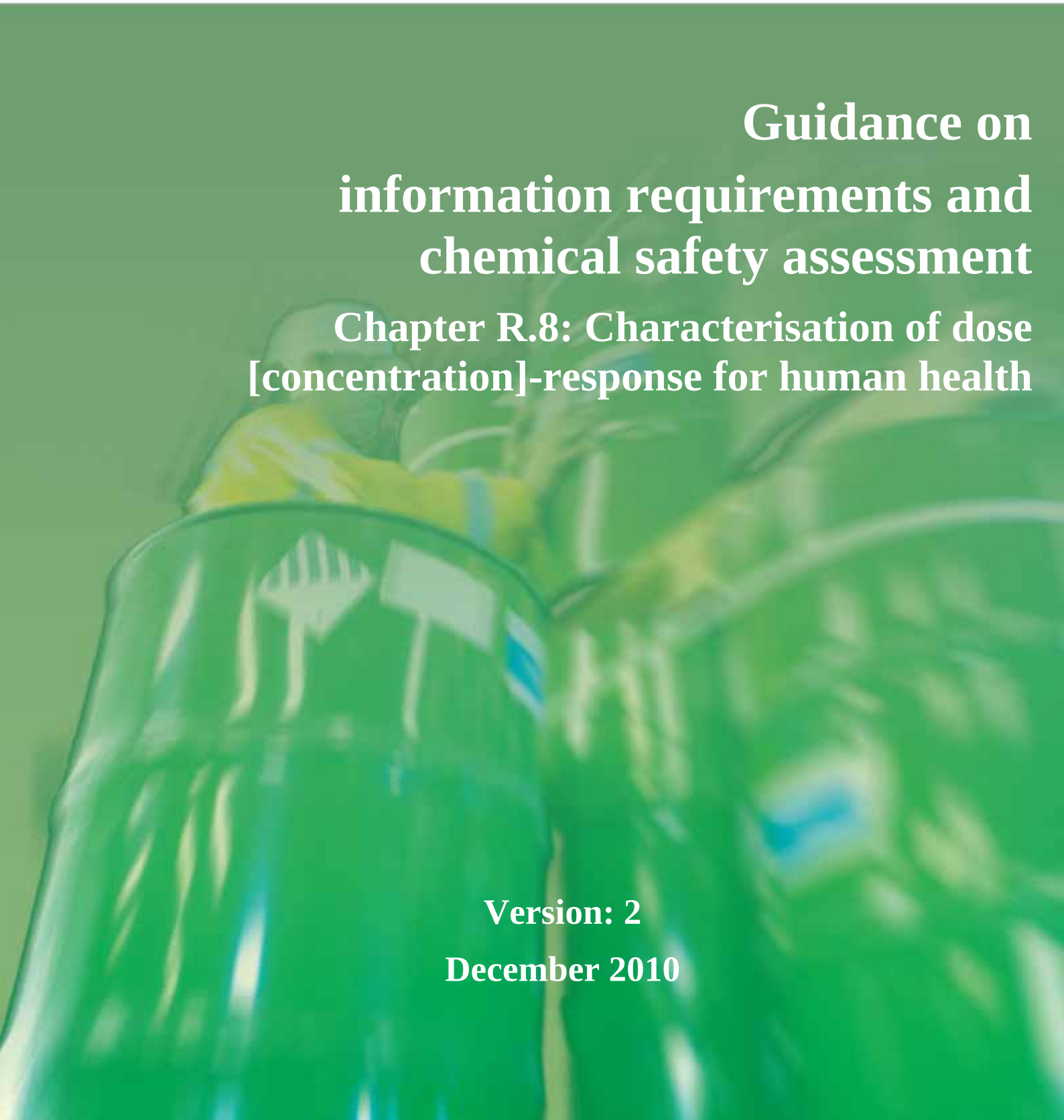


Guidance on information requirements and chemical safety assessment

Chapter R.8: Characterisation of dose [concentration]-response for human health

Version: 2

December 2010

Guidance on information requirements and chemical safety assessment
Chapter R.8: Characterisation of dose [concentration]-response for human health

Reference: ECHA-2010-G-19-EN
Publ.date: December 2010
Language: EN

If you have questions or comments in relation to this document please send them (indicating the document reference, issue date, chapter and/or page of the document which your comment refers to) using the Guidance feedback form. The feedback form can be accessed via the ECHA Guidance website or directly via the following link:
https://comments.echa.europa.eu/Comments/FeedbackGuidance.aspx

