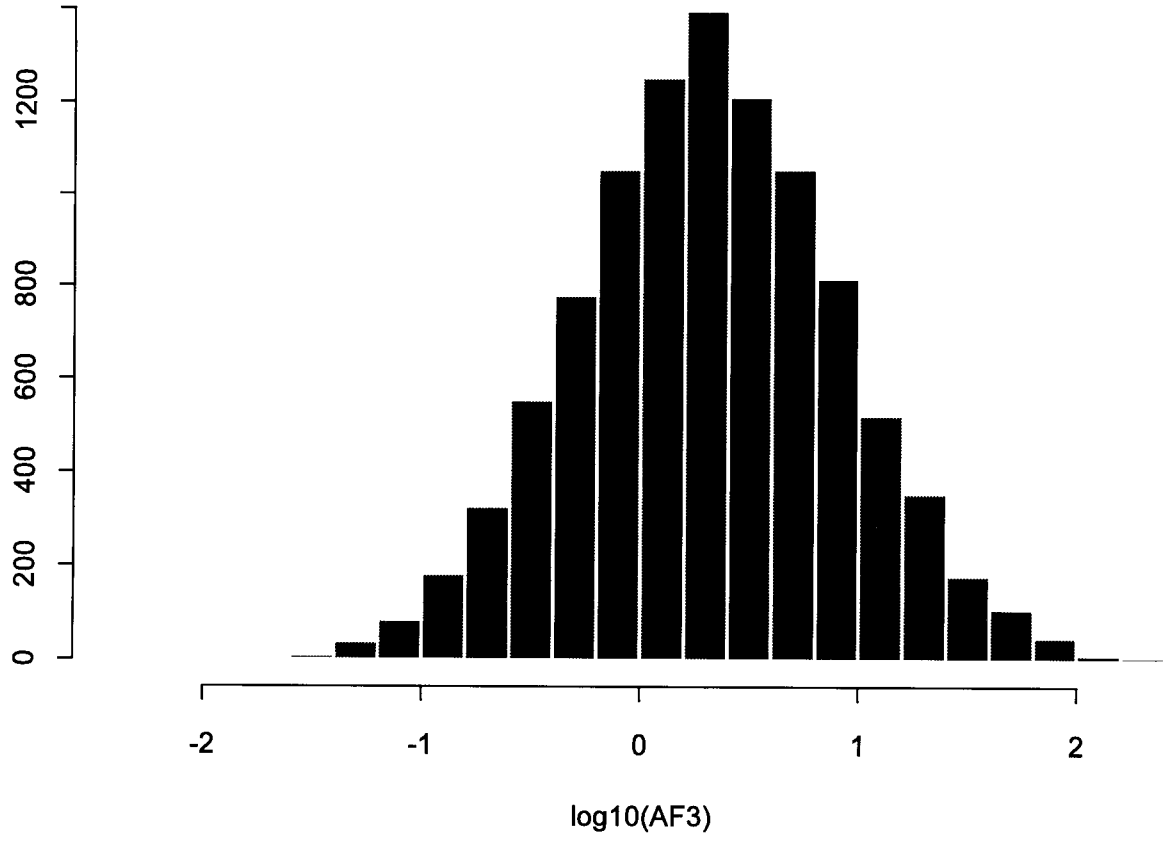

Fig. 2: Distribution of semi-chronic to chronic assessment factor. In this distribution the often applied default factor of 10 coincides with the 88th percentile

3.2.5 Dose-response curve

The dose-response curve is not used to derive a NOAEL. Apart from ensuring that the number and spacing of data points is adequate to provide a reasonable estimate of the NOAEL, all other data points are ignored. In theory, the steeper the slope of the curve, the smaller the assessment factor can be and vice versa. The threshold dose can be under- or overestimated, respectively. There is no scientific basis for any value of a default factor to account for uncertainty in the NOAEL, nor any distribution. Alternative methods such as the Benchmark-dose concept do take the shape of the dose-response curve into account. The extrapolation from the LOAEL to the NOAEL may be regarded as part of the dose-response analysis.

LOAEL to NOAEL

The use of historical LOAEL/NOAEL-ratios to estimate a NOAEL from a LOAEL (Dourson and Stara, 1983; Kramer et al., 1995) is questionable. Usually, doses in toxicological tests are spaced in fixed intervals and the observed distribution of LOAEL/NOAEL ratios therefore primarily reflects the historical frequency of use of various dose spacing (Baird et al., 1996). There is no guarantee whatsoever that extrapolation of a LOAEL with any factor will yield an estimate of the NOAEL. Therefore this factor can only be assigned using expert judgement in which the shape of the dose-response curve and the magnitude of the effect at the LOAEL is taken into account.

3.2.6 Route-to-route extrapolation

In the case relevant data are lacking on exposure routes of interest, route-to-route extrapolation is used in the risk assessment. The currently applied route-to-route extrapolation methodology is an easy, straightforward way to determine a dermal or inhalation NAEL based on an oral NOAEL. To account for differences between routes of exposure, data on absorption or acute toxicity are often used. However, these methodologies are not validated.

A study was performed that was aimed at an evaluation of route-to-route extrapolation on the basis of (estimates of) absorption or acute toxicity data (Wilschut et al., 1998). By using experimental repeated-dose toxicity data, it was tried to establish a factor to account for other factors than bioavailability which are generally not taken into account in route-to-route extrapolation.

Data were primarily gathered on dermal and respiratory repeated-dose toxicity. An extrapolation factor, defined as the factor that is applied in route-to-route extrapolation to account for differences in the expression of systemic toxicity between exposure routes, was determined for each substance by using data on absorption and acute toxicity data. As experimental data on absorption are often not available, default values for absorption were also used to determine an extrapolation factor.

Despite a rather large overall database, it was remarkable that relatively few data could be used for the evaluation. Therefore, conversions were performed to include data that initially were considered less suitable for data analysis: interspecies extrapolation based on caloric demands was introduced, and a factor 3 was applied in case a LOAEL instead of NOAEL was available. The choice of NOAELs for different exposure routes known for a substance, suitable for analysis was primarily based on the same effect. However, this criteria could not be maintained.

It appeared that for oral to respiratory route-to-route extrapolation (n=28), the predicted NAEL was often higher than the observed NOAEL. So, the substance was considered less toxic after extrapolation when compared with experimental observations. Based on the 95th-

percentile of the log-normal distribution of the ratios between the predicted NAEL and the observed NOAEL, uncertainty factors ranging from 75 to 201 for the different extrapolation methodologies were found.

For oral to dermal route-to-route extrapolation (n=25), the predicted dermal NAEL was often lower than the observed NOAEL, e.g., the substance was often considered to be more toxic after extrapolation when compared with experimental observations. Uncertainty factors ranging from 2.7 to 35 were found for the different extrapolation methodologies.

Given the implications of the use of these uncertainty factors, (inter)national discussion on the results of the study of Wilschut et al. (1998) would be of value for optimal tuning in view of future application. As a result of such discussions more data may become available. It should be noted that the reliability of the data is questionable as the influence of the several assumptions made in order to derive comparable data on the ratio of the predicted NAEL and the NOAEL is unknown.

For both extrapolations, the results were hardly influenced by the assumptions made on absorption, indicating that other factors may be important in route-to-route extrapolation, and/or the reliability of the estimates of absorption used in this study was poor.

Given these findings, it was concluded by Wilschut et al. (1998) that the development of scientifically based principles and procedures for route-to-route extrapolation appears to be a difficult task without the availability of adequate experimental data. In this study, scientific justification for the application of route-to-route extrapolation was not derived. As only a limited number of data on toxicity after repeated dermal and inhalatory exposure were found after an extensive search, it is doubtfull whether such experimental data indeed do exist. Insight in the reliability of route-to-route extrapolation methodologies may then only be obtained if more suitable experimental data become available.

In summary:
Several options can be considered to deal with the issue of route-to-route extrapolation:

1. Assessment factors are used to account for uncertainties in the route-to-route extrapolation;
2. Repeated-dose toxicity studies with exposure routes relevant for the risk characterisation are performed until validated and reliable route-to-route extrapolation methodologies will be available. The role of PBPK modelling should also be further investigated here.

The choice between these options is a regulatory one and depends on the desired degree of reliability of the risk assessment. At this moment, the application of route-to-route extrapolation heavily depends on expert judgement.

3.2.7 Type of critical effect

The type of critical effect should be taken into account. Assessment factors may be applied by

expert judgement depending on each individual case. By default it can be assumed that no extra correction is necessary.