European Chemicals Agency

Mailing address: P.O. Box 400, FI-00121 Helsinki, Finland
Visiting address: Annankatu 18, Helsinki, Finland

PREFACE

This document describes the information requirements under REACH with regard to substance properties, exposure, use and risk management measures, and the chemical safety assessment. It is part of a series of guidance documents that are aimed to help all stakeholders with their preparation for fulfilling their obligations under the REACH regulation. These documents cover detailed guidance for a range of essential REACH processes as well as for some specific scientific and/or technical methods that industry or authorities need to make use of under REACH.

The guidance documents were drafted and discussed within the REACH Implementation Projects (RIPs) led by the European Commission services, involving stakeholders from Member States, industry and non-governmental organisations. These guidance documents can be obtained via the website of the European Chemicals Agency (http://echa.europa.eu/reach_en.asp). Further guidance documents will be published on this website when they are finalised or updated.

This document relates to the REACH Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006[1]

1 Corrigendum to Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC (OJ L 396, 30.12.2006); amended by Council Regulation (EC) No 1354/2007 of 15 November 2007 adapting Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) by reason of the accession of Bulgaria and Romania (OJ L 304, 22.11.2007, p. 1).

DOCUMENT HISTORY

Version	**Comment**	**Date**
Version 1	First edition	May 2008
Version 2	The present version includes the following new sections: - R.8.1.2.8 - Appendix R.8-15 - Appendix R.8-16 These new sections aim to explain in detail how to evaluate the data, extract the dose descriptors and derive the DNEL/DMEL when human data are used.	December 2010

GUIDANCE FOR IMPLEMENTING THE UPDATES

The update of this Guidance provides additional information on how to derive a DNEL/DMEL when human data are used as a starting point. It also presents the criteria to be applied in case human and animal data are available for the same endpoint/exposure patterns.

A registrant having already finalised the derivation of DNELs/DMELs based on Chapter R.8 as published in May 2008 may therefore wish to take the following advice into account:

- Carefully read the document history to be informed on what has been updated
- Check whether the changes in the guidance put into question the criteria used for the derivation of DNELs/DMELs for a given substance.

In case that the changes affect the DNEL/DMEL values already defined, the registrant is advised to evaluate to what extent an update of the existing Chemical Safety Report is required.

Convention for citing the REACH regulation

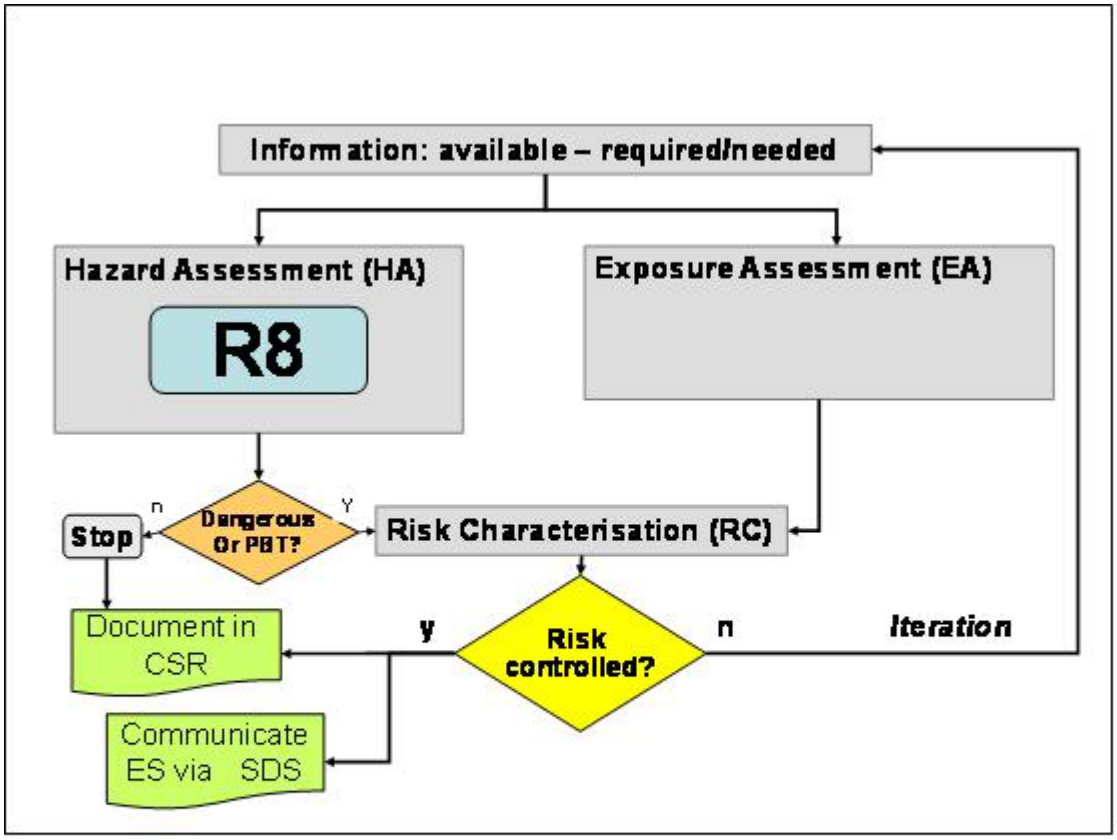
Where the REACH regulation is cited literally, this is indicated by text in italics between quotes.

Table of Terms and Abbreviations

See Chapter R.20

Pathfinder

The figure below indicates the location of Chapter R.8 within the risk assessment process.



CONTENTS

TABLES

FIGURES

EXAMPLES

APPENDICES

R.8 CHARACTERISATION OF DOSE/CONCENTRATION-RESPONSE FOR HUMAN HEALTH

R.8.1 Introduction

This section will give brief overviews of

- the requirements for dose/concentration-response in the context of feeding into the risk characterisation according to REACH,
- aspects needed to be considered when deriving DNELs (Derived No-Effect Levels) for threshold effects
- what to do when no DNEL can be derived, including, where possible for some non-threshold effects, aspects to be considered when deriving DMELs (Derived Minimal Effect Levels)
- the steps involved in this process.

It is clear that there is need for a high level of expertise to be able to follow this procedure. Detailed explanations of each step in this procedure are given in the following Sections R.8.2 to R.8.7.

R.8.1.1 Overview of legislative requirements

Under REACH manufacturers, importers and downstream users should ensure that they manufacture / place on the market / use substances in such a way that they do not adversely affect human health. REACH Annex I sets out how manufacturers and importers are to assess and document that the risks arising from the substance they manufacture or import are controlled during manufacture and their own use(s) and that others further down the supply chain can control the risks. REACH (Annex I, 1.0.1) defines the Derived No-Effect Level (DNEL), i.e. the level of exposure above which humans should not be exposed. In the risk characterisation, the exposure of each human population known to be or likely to be exposed is compared with the appropriate DNEL. The risk to humans can be considered to be controlled if the exposure levels estimated do not exceed the appropriate DNEL.

Where required, DNEL(s) shall, where possible and taking data availability into account, be derived for all substances subject to registration that are manufactured/imported/used in quantities of 10 tonnes or more per year, as part of the chemical safety assessment (CSA). DNEL(s)[2] should be documented in the chemical safety report (CSR). In case an exposure assessment and risk characterisation is required, the DNEL is subsequently to be used:

- a. in the risk characterisation part of the CSA, and
- b. for hazard communication, via extended SDS.

2 A DNEL may not be possible to set for non-threshold effects (see below). DNELs may not need to be derived for certain uses outside the scope of REACH, e.g., for cosmetics.

With respect to the derivation of DNEL(s), REACH (Annex I, 1.4.1) specifies that:

"(a) DNEL(s) shall be established for the substance, reflecting the likely route(s), duration and frequency of exposure. For some endpoints, especially mutagenicity and carcinogenicity, the available information may not enable a threshold[3], and therefore a DNEL, to be established. If justified by the exposure pattern(s), a single DNEL may be sufficient. However, taking into account the available information and, where available, the exposure scenario(s) in Section 9 of the chemical safety report[4] it may be necessary to identify different DNELs for each relevant human population (e.g. workers, consumers and humans liable to exposure indirectly via the environment) and possibly for certain vulnerable sub-populations (e.g. children, pregnant women) and for different routes of exposure. A full justification shall be given specifying, *inter alia*, the choice of the information used, the route of exposure (oral, dermal, inhalation) and the duration and frequency of exposure to the substance for which the DNEL is valid. If more than one route of exposure is likely to occur, then a DNEL shall be established for each route of exposure and as appropriate, also combined exposure through different routes needs to be addressed. When establishing the DNEL, the following factors shall, *inter alia*, be taken into account:

a. the uncertainty arising, among other factors, from the variability in the experimental data and from intra- and inter-species variation;
b. the nature and severity of the effect;
c. the sensitivity of the human (sub-)population to which the quantitative and/or qualitative information on exposure applies."