3.2.8 Confidence in the database

The size, quality, completeness, and consistency of the database should be considered. The schemes available indicate that the assessment factor should be higher than unity in the case one is less confident about the database and may run up to 100. On the other hand, in case of a very reliable database, a factor less than one may be applied (US-EPA). This assessment factor can only be assigned on the basis of expert judgement, preferably made transparent through the application of a set of criteria. It may be argued that a database necessitating very high assessment factors are probably inadequate for the risk assessment altogether. ECETOC proposed to distinguish between a high, medium and low degree of confidence. The default (high confidence) is 1. A medium degree of confidence would warrant a higher assessment factor, running up to 10 for a low confidence database.

Criteria should be developed to make expert decisions transparent before implementing this approach.

3.3 Combining assessment factors

In the standard procedure for deriving acceptable limit values, various assessment factors are multiplied to obtain an overall assessment factor. However, multiplication of assessment factors implies a piling up of worst case assumptions: the probability of simultaneous occurrence of worst case situations for the same chemical will be smaller than that of a single worst case situation to occur. Therefore, the more extrapolation steps are taken into account, the higher the level of conservatism.

The piling-up of worst-case assumptions can be avoided by using probability distributions (Baird et al., 1996; Slob and Pieters, 1997a). In this method each assessment factor is considered uncertain and characterised as a random variable with a distribution. Propagation of the uncertainty can be evaluated using Monte Carlo simulation yielding a distribution of the overall assessment factor. This method requires characterisation of the distribution of each assessment factor, as was attempted in this chapter, and of possible correlations between them. As a first approach it can be assumed that all factors are independent.

The database derived distributions of assessment factors can be used to establish default point values by taking for example the P_{95} of these distributions. These P_{95}-values can be multiplied to arrive at an overall factor. This is shown in the second example of Chapter 5 (section 5.3.2). As argued above, conservatism can be avoided by probabilistic multiplication of the derived distributions. An example of this approach is shown in section 5.3.3.

4. NEW CONCEPTS IN DERIVING HUMAN LIMIT VALUES

4.1 Introduction

In the standard procedure for deriving Human Limit Values (HLVs), such as ADI, TDI, RfD, or HBORV from animal study data, the NOAEL is divided by a number of assessment factors according to equation 4.1:

$$ADI, TDI, RfD = \frac{NOAEL}{AF_1 . AF_2 . AF_3 .....} \qquad (4.1)$$

The assessment factors are assumed to be independent from each other (see also 3.1). Because of this multiplication the standard method for deriving HLVs is generally considered to be conservative. Indeed, when each individual assessment factor by itself is regarded to reflect a worst case situation, their product, i.e. the overall assessment factor, will tend to be overly conservative. However, the degree of conservatism in the limit value in any particular assessment is unknown, which hampers risk managers to appraise possible health risks against other (e.g. economic) interests.

On the other hand, the uncertainty in the numerator, the No-Observed-Adverse-Effect Level (NOAEL) as an estimate of the "true" No-Adverse-Effect Level ($NAEL_{true}$) in the animal is completely ignored. The NOAEL is defined as the highest dose level at which no statistically significant effects occur for all endpoints that are considered toxicologically relevant. This NOAEL is not the same as the $NAEL_{true}$. Suppose there is a (true, but unknown) threshold dose below which the substance does not evoke any adverse effects. Depending on the study design, the NOAEL resulting from a statistical analysis of the data can be lower or higher than this dose. The potential deviation of the NOAEL from the $NAEL_{true}$ cannot be quantified. The latter uncertainty may be substantial and ignoring it may introduce an anti-conservative or an additional conservative element in the derivation of acceptable exposure limits.

This chapter will further examine the uncertainties present in both the numerator and the denominator of equation 4.1. To this end first a conceptual framework will be presented as worked out by by Slob and Pieters (1997a). In subsequent sections this concept will be operationalised. The practical implications of this operationalisation is shown in an example in section 5.3.3.

4.2 The concept

Assume there is a true No-Adverse -Effect Level in the sensitive human, or $NAEL_{sens.human}$. Since this $NAEL_{sens.human}$ usually is derived from animal data, extrapolation factors have to be applied. To that end the true factor (EF_{true}) is defined. as an alternative for the assessment factor (AF, see equation 4.1). The $EF_{true,\ interspec}$ is defined as the ratio between the 'true', but unknown, NAEL in the animal ($NAEL_{true,animal}$) and the 'true', but unknown, NAEL of the average human ($NAEL_{true,human}$). Any particular compound has its own $EF_{true,interspec}$.

$$EF_{true,interspec} \equiv \frac{NAEL_{true,animal}}{NAEL_{true,human}} \qquad (4.2)$$

Similarly, the intraspecies EF_{true} is defined as

$$EF_{true,intraspec} \equiv \frac{NAEL_{true,human}}{NAEL_{true,sens.human}} \qquad (4.3)$$

Clearly, for a particular compound we have

$$NAEL_{true,sens.human} = \frac{NAEL_{true,animal}}{EF_{true,interspec}\, EF_{true,intraspec}} \qquad (4.4)$$

Although equation 4.4 has the same appearance as the standard equation (4.1), it fundamentally differs in interpretation: all entities in (4.4) refer to true but unknown values.

For the operationalisaton of this concept, the question therefore is how to estimate the $NAEL_{animal}$ and the EFs and the uncertainty distribution associated to each of them. The next section will deal with the best approximation of the distribution of the $NAEL_{animal}$. With regard to the EFs is can be argued that, alhough the value of the EFs are unknown for specific compounds, the extrapolation factors for the universe of all compounds must have a specific distribution. One might be able to estimate that distribution from historical data (e.g. from drugs). Ideally this should be done on the basis of ratios of the best approximations of the $NAEL_{true}$. More crude estimates of the distributions of EFs can be obtained on the basis of NOAELs as was done in the previous chapter. It was argued that the database derived distributions thus obtained are wider than would be obtained on the basis of the $NAEL_{true}$.

4.3 Estimation of the No-Adverse-Effect Level in the animal

4.3.1 The NOAEL and the true NAEL

The numerator in equation 4.1, the NOAEL, is defined as the highest dose level at which no statistically significant effects occur (for all endpoints that are considered toxicologically relevant). As pointed out in the introduction to this chapter, it is important to keep in mind that the NOAEL is not the same as the $NAEL_{true}$ and that, although the NOAEL could be considered an estimate of the true threshold dose, the quality (precision) of the estimate cannot be assessed.

Other objections against the use of the NOAEL have been discussed extensively elsewhere (e.g., Crump, 1984; McColl, 1989; Beck et al., 1993). In general there is a call for consideration of the dose-response relationship as a whole. One of the alternatives proposed has been the benchmark approach (Crump, 1984; US-EPA, 1995).

4.3.2 The benchmark dose concept

In the benchmark approach a regression function fitted on response data is used to estimate the dose at which adverse effects start to arise. Using regression models for describing the dose-effect relation has two advantages. Firstly, a 'modifying factor' to account for the steepness of a dose-effect curve is redundant and secondly, there is no need to extrapolate a LOAEL to a NOAEL. The latter may be considered as a major advantage since there is no scientific justification for the use of an assessment factor for LOAEL-NOAEL extrapolation.

In the benchmark concept one needs to postulate a critical effect size (CES) below which there is no reason for concern. The CES for a critical endpoint is defined as:

CES ≡ value of effect-size below which there is no reason for concern

and the associated Critical Effect Dose (CED) as:

CED ≡ dose at which the average animal shows the (postulated) critical-effect-size defined for a particular endpoint.

A drawback of using dose-effect curves for the evaluation of toxicity is that current toxicological and biological knowledge does not provide sufficient basis to unequivocally establish the breaking point between non-adverse and adverse effect size for most endpoints. Since a single "universal" CES does not seem a realistic option, a value must be chosen for each separate endpoint. A wide-spread implementation and acceptance of the value of CES for each of the (most relevant) endpoints would require international consensus on this issue.

The CED is referred to here as a true, unknown value, which we can only estimate with a certain degree of precision, if data for the endpoint of concern are available. The true No-Adverse-Effect-Level ($NAEL_{true}$) may then be defined as the lowest CED of all endpoints:

$NAEL_{true}$ ≡ minimum of all CEDs.

The definition of the $NAEL_{true}$ refers to a true, unknown value, which can only be estimated with a certain degree of precision, if data are available for the endpoint involved. The definition of the NAEL refers not only to an unknown, but also to a rather theoretical value, since it is unknown to what endpoint it is associated. In practice, one can never be sure whether information on all relevant endpoints for the compound studied is present. Furthermore, the lowest CED in two situations (e.g., animal versus human) may not refer to the same endpoints. For example, rats may be most sensitive to endpoint A, but humans to enpoint B. EFs, as discussed in the previous section, therefore should be applied and can best be approximated by the ratio of CEDs per endpoint. The variation of EFs between all endpoints and all substances should preferably be expressed by the distribution of these CED-ratios rather than by the distribution of NOAEL-ratios. Unfortunately, quantitative knowledge on these distributions of CED-ratios, or benchmark doses, is scarce.

A drawback on the use of dose-response modelling from a practical point of view, is that most toxicity data are not suitable for curve-fitting procedures (Crump, 1984; Beck et al., 1993; Woutersen et al., 1997). A typical study design as agreed upon in for example OECD guidelines considers three dose groups and a control. Ideally, more dose groups should be used with each dose group comprising less animals. See Slob and Pieters (1997b) for elaboration on this issue.

4.3.3 The probabilistic approach towards the CED

When for a particular endpoint data are available that allow for fitting a regression function, the CED may be estimated. Depending on the quality of the data, this estimate has a certain degree of imprecision. To take this into account, Crump (1984) proposed to calculate the lower 95%-confidence limit of the estimated CED. Slob and Pieters (1997a) proposed to find the complete uncertainty distribution of this estimate by bootstrapping: once a regression model has been fitted, Monte Carlo sampling is used to generate a large number of new data sets from this regression model, each time with the same number of data points per dose group as observed animals in the real experiment. For each generated data set the CED is re-estimated. Taking all these CEDs together results in the required distribution.

4.4 The probabilistic approach towards the HLV

Since for each EF a certain distribution over all endpoints and substances is assumed it is possible to extrapolate any CED from one situation to the other. Thus, instead of choosing a single (most sensitive) endpoint from the animal data, each CED-distribution that is associated to a relevant endpoint is extrapolated to the distribution of the associated CED in the sensitive human ($CED_{sens.human}$) by probabilistic combination with the distributions of each EF. This results in a series of distributions for $CED_{sens.human}$, each related to another endpoint. Then this complete set of distributions can be considered as a basis for deriving a HLV, for example by choosing the lowest of each distribution's first percentile. It is noted that the assumption of complete independence of the various distributions of EFs will also be applied here. It has been argued that this worst-case assumption may not been valid (Calabrese et al., 1993). In case correlations can be demonstrated and quantified the method can allow for these by introducing correlation coefficients.

Fig. 3. illustrates the proposed approach.

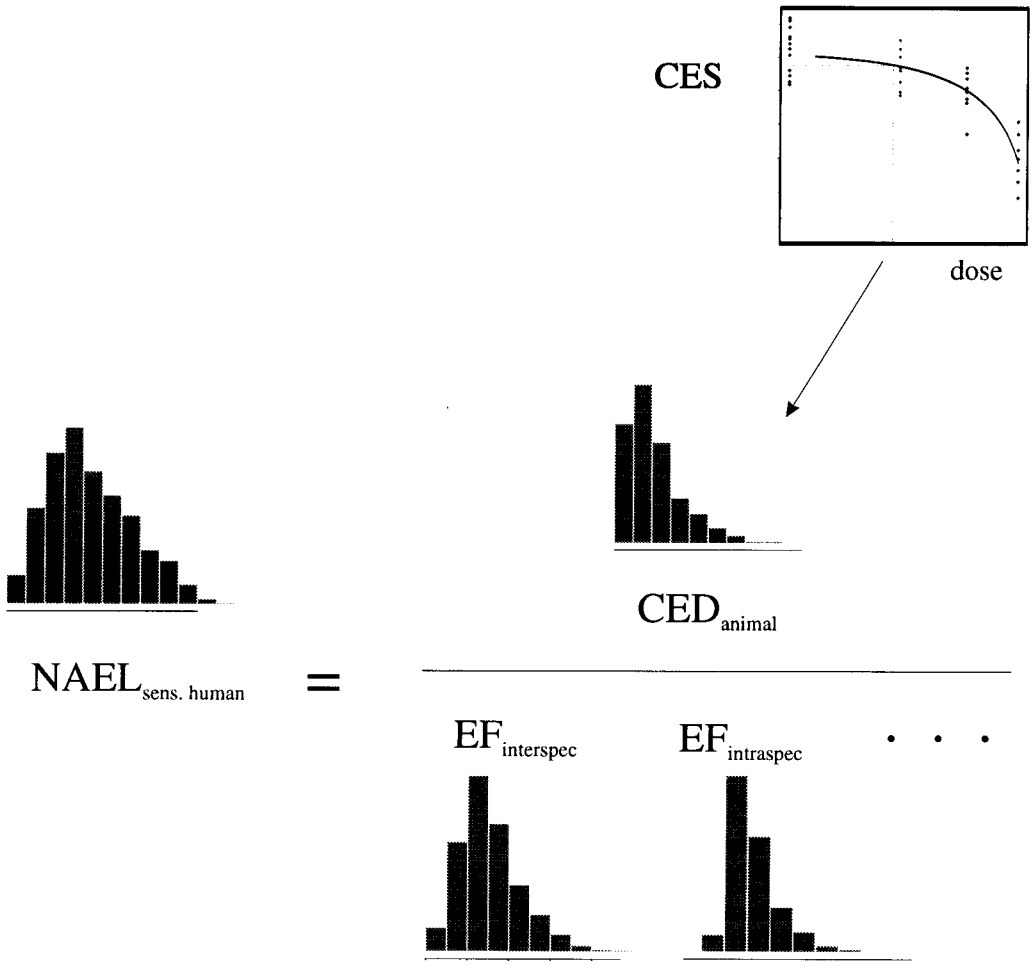

Fig. 3: The probabilistic determination of a Human Limit Value

4.5 Conclusions

The approach as discussed above differs from the benchmark approach (Crump, 1984) in various ways. Crump introduced the "Benchmark Dose Level (BMDL)", defined as the lower 95%-confidence limit of the CED, as a starting point for extrapolation to the (sensitive) human. By dividing the BMDL by assessment factors for interspecies and intraspecies variation (default values of ten), a HLV can be derived. Instead of this, it is first of all proposed to use the entire distribution of the CED instead of the lower 95%-confidence limit of the Critical Effect Dose (CED). Secondly, it is proposed to combine this CED distribution with distributions of extrapolation factors in a probabilistic way. The result of the probabilistic combination of distributions is in the form of an assessment distribution, so that the degree of conservatism is quantifiable in any particular assessment. As a matter of fact, this approach allows for deriving a HLV as a function of an *a priori* chosen degree of conservatism. In addition, the approach allows for estimating the lower and upper bounds for possible health effects in the sensitive population at a given exposure level.

5. AN EXAMPLE

5.1 Introduction

The different approaches discussed in previous chapters will now be applied to an example substance EXA. EXA has an oral NOAEL of 400 $mg.kg_{bw}^{-1}.d^{-1}$. This NOAEL was derived in a 3-months test (semi-chronic) in rats. This example will only include extrapolation from experimental animals to average humans (AF_1), from average humans to sensitive humans (AF_2) and from a semi-chronic toxicity test to a chronic test (AF_3). In all approaches the target is to protect all human beings and therefore the 95th percentile (P_{95}) of distributions of assessment factors is selected for the derivation of the Human Limit Value (HLV) for man.

Disclaimer:
These example calculations are presented with the purpose of showing the procedures in each of four approaches. It should be recognised that no general conclusions can be attached to the quantitative outcome for this particular substance.

The different approaches applied are:
1. Multiplication of assessment factors (traditional approach using default factors of 10);
2. Multiplication of the database-derived default assessment factors, proposed in section 3.2;
3. Probabilistic multiplication of database-derived assessment factors, proposed in section 3.2;
4. Determination of the distribution of a Critical Effect Dose and, combining this with distributions of database-derived assessment factors, the probabilistic estimation of the HLV as discussed in chapter 4.

In all examples, the following equation is used to derive the HLV:

$$HLV = \frac{NOAEL}{AF_{tot}} \tag{5.1}$$

The overall assessment factor AF_{tot} is calculated from the following formula:

$$AF_{tot} = AF_1 . AF_2 ..... AF_n \tag{5.2}$$

5.2 Input data

In an OECD-test, groups of 20 rats of both sexes received through their diet doses of 0, 400, 1200 or 4800 $mg.kg_{bw}^{-1}.d^{-1}$ for 90 days. Increased mortality was observed in female rats at the highest dose. The main other effects observed were on body weight, lactate dehydrogenase levels and histopathology of the urinary bladder (mucosal hyperplasia). The mean effects data are shown in Table 10, but individual data for both sexes combined were used for modelling.

For the purpose of this example all methods discussed will concentrate on the effect on lactate dehydrogenase levels.