From this it follows that, based on an integration of all available and relevant human health hazard data, the DNEL can be considered as an 'overall' No-(Adverse-)Effect-Level (N(A)EL) for a given exposure (route, duration, frequency), accounting for uncertainties/variability in these data and the human population exposed. Whereas the former legislation on new and existing substances required a comprehensive risk assessment and a risk characterisation (RC) for all relevant toxicological effects, REACH requires a RC for the leading health effect (i.e., the toxicological effect that results in the most critical DNEL) for a given exposure pattern (duration, frequency, route and exposed human population) associated with an exposure scenario (ES). It is to be noted that one exposure pattern can fit to more than one ES.

For workplace exposure, there may already exist occupational exposure limits (OELs). Under certain circumstances OELs and/or the underlying information used for setting the OELs can be used to derive DNEL. See APPENDIX R. 8-13 for further guidance.

The exposure/DNEL comparison (as prescribed in REACH Annex I, 6.3 and 6.4) in principle presents a simple tool for RC, especially for downstream users who do not have the hazard data at their disposal. For any exposure scenario the risk to humans can be considered to be controlled if exposure levels do not exceed the appropriate DNEL (REACH Annex I, 6.4).

3 The threshold concept is based on the assumption that substances can only cause (noncancer) toxicological effects if the dose exceeds a certain level.

4 Note that exposure scenarios and exposure assessment are only needed for substances manufactured/imported/used in quantities >10 t/y and classified as dangerous according to Directive 67/548/EEC, or being vPvB/PBT.

Although under REACH the assessment of control of risk to humans is principally based on the exposure/DNEL comparison, it may not always be possible to derive DNEL(s) for an end-point. This is the case when:

A substance exerts its effect by a threshold mode of action, but the available data do not allow to reliably identify the threshold	This might be the case for the endpoints sensitisation and irritation.
A substance exerts its effect by a non-threshold mode of action. In that case it is generally assumed, as a default assumption that even at very low levels of exposure residual risks cannot be excluded. Consequently, a dose without potential effects cannot be established	This might especially be the case for the endpoints mutagenicity and carcinogenicity when involving a non-threshold mechanism (REACH Annex I, 1.4.1) (see Section R.7.7). It is to be noted that, as a consequence of the uncertainties in establishing an exposure level adequately controlling risk for these non-threshold substances, a substantially different approach is needed in relation to assessing and expressing risks (see Section R.8.5) as well as for their risk management. It should be noted that for carcinogens and mutagens used in the workplace, the Carcinogens and Mutagens Directive (2004/37/EC) also applies, including the hierarchy of risk management set out in that directive.
Test data (for one or more endpoints) are absent	There are 4 justified cases when no test data are needed, which are described in the following table.

For the two first cases above (two first rows) there is still need for a qualitative/semi-quantitative safety assessment in line with REACH Annex 1. Note that a DNEL-based safety assessment still may be needed for other end-points (with a DNEL), especially when it concerns other routes of exposure. Regarding the third case, the type and level of safety assessment needed depends on the case, and includes situations where;