Table 10: Results of the semi-chronic test of EXA[1]

Dose ($mg.kg_{bw}^{-1}.d^{-1}$)	Survival		mean body weight (g)		mean LDH level[2] (bb/ml)		incidence of UBH	
	m	f	m	f	m	f	m	f
0	19/20	19/20	480	264	1893	1427	0/20	0/20
400	20/20	20/20	480	266	2075	1584	0/20	0/20
1200	20/20	18/20	453*	274	2442*	1971*	2/20[3]	3/20[3]
4800	19/20	15/20*	346*	252*	4637*	3866*	15/20[4]	13/20[4]

[1] LDH = lactate dehydrogenase, UBH = urinary bladder mucosal hyperplasia, m = males, f = females

[2] LDH-levels were determined in 10 rats/sex/dose.

[3] very slight

[4] very slight (3m, 9f), slight (3m, 3f), moderate (6m, 1f), marked (3m, 0f)

* = statistically significant (t-test)

5.3 The derivation of the HLV

5.3.1. Multiplication of assessment factors (traditional approach)

The default assessment factors used traditionally take on values of 10:

10 for AF_1 = interspecies factor
10 for AF_2 = intraspecies factor
10 for AF_3 = factor for duration of exposure

The HLV will be 400/(10.10.10) = 0.4 $mg.kg_{bw}^{-1}.d^{-1}$
The overall assessment factor is 1000.

Comment: In case sufficient data on a substance are available, default values may be replaced by more toxicity profile derived (data derived) assessment factors. An example of such an approach is shown in Annex I.

5.3.2. Multiplication of database-derived assessment factors

In this approach, the default assessment factors take on the values as derived above on the basis of an analysis of the historical data available. Note again, that the values of the individual assessment factors are considered to be the 95th percentile of the total distribution.

76 (4 . 19) for AF_1 = interspecies factor based on the extrapolation from rats to humans
10 for AF_2 = intraspecies factor
20 for AF_3 = factor for duration of exposure
The HLV will be 400/(76.10. 20) = 0.026 $mg.kg_{bw}^{-1}.d^{-1}$
The overall assessment factor is 15200.

Comment: The multiplication of P_{95}-values of broad, database-derived assessment factors leads to an overall factor which is far beyond the P_{95} of the overall assessment factor and therefore is far too conservative.

5.3.3. Probabilistic multiplication of assessment factors

The same formulae apply as stated above, but now each assessment factor either has a discrete value in case of known values (e.g. in case of the factor indicating the quality of the database) or a distribution of values caused by uncertainty due to empirical inaccuracy or lack of empirical data. The propagation of the uncertainty in each assessment factor can be performed by Monte Carlo simulation. This analysis is performed by sampling randomly from the distributions specified for each assessment factor and combine these values to a distribution for the overall assessment factor. It is assumed here that the factors are not correlated. The data relevant to the Monte Carlo simulation for EXA are shown in Table 11. Propagation of the uncertainty was investigated using 1000 runs.

The histograms of the distributions in Table 11 are shown in figures 1 (interspecies factor) and 2 (semi-chronic to chronic factor). The distribution of the overall assessment factor $AF_1.AF_2.AF_3$ is shown in Figure 4. This assessment factor AF_{tot} has a median of 80 and a P_{95} of 3300.

Based on the P_{95} the HLV will be 400/3300 = 0.12 $mg.kg_{bw}^{-1}.d^{-1}$

Table 11: Input for the Monte Carlo simulation for EXA

Assessment factor	Distribution	Geometric mean	Geometric standard deviation
AF_1: interspecies, kinetics	discrete value	4	-
interspecies, residual	lognormal	1	6
AF_2: intraspecies	discrete value	10	-
AF_3: duration of exposure	lognormal	2	4
semi-chronic to chronic			

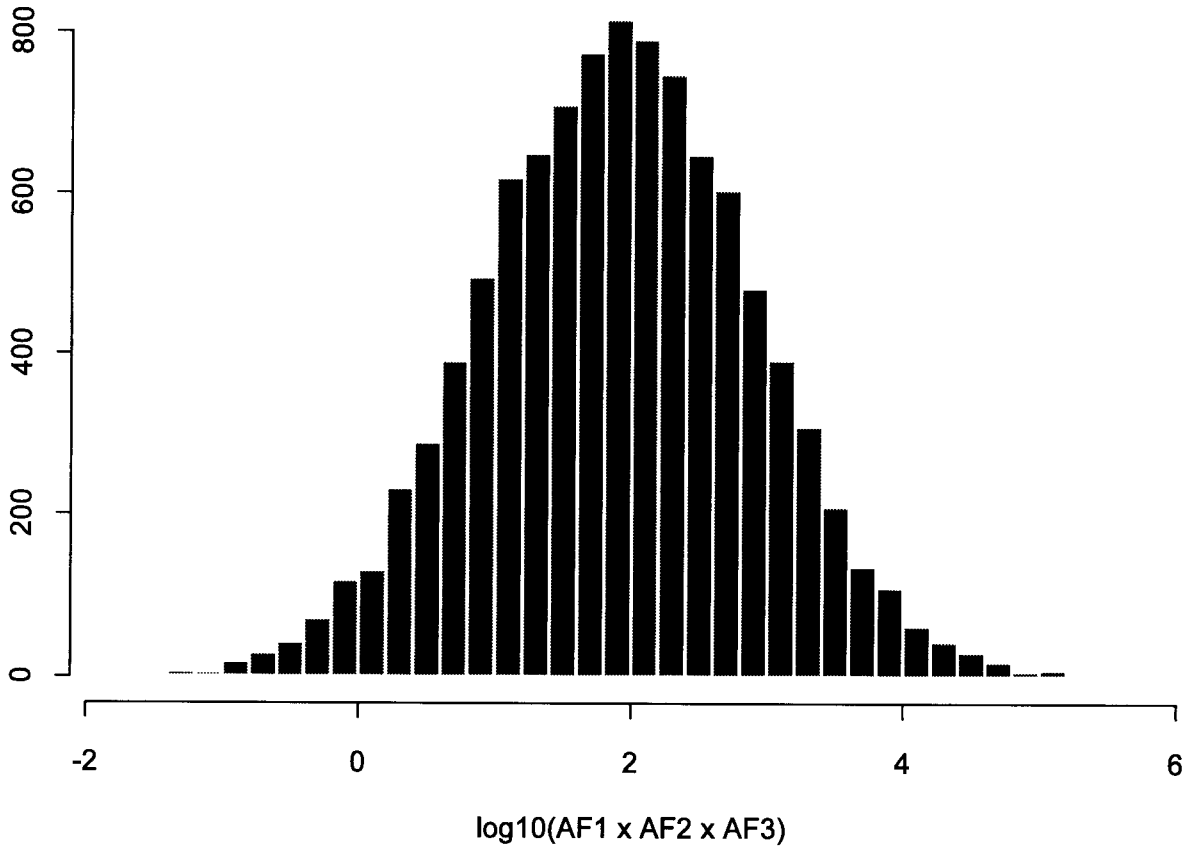

Fig. 4: Distribution of AF_{tot}: the median is 80 and the GSD is 9.6: it follows that P_{87} is 1000, P_{90} is 1452, P_{95} is 3303 and P_{99} is 15424.

Comment: This method characterises partly the probabilistic uncertainty in the HLV as derived from experimental data. The uncertainty in the experimental NOAEL itself is ignored.

5.3.4 Probabilistic estimation of the HLV

In this example the probabilistic approach as discussed in chapter 4 will be applied to the example substance including the distribution in the NAEL, estimated from the dose-response curve.

The dose-response curves fitted to the male and female LDH data are shown in Figure 5. No difference in sensitivity between males and female rats is apparent and for both sexes a CED_{animal} of 967 $mg.kg_{bw}^{-1}.d^{-1}$ is derived at a CES of 20%. The CES of 20% is chosen by expert judgement and therefore is a matter of debate. The associated uncertainty distribution of the CED at a CES of 20%, obtained by Monte Carlo analysis, is shown in Figure 6. This distribution is combined with the (distributions of) the assessment factors AF_1, AF_2 and AF_3, as shown in Table 11. This results in the distribution of the CED for the sensitive human population, $CED_{sens.human}$, which is shown in Figure 7.

Taking the HLV to be the P_5 of the $CED_{sens.human}$, the HLV is 0.25 $mg.kg_{bw}^{-1}.d^{-1}$.
The overall assessment factor relative to the NOAEL of 400 $mg.kg_{bw}^{-1}.d^{-1}$ is 1600.

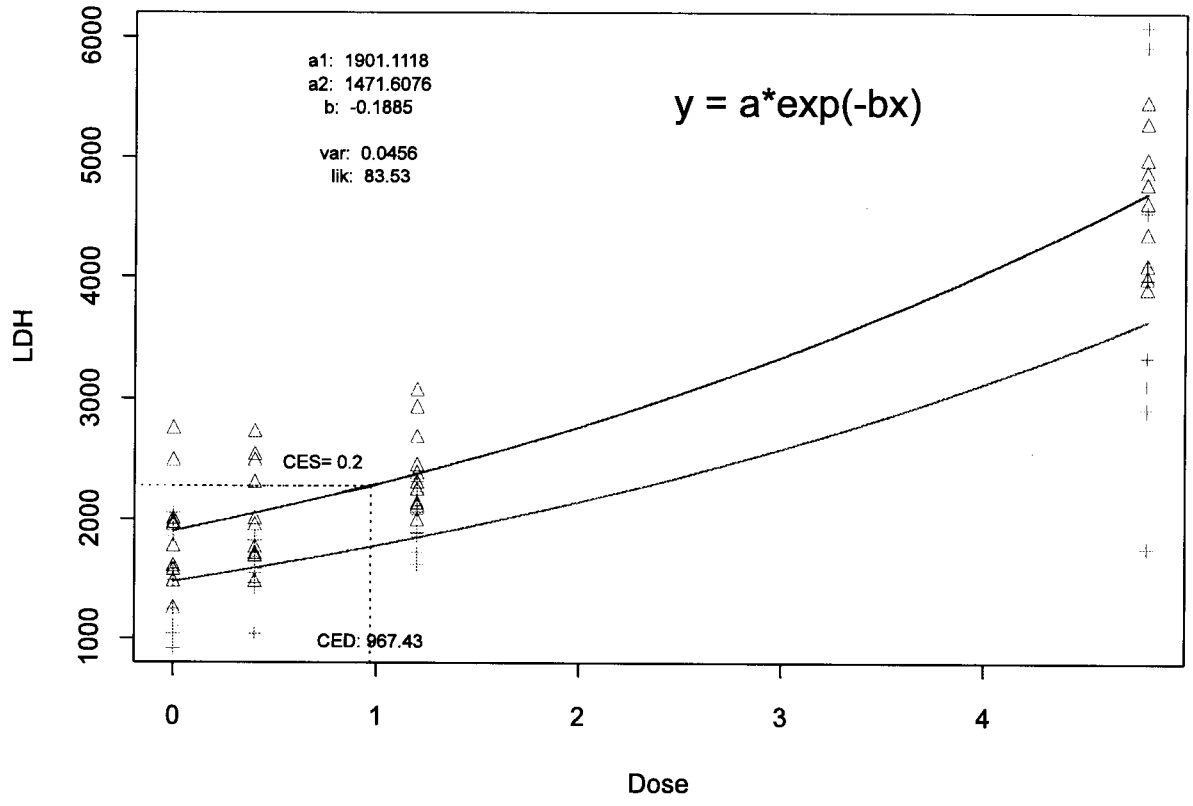

Fig. 5: Regression curves for male (triangles) and female (plusses) rats exposed for 3 months to EXA

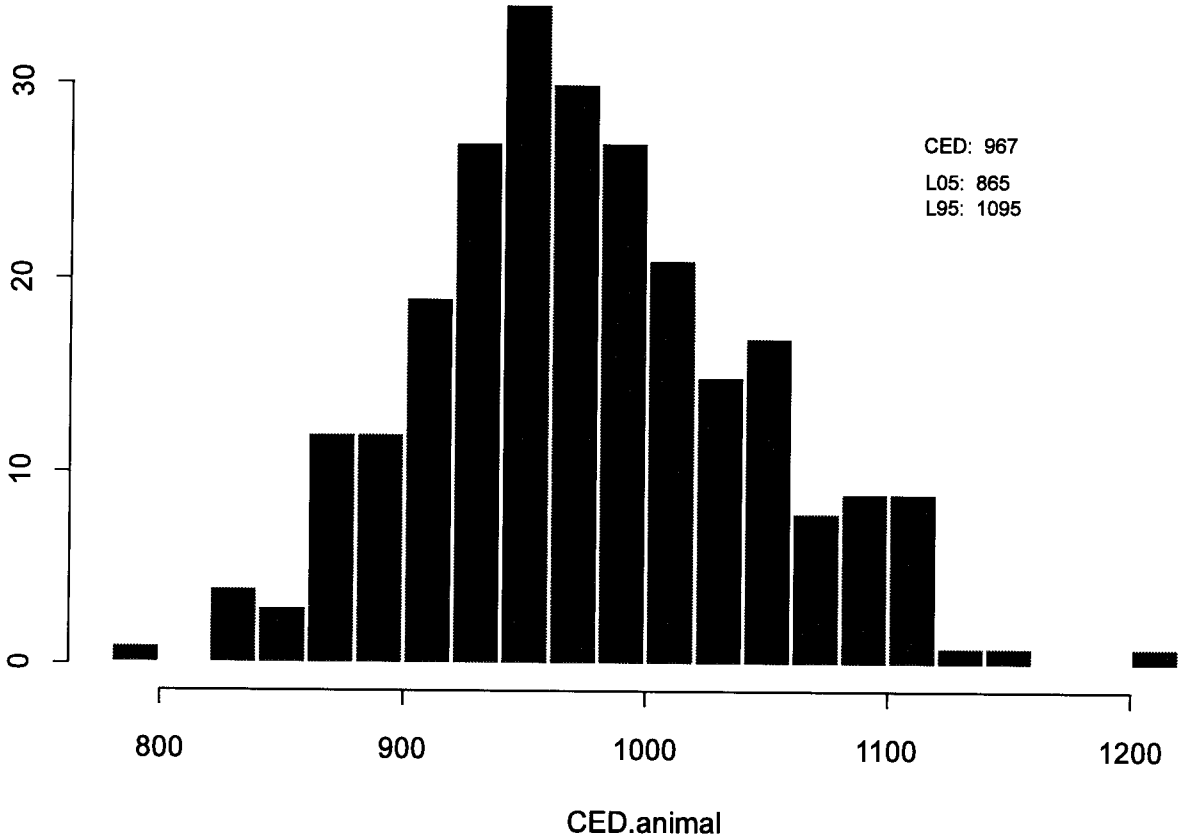

Fig. 6: Uncertainty distribution for the CED_{animal} for EXA at a CES of 20%. L05 is 5% lower confidence limit, L95 is 95% upper confidence limit.

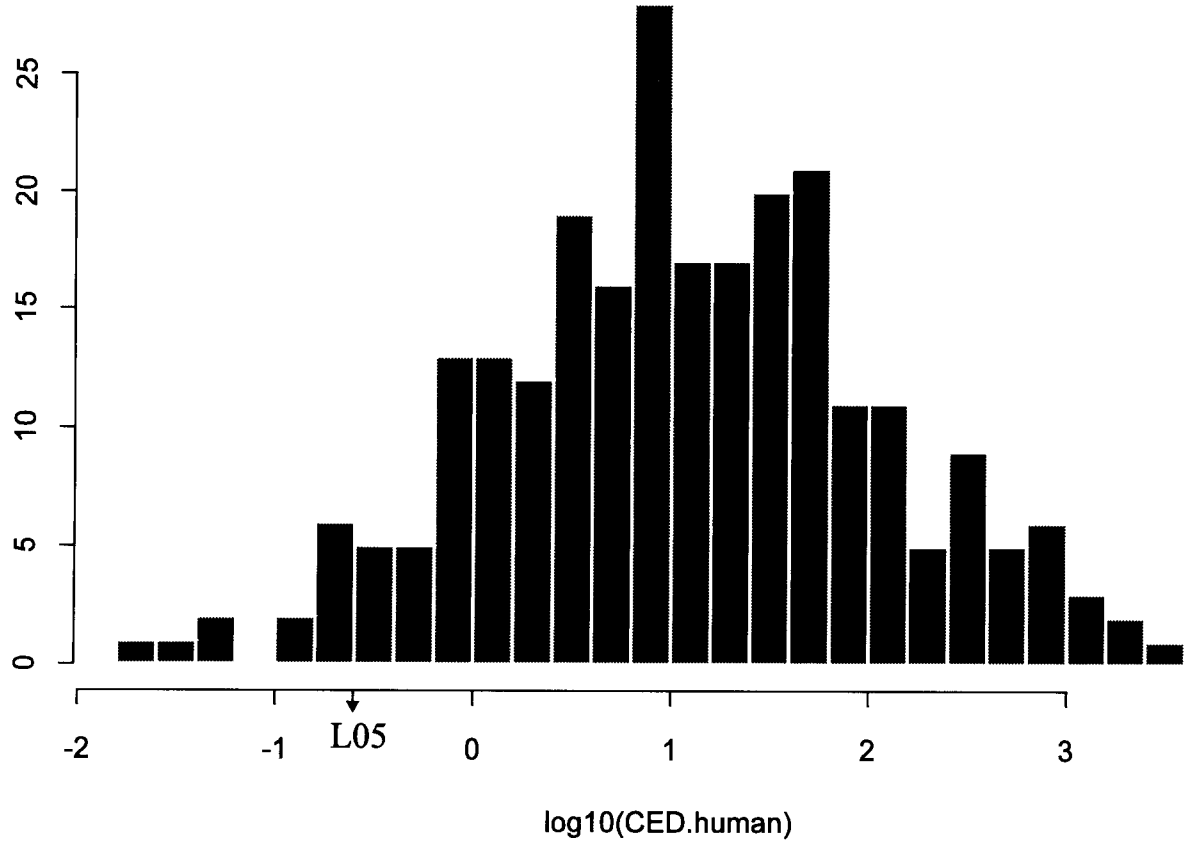

Fig. 7: Distribution of the dose level of EXA that would evoke a 20% adverse effect in the sensitive human subpopulation. The 5th percentile of this distribution (L05 on x-axis) is 0.25 $mg.kg_{bw}^{-1}.d^{-1}$.