Testing could be omitted, based on exposure arguments;	This relates to REACH Annex XI-3 and column 2 (specific rules for adaptations from column 1) in Annexes VIII-X. More detailed guidance on exposure-based waiving this can be found in Chapter R.5.One might argue that due to the no/low human exposure to a substance, the derivation of a DNEL is superfluous since the outcome of the risk assessment will in any case be negligible risk. However, for substances > 10 t/y a DNEL should be set based on the available information.
Testing could be omitted, because testing was technically not possible as a consequence of the properties of a substance;	The properties of a substance, e.g. high volatility/reactivity, can make testing dangerous or impossible. Reference to guidance on this issue to be added (adaptation of testing/ test methods). See also Chapter R.5
A substance is registered as on-site isolated intermediate[5];	Any available existing information on e.g. human health effects shall be submitted, without any additional testing. So, if no data or no adequate data are available, it may not be possible to derive a DNEL[5]. However, for on-site isolated intermediates a strict control is required (REACH, Article 17). If a substance is used only as on-site isolated intermediate and/or as transported isolated intermediate, then no DNEL is required even if information is available that would make it possible to derive one. However, the registrant might wish to develop a DNEL to help documenting that the strictly controlled conditions are sufficient.
A substance is registered as transported isolated intermediate[5].	Any available existing information on e.g. human health effects shall be submitted, together with the information specified in Annex VII for transported isolated intermediates in quantities of more than 1,000 t/y. So, if no data or no adequate data are available, it may not be possible to derive a DNEL[5]. However, for transported isolated intermediates exposure of the general population is not an issue. Given that a strict control is required, only very limited exposure may be expected for specific workers. (REACH, Article 18).

If it is not possible to derive a DNEL, then REACH (Annex I) requires, that "this shall be clearly stated and justified" (section 1.4.2) and that "a qualitative assessment of the likelihood that effects are avoided, when implementing the exposure scenario, shall be carried out" (section 6.5), in the risk characterisation part of the CSA.

REACH (Annex I, 1.1.2) only refers to a **qualitative or semi-quantitative approach** for human health effects for which no DNEL(s) can be derived (e.g. non-threshold carcinogens as illustrated in the above situation 2).

In a strictly **qualitative approach** (for e.g., genotoxic substances (i.e. non-threshold mutagens) without information on in vivo carcinogenicity) estimation of specific levels of risk for a given exposure pattern is not possible and emphasis is placed on assessing the adequacy of control of exposure in the human population of interest (e.g. workers, consumers, or humans exposed indirectly via the environment). The qualitative risk characterisation approach operates with more qualitative measures for the potency of the substance used for developing exposure scenarios with appropriate risk management measures (RMMs) and operational conditions (OCs).

[5] Note that no CSA/CSR (and thereby DNEL derivation) is formally required as part of intermediate registrations.

It is based on the principle that the more severe the nature of the hazard, the stricter the RMMs/OCs needed. This approach, in particular for high hazard substances is to some extent similar to the ALARA-principle (*as-low-as-reasonably-achievable*) originally used in the area of radiation protection. For further details of this approach, see Sections R.8.6. and E.3.4).

However, it can be useful to include in this qualitative assessment an additional **semi-quantitative** element in order to assess the likelihood that effects are avoided (as required in Annex I, Section 6.5). Thus, when no DNEL can be derived, the registrant has to conduct "a qualitative assessment of the likelihood that effects are avoided when implementing the exposure scenario" (REACH Annex I, Section 6.5). No DNEL can be derived for non-threshold mutagens/carcinogens as it is assumed that a no-effect-level cannot be established for these substances (either because there is no threshold or the threshold level cannot be determined). In such cases, and assuming that there are data allowing it, the registrant should develop a **DMEL** (derived minimal effect level), a reference risk level which is considered to be of very low concern. DMEL derived in accordance with the guidance should be seen as a tolerable level of effects and it should be noted that it is not a level where no potential effects can be foreseen. If a DMEL is not derived, the registrant should find other means for assessing/judging "...the likelihood that effects are avoided when implementing the exposure scenario" (Annex I, section 6.5).

Although not strictly required under REACH, depending on the reliability and quality of the available data (being from epidemiological studies, from animal studies and/or from alternative methods such as read-across), it is strongly recommended to develop a DMEL if data are available to allow it.

This guidance document sets out two (default) methodologies which can be applied for deriving a DMEL. For some substances very detailed information may be available (e.g. detailed data on the exact kinetics and mechanism of the carcinogenicity or detailed dose-response data). In such cases the applicant is of course allowed to use more sophisticated models simulating the low dose behaviour of such substances, based on a solid justification.

It is important to stress that a DMEL is not equivalent to a DNEL. A DNEL expresses a derived value below which exposures should be controlled – with the underlying assumption that such an exposure level would be below a no-effect-level. For non-threshold effects, the underlying assumption is that a no-effect-level cannot be established and a DMEL therefore expresses an exposure level corresponding to a low, possibly theoretical, risk.