6. CONCLUSIONS

6.1 MOS versus NOAEL

The human risk characterisation according to the EU Technical Guidance Documents for new and existing substances is based on the NOAEL/exposure ratio, i.e. the Margin of Safety (MOS). It can be concluded that more explicit guidance for human risk characterisation is needed to reduce variability and to achieve higher transparency and consistency between risk assessments of individual Member States. A formal, harmonised set of default assessment factors for the derivation of a Human Limit Value (HLV) or for judgement of the MOS is considered essential here.

6.2 Assessment factors

In this discussion document the quantification of assessment factors is adressed. Historical data were analyzed to determine the distribution of the factors for interspecies differences and for duration of exposure and the probabilistic approach towards the combination of factors is described. Various approaches towards human risk characterisation with regard to the application of assessment factors have been summarised and evaluated. The three methods distinguished comprise:

1. Multiplication of assessment factors (traditional approach using default factors of 10);
2. Multiplication of the database-derived default assessment factors;
3. Probabilistic multiplication of distributions of database-derived assessment factors.

In all these methods correlations between factors are not taken into account.

From the analysis of the various assessment factor approaches the following can be concluded:

1. For interspecies extrapolation, allometric scaling on the basis of caloric demands (the 0.75 power of body weight) is considered preferable above scaling on body weight.
2. The traditional extrapolation approach, based on more or less arbitrary factors of 10, is simple to apply but obscures the relative contributions of scientific arguments and policy judgements. Inflexible and arbitrary assessment factors may result in limited utilisation of existing knowledge (Beck et al., 1993). The two other default approaches as well as the application of a toxicity profile derived factor (Hakkert et al., 1995) make better use of the data available and therefore are more directed to a scientific estimation of safe exposure levels for humans.
3. The worst case character of the traditional default assessment factors is doubtful in view of the data analyzed. The P_{95}-values of the proposed distributions for the interspecies factor and the semi-chronic to chronic timescale factor are considerably higher than 10 and the limited data on intraspecies variation also indicate that a default factor of 10 may not be sufficient. However, it is also noted that the real distributions may be less wide than observed on the basis of historical data.
4. The derivation of approximations of the distribution of assessment factors from historical data, i.e. on the basis of NOAEL-ratios, has limitations. NOAEL-ratios are assessed

without always knowing the quality of the underlying data. Furthermore, it should be recognised that the use of the NOAEL instead of the $NAEL_{true}$ brings along the variation (error) in the NOAELs. The NOAELs are only rough estimates of the true NAEL. It is noted for example that true risks at the NOAEL may vary from 1 to over 10% (Leisenring & Ryan, 1992; US-EPA, 1995). Therefore, the geometric standard deviations of the NOAEL-ratios assessed in the studies will overestimate the variation among the ratios of the CEDs. Unfortunately, it is impossible to quantify the measurement error of a NOAEL and to correct for this (Slob & Pieters, 1997a).

5. The application of assessment factors derived from current estimated distributions of assessment factors (examples 2 and 3) may lead to very wide distributions of the overall assessment factor unless chemical-specific data can be introduced (Annex I). This large variation can be expected, but will also arise from the conservatism in the method for the derivation of the assessment factors used (see point 2 above).
6. Probabilistic multiplication of distributions of assessment factors (example 3) is preferred above simple multiplication of percentiles to avoid extreme conservatism without indication how conservative it may be. The result of the probabilistic combination of distributions is in the form of an assessment distribution, so that the degree of conservatism is quantifiable in any particular assessment. As a matter of fact, this approach allows for deriving a HLV as a function of an *a priori* chosen degree of conservatism. In addition, the approach allows for estimating the lower and upper bounds for possible health effects in the sensitive population at a given exposure level.
7. A prerequisite for the application of methods 2 and 3 is that consensus is reached on default distributions for the assessment factors.

6.3 The benchmark dose concept

Furthermore, the framework presented in Slob and Pieters (1997a) is described and may be considered as the 'ideal' approach for deriving exposure limits and for quantification of the risk of exceeding these limits. This method is considered as a complete probabilistic approach, offering the possibilities of comparing the various uncertainties involved in typical risk assessment, including the uncertainty in the exposure estimate, the uncertainty in the toxicological starting point (benchmark dose concept), and the uncertainty in assessment factors.

The benchmark-dose concept takes into account the information on the dose-response and the uncertainties in the estimation of the "true" experimental threshold in the animal, depending on the quality of the particular study from which the data are used. The method presented also allows for estimating the lower and upper bounds for possible health effects in the sensitive population at a given exposure level. The use of assessment factors for LOAEL-NOAEL extrapolation and for the 'steepness' of a dose-effect curve are completely redundant applying the benchmark-dose. Additionally to the uncertainties in deriving distributions of assessment factors as described above, this method has the following drawbacks:

1. Consensus needs to be reached on the definition of Critical Effect Sizes for all

toxicological endpoints that may be relevant. Current toxicological and biological knowledge does not provide sufficient basis to unequivocally establish the breaking point between non-adverse and adverse effect sizes for most endpoints.
2. The method is less straightforward than the NOAEL-approach and requires some statistical experience in fitting mathematical dose-response models to data.
3. The data available in many cases will not allow modelling, since they were not generated with the intention of dose-response modelling. A study design with more, but smaller dose groups may be helpful here.

6.4 The way forward

It should be considered to replace the traditional application of default assessment factors of 10 (example 1) by more database-derived defaults and distributions of assessment factors. Probabilistic multiplication of distributions of these factors (example 3) is then preferred above simple multiplication of percentiles (factors) to avoid extreme conservatism. It can be considered to replace the current methods by this framework in future:

- *Determination of the value of assessment factors:* The value of various assessment factors as well as its combination by multiplication is a matter of current debate. Consensus needs to be reached on the distributions of assessment factors for e.g., inter- and intraspecies extrapolation and duration of exposure.
- *Probabilistic approach towards the combination of assessment factors:* If consensus can be reached on these distributions, they may be combined by a probabilistic approach.

This can be considered as a first step for the implementation of a complete probabilistic method of risk assessment, but more work is needed to fully override the traditional method:

- *Determination of the CED:* Parallel to the determination of the NOAEL, experience can be gained in defining CESs and in modelling the dose-response data to derive CEDs and to establish distributions of EFs on the basis of CED-ratios.
- *Probabilistic approach towards the determination of the HLV:* Subsequently, experience can be gained in arriving at the distribution of a particular CED and in the probabilistic derivation of the HLV.
- *Probabilistic approach towards risk characterisation and risk management:* Risk assessors and risk managers continuously needs to examine the pros and cons of each of the above stages and their implications for decision making.

In this way further research in this area will benefit from the experience gained. At the same time both risk managers and the public will have to adapt themselves to the new type of judgements they face: rather than relying on one point estimate for the HLV, a decision has to be made on the acceptable degree of confidence in the estimation of the HLV as well as in the estimation of exposure. An important aim should be to avoid overconservative estimates which may lead to high costs for risk reduction measures.

6.5 Recommendations for further research

The approach presented still contains a lot of uncertainties. It is recommended to direct future research on the elucidation of the major aspects of these uncertainties along the following lines:

1. Interspecies extrapolation: additional data analysis is needed to investigate the variability in the available allometric scaling methods to account for pharmacokinetic differences. This needs a comprehensive comparison of available NOAELs or preferably CEDs of different species for different endpoints. Research is also needed to gain insight into pharmacodynamic extrapolation.
2. Intraspecies extrapolation: research into human variability in sensitivity to chemical substances so far is very limited both with regard to pharmacokinetic and to pharmacodynamic variation. Data analysis should be continued towards the derivation of the distribution of these factors.
3. The assessment factor adjusting for differences in time scale between the experimental result and the exposure scenario considered does not differentiate among substances, e.g. with regard to mechanism of toxicity or effect. The evaluation of databases with regard to the influence of study design, the interpretation of tests with regard to the derivation of the NOAELs or preferably CEDs and the correction for interspecies variation is recommended. The inhalatory and dermal route also have not been sufficiently addressed and this research should be extended.
4. The probabilistic approach towards the derivation of the Critical Effect Dose has the drawback that consensus needs to be reached on the definition of Critical Effect Sizes for all toxicological endpoints that may be relevant. Research in this area should address the adversity of changes in effect parameters observed in experimental animals in order to be able to define the Critical Effect Size.

REFERENCES

Anonymous (1995) Occupational risk assessment of pesticides, Part I, Guidance for the setting of Acceptable Operator Exposure Levels, Workshop held on April 27, 1995, The Hague, The Netherlands.

Baird, S.J.S., Cohen, J.T., Graham, J.D., Shlyakhter, A.L. and Evans, J.S. (1996) Noncancer risk assessment: a probabilistic alternative to current practice. Hum. Ecol. Risk Assessment 2: 79-102

Beck, B.D., Conolly, R.B., Dourson, M.L., Guth, D., Hattis, D., Kimmel, C. and Lewis, S.C. (1993) Improvements in quantitative noncancer risk assessment. Fundam. appl. Toxicol. 20: 1-14.

Bigwood, E.J. (1973) The acceptable daily intake of food additives. CRC Crit. Rev. Toxicol., June, pp. 41-93.

Calabrese, E.J., Gilbert, C.E. (1993) Lack of total independence of uncertainty factors (Ufs): Implications for the size of the total uncertainty factor. Regulat. Toxicol. Pharmacol. 17: 44-51.

Calabrese, E.J. (1995) Uncertainty factors and interindividual variation. Regulat. Toxicol. Pharmacol. 5: 190-196.

Crump, K.S. (1984) A new method for determining allowable daily intakes. Fund. Appl. Toxicol. 4: 854-871.

Daly, A.K., Cholerton, S., Gregory, W. and Idle, J.R. (1993) Metabolic polymorphisms. Pharmac. Ther. 57: 129-160.

Dourson, M.L., Stara, J.F. (1983) Regulatory history and experimental support of uncertainty factors. Regulat. Toxicol. Pharmacol. 3: 224-238.

Dourson, M.L. and Lu, F.C. (1995) Safety/risk assessment of chemicals compared for different expert groups. Biomed. environ. Sci. 8: 1-13.

Dourson, M.L. (1996) Evolution of science-based uncertainty factors in non-cancer risk assessment. Regulat. Toxol. and Pharmacol. 24: 108-120.

EC (1992) Council Directive of 30 April 1992 amending for the seventh time Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances (92/32/EEC). Official Journal of the European Communities, L154.

EC (1993a) Council Regulation (EC) 793/93 of 23 March 1993 on the evaluation and control of the risks of existing substances. Official Journal of the European Communities, L84.

EC (1993b) Commission Directive 93/67/EEC of 20 July 1993, laying down the principles for the assessment of risks to man and the environment of substances notified in accordance with Council Directive 67/548/EEC. Official Journal of the European Communities, L227.

EC (1994) Commission Regulation (EC) 1488/94 of 28 June 1994, laying down the principles for the assessment of risks to man and the environment of existing substances in accordance with Council Regulation (EEC) No. 793/93. Official Journal of the European Communities, L161.

EC (1996a) Technical Guidance Documents in support of Directive 93/67/EEC on risk assessment of new notified substances and Regulation (EC) No. 1488/94 on risk assessment of existing substances, Parts I,II,III, IV. Luxembourg, Office for Official Oublications of the European Community L-2965, EC catalogue numbers CR-48-96-001, 002, 003, 004-EN-C.

EC (1996b) EUSES, the European Union System for the Evaluation of Substances. Bilthoven, The Netherlands, National Institute of Public Health and the Environment (RIVM). Available through the European Chemicals Bureau, Ispra, Italy.

ECETOC (1995) Assessment factors in human health risk assessment. Brussels, Belgium, European Centre for Ecotoxicology and Toxicology of Chemicals. ECETOC Technical Report no. 68.

Gevel, I.A. van de and Hakkert, B.C. (1997) Methods for interspecies extrapolation: an evaluation. Zeist, The Netherlands, TNO Nutrition and Food Research, TNO-report V97.529 (DRAFT).

Hakkert, B.C., Stevenson, H., Bos, P.M.J. and van Hemmen, J.J. (1996) Methods for the establishment of Health-Based Recommended Occupational Exposure Limits for existing substances. Zeist, The Netherlands, TNO Nutrition and Food Research, TNO report V96.463.

Hattis, D., Erdreich, L. and Ballew, M. (1987) Human variability ans susceptability to toxic chemicals - a preliminary analysis of pharmacokinetics data from normal volunteers. Risk Analysis 7: 414-426.

Health Council (1985) Advies inzake uitgangspunten voor normstelling; De inzichtelijke opbouw van advieswaarden voor niet-mutagene, niet-carcinogene en niet-immunotoxische stoffen. The Hague, The Netherlands. Health Council Report nr. 31 (in dutch).

Health Council (1996) Toxicology-based recommended exposure limits. Rijswijk, The Netherlands. Health Council Report No. 1996/12.

IPCS (1994) Assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits. Geneva, Switzerland, International Programe on Chemical Safety, WHO/UNEP/ILO, Environmental Health Criteria 170.

Kalberlah, F., Hassauer, M. and Schneider, K. (1997) Überprüfung der Datenlage im Hinblick auf eine Verbesserung der Dosis-Wirkungsbeziehung als basis zur Quantifizierung von Sicherheitsfaktoren. Freiburg, Germany, Forschungs- und Beratungsinstitut Gefahrstoffe, FoBiG GmbH (im Auftrag des Umweltbundesamtes). Bericht zum F+E-Vorhaben 11606113. Draft.

Kramer, H.J., van den Ham, W.A., Slob, W. and Pieters, M.N. (1995) Derivation of conversion factors to estimate indicative chronic NOAEL from short-term toxicity data. Bilthoven, RIVM, report no. 620110001.

Kramer, H.J., van den Ham, W.A., Slob, W. and Pieters, M.N. (1996) Conversion factors estimating indicative chronic No-Observed-Adverse-Effect Levels from short-term toxicity data. Regulat. Toxicol. Pharmacol. 23, 249-255.

Lehman, A.J., Fitzhugh, O.G. (1954) 100-Fold margin of safety. Assoc. Food Drug Off. U.S.Q. Bull. 18:33-35.

Leisenring, W. and Ryan, L. (1992) Statistical properties of the NOAEL. Regulat. Toxicol. Pharmacol. 15: 161-171.

Lewis, S.C., Lynch, J.R., Nikiforov, A.I. (1990) A new approach to deriving community exposure guidelines from "No-observed-adverse-effect levels". Regulat. Toxicol. Pharmacol. 11: 314-330.

Lu, F.C. (1979) Assessment at an international level of health hazards to man of chemicals shown to be carcinogenic in laboratory animals. In: Coulston, E.F. (ed.), Regulatory aspects of carcinogenesis, New York, Academic Press, pp. 315-328.

McColl, R.S. (1989) Biological safety factors in toxicological risk assessment. Ottawa, Canada, Report produced for the Environmental Health Directorate, Health Protection Branch, Health and Welfare Canada. Ministry of Supply and Services Canada, Cat. H49-49/1990E, ISBN 0-662-17638-3.

Naumann, B.D., Weidemann, P.A., Dixit, R., Grossman, S.J., Shen, C.F. and Sargent, E.V. (1997) Use of toxicokineticand toxicodynamic data to reduce uncertainties when setting occupational exposure limits for pharmaceuticals. Hum. Ecol. Risk Assessment 3: 555-565.

McNamara, B.P. (1976) Concepts in health evaluation of commercial and industrial chemicals. In: Advances in modern technology, 1, part1, New concepts in safety evaluation, Mehlman, M.A., Shapiro, R.E. and Blumenthal, H., eds. Washington DC, USA, Hemisphere, pp. 61-140.

Nessel, C.S., Lewis, S.C., Stauber, K.L. and Adgate, J.L. (1995) Subchronic to chronic exposure extrapolation: toxicologic evidence for a reduced uncertainty factor. Hum. Ecol. Risk Assessment 1: 516-526.

Renwick, A.G. (1991) Safety factors establishment of acceptable daily intakes. Food Add. Contamin. 8: 135-140.

Renwick, AG (1993a) Data-derived safety factors for the evaluation of food additives and environmental contaminants. Food Add. Contamin. 10: 275-305.

Renwick, AG (1993b) An analysis of the risk of exceeding the acceptable or tolerable daily intake, Regulat. Toxicol. Pharmacol. 18, 463-480.

Rulis, A.M. and Hattan, D.G. (1985) FDA's priority based assessment of food additives. Regulat. Toxicol. Pharmacol. 5: 152-174.

Schaddelee, M. (1997) Interspecies en intraspecies variabiliteit. Bilthoven, RIVM, Trainee report, RIVM.

Slob, W. (1994) Uncertainty analysis in multiplicative models. Risk Analysis 14: 571-576.

Slob, W. and Pieters, M.N. (1997a) A probabilistic approach for deriving acceptable human intake limits and human risks from toxicological studies: general framework. Bilthoven, RIVM, report no. 620110005.

Slob, W. and Pieters, M.N.(1997b) Few large, or many small groups? An evaluation of toxicological study designs using computer simulations. Bilthoven, RIVM, report no. 620110006.

Sonich-Mullin, C. (1995) Harmonisation of approaches to the assessment of risk from exposure to chemicals. Risk Newsletter, 4th quarter 1995, pp. 6-7.

Stevenson, H., Bos, P.M.J. and de Raat, W.K. (1995a) Review of applied factors to derive health-based resommended exposure levels. Zeist, The Netherlands, TNO Nutrition and Food Research Institute, TNO-report V95.092.

Stevenson, H., Bos, P.M.J. and van Hemmen, J.J. (1995b) Possibilities for a harmonised approach on assessment factors. Zeist, The Netherlands, TNO Nutrition and Food Research Institute, TNO-report V95.697.

US-EPA (1995) The use of the Benchmark Dose approach in health risk assessment. Washington, USA, US Environmental Protection Agency, Risk Assessment Forum EPA/630/R-94/007.

Vettorazzi, G. (1980) Safety factors and their application in the toxicological evaluation. In: The evaluation of toxicological data for the protection of public health. Oxford, UK, Pergamon, pp. 207-223.

Vollmer, G., Giannoni, L., Sokull-Klüttgen, B. and Karcher, W. (1996) Risk assessment, theory and practice. Proceedings of the Workshop Risk Assessment: a Workshop on Practical Experience, European Chemicals Bureau, Ispra, Italy, 27-28 March, 1996. Luxembourg, Office for Official Publications of the European Communities.

Weil, C.S. and McCollister, D.D. (1963) Relationship between short- and long-term feeding studies in designing an effective toxicity test. Journal of Agric. Food Chem. 11: 486-491.

WHO (1979) Agreed terms on health effects evaluation and risk and hazard assessment of environmental agents. Internal report of a working group. Geneva, World Health Organisation, EHE/EHC/79.19.

Wilschut, A., Houben, G.F., Hakkert, B.C. (1998) Evaluation of route-to-route extrapolation in health risk assessment for dermal and respiratory exposure to chemicals. Zeist, TNO Nutrition and Food Research Institute, TNO report V97.520.

Woutersen, R.A., Opdam, J.J.G. and Stevenson, H. (1997) Health-based recommended (occupational) exposure limits: value of the benchmark-dose approach in comparison with the No Observed Adverse Effect Level method. Zeist, The Netherlands, TNO Nutrition and Food Research, TNO-report V97.164.

ANNEX

THE ASSESSMENT FACTOR APPROACH FOR ESTABLISHMENT OF A TOXICITY PROFILE DERIVED OVERALL ASSESSMENT FACTOR VERSUS THE TRADITIONAL APPROACH USED IN RISK ASSESSMENT.

Table: Toxicity profile of active substance

Study	NOAEL $mg.kg_{bw}^{-1}.d^{-1}$	LOAEL $mg.kg_{bw}^{-1}.d^{-1}$	Effects[1]
Inhalation toxicity			
subacute, rat (28-days) (exposure: 6 hours/day 5 days/week)	100 (mg/m^3)	1000 (mg/m^3)	LOAEL: liver effects: increased weight, clinical chemistry Higher doses: effects on liver and red blood cell parameters Local effects: none
Oral toxicity			
subacute, rat (28-days)	50	500	LOAEL: liver effects: clinical chemistry Higher doses: effects on liver, kidneys and red blood cell parameters
semi-chronic, rat (90-days)	10	90	LOAEL: liver effects: increased weight, clinical chemistry and effects on red blood cell parameters Higher doses: effects on liver, kidneys, (among which microscopical changes), urinalysis and red blood cell parameters
chronic, rat (104 wk)	5	30	LOAEL: effects on the liver and kidneys (increased weight, clinical chemistry and urinalysis) Higher doses: effects on liver, kidneys (among which microscopical changes) and red blood cell parameters Tumours: liver tumours (benign) were observed at dose levels of 150 $mg.kg_{bw}^{-1}.d^{-1}$ and higher dose levels
semi-chronic, dog (90-days)	4	40	LOAEL: effects on body weight and red blood cell parameters Higher doses: effects on body weight, kidneys and red blood cell parameters
Carcinogenicity			
carcinogenicity, mouse (104 wk)	carcino-genicity study	carcino-genicity study	the test substance produced liver tumours (benign and malign) at dose levels of 300 $mg.kg_{bw}^{-1}.d^{-1}$ and above.
Teratogenicity			
gavage, rat	80 200	200 500	maternal toxicity developmental toxicity (200 $mg.kg_{bw}^{-1}.d^{-1}$ highest dose) no teratogenic effects were observed
gavage, rabbit	40 100 250	100 250	maternal toxicity developmental toxicity teratogenic effects (hydrocephaly/encephalocele; 250 $mg.kg_{bw}^{-1}.d^{-1}$ highest dose)

Table: Toxicity profile of active substance			
Study	NOAEL $mg.kg_{bw}^{-1}.d^{-1}$	LOAEL $mg.kg_{bw}^{-1}.d^{-1}$	Effects[1]
Reproduction toxicity			
oral, rat (2-generation study)	7 50 ≥250	50 250	parental toxicity developmental toxicity reproduction toxicity (250 $mg.kg_{bw}^{-1}.d^{-1}$ highest dose)
Genotoxicity			
in vitro, bacteria *in vitro*, mammalian cells *in vivo*, mammalian	negative negative negative		conclusion: the test substance has no genotoxic potential

[1] Only the adverse effects that are regarded as most important and those that determine the NOAEL are mentioned

The following factors are used for the establishment of a level for chronic inhalation exposure of the general population on basis of the subacute inhalation study in the rat:

1 *Interspecies differences*: This assessment factor is normally composed of two factors, one accounting for difference in caloric demands (experimental species factors of 4, 1.4 and 7 are used for rats, dogs and mice, respectively), and a default value of 3 accounting for remaining uncertainty. In the case an inhalation study is used as starting point, no factor for caloric demand is used, because animals breath according to their caloric demand.

2 *Intraspecies differences*: A default value of 10 is used, which compensates for differences in sensitivity within the general population.

3 *Difference in duration* of exposure between the experimental conditions and anticipated exposure pattern: The present toxicity profile demonstrates that after oral exposure in the rat, a factor of 5 could be used for extrapolation of results from subacute to semi-chronic exposure ($NOAEL_{subacute}$ of 50 $mg.kg_{bw}^{-1}.d^{-1}$ versus $NOAEL_{semi\text{-}chronic}$ of 10 $mg.kg_{bw}^{-1}.d^{-1}$) and a factor 2 could be used for the extrapolation of results from semi-chronic to chronic exposure ($NOAEL_{semi\text{-}chronic}$ of 10 $mg.kg_{bw}^{-1}.d^{-1}$ versus $NOAEL_{chronic}$ of 5 $mg.kg_{bw}^{-1}.d^{-1}$). Combination of these factors results in a factor of 10 for the extrapolation of results from subacute to chronic exposure. It is noted that this factor is considerably lower than the default value of 100, which is traditionally used.

4 *Critical effect*: The critical effect of the present substance does not need compensation for the type of critical effect; therefore an assessment factor 1 is used.

5 *Dose-response*: In case of the present substance, the available dose-response relationships do not justify compensation for the steepness of shallowness of the curve; therefore a factor of 1 is used.

6 *Confidence in the data base*: A factor may be used for limitation of the entire toxicological data base. In case of the present substance, there are no indications for such a factor.

In accordance with the above mentioned considerations the (overall) assessment factors in the table that follows are applicable on the 28-day inhalation study in the rat. It is assumed here that route-to-route extrapolation on the basis of the chronic oral test is not possible.

Aspect	Assessment factor approach	Traditional approach
1. interspecies	3	10
2. intraspecies	10	10
3. duration of exposure	10	100
4. critical effect	1	n.i.
5. dose response	1	n.i
6. confidence of the data base	1	n.i
overall assessment factor	300	10000

n.i.: not indicated.