Furthermore, it should be stressed that for carcinogens and mutagens, the Carcinogens and Mutagens Directive (2004/37/EC) requires that workplace exposures are avoided/minimised as far as technically feasible. As REACH does not overrule the Carcinogens and Mutagens Directive, the approach to controlling workplace exposure should therefore comply with this minimisation requirement.

The DMEL approach is useful when preparing chemical safety assessment to judge the remaining/residual likelihood of risks (to workers, consumers, or humans via the environment). Based on such judgement the registrant may need to refine the way he uses or recommends to use the substance by revising the relevant tentative exposure scenario(s) for use of the substance.

Contrary to the risk assessment for threshold effects, by definition for non-threshold mutagens and carcinogens a dose without a theoretical cancer risk cannot be derived.

Therefore the establishment of a reference risk level for the DMEL clearly is of societal concern and needs policy guidance. Although there is no EU legislation setting the 'tolerable' risk level for carcinogens in the society, cancer risk levels have been set and used in different contexts (See APPENDIX R. 8-14 for various values previously applied within and outside the EU). Based on these experiences, cancer risk levels of 10^{-5} and 10^{-6} could be seen as indicative tolerable risks levels when setting DMELs for workers and the general population, respectively. Alternatively, a reference exposure level judged to be of very low concern can be obtained by applying a large assessment factor to a suitable starting point from a rodent long-term cancer bioassay or from reliable human epidemiological studies.

In conclusion, for threshold substances, a DNEL is a level of exposure which should not be exceeded, indicating control. For non-threshold substances, a DMEL is a risk-related reference value that should be used to better target risk management measures. Exposure levels below a DMEL are judged to be of very low concern, due to a high likelihood that effects are avoided for the particular Exposure Scenario under consideration.

R.8.1.2 Overview of aspects to be considered in derivation of DNEL(s) / DMEL(s)

Based on the specification given in REACH (Annex I, Section 1.4.1), several aspects need to be considered in deriving DNEL(s). These are addressed below. It is to be noted that most aspects (in particular on uncertainty/variability, populations, and routes) also apply to the derivation of DMEL(s).

R.8.1.2.1 Data requirements

The derivation of DNELs is required for the chemical safety assessment (CSA) of substances manufactured/imported/used in quantities from 10 t/y onwards. For each tonnage level standard data requirements have been specified in REACH (Annex VII-X, in conjunction with Annex XI), but REACH also requires that any other relevant hazard information that is available (i.e. on other endpoints and/or from other test and non-test methods) is taken into account. Even at the lower tonnage levels the data requirements include several studies that should allow the derivation of a quantitative estimate of the dose without adverse effects, i.e. a NOAEL (e.g. 28/90 repeated dose toxicity study, screening reproductive/developmental toxicity study) and thus the derivation of a DNEL. However, the derivation of a DNEL for lifetime exposure from the minimal dataset required for the 10-100 t/y band, by default application of assessment factors for several extrapolation steps, including duration extrapolation, involves considerable uncertainty. As further toxicological information is requested at each higher tonnage level or becomes available in the scientific literature, more robust estimation becomes possible. DNEL(s) should therefore be reconsidered if further information becomes available at higher tonnage levels.

For derivation of DNELs, all available hazard information needs to be evaluated (see Chapter R.7) and, where possible, dose descriptors (N(L)OAEL, benchmark dose, etc.) need to be established (see Section R.8.2). It is to be noted that under REACH the data may originate from experiences from humans (e.g., case reports or epidemiological studies), studies with experimental animals, in vitro studies and non-testing sources ((Q)SAR), read across or chemical categories) – see Chapters R.6 and R.7.

R.8.1.2.2 Uncertainty/variability

REACH requires differences between effect assessment data and the real human exposure situation to be addressed, taking into account variability and uncertainty within and between species.

In order to address these differences, assessment factors (AF) should be applied. The applied AFs only correct for uncertainties/variability in the effect data, not for exposure uncertainties.

R.8.1.2.3 Populations

DNELs may have to be derived for workers and the general population. The general population includes consumers, and humans exposed via the environment, with the DNEL usually being identical for consumers and human via the environment. Under certain circumstances it might also be necessary to derive DNELs for certain subpopulations, i.e. covering a particular higher sensitivity (e.g. in case of indication of higher sensitivity of children for certain end-points). Another reason for assessing whether a DNEL covers a specific sub-population is when there is a specific exposure of this sub-population, e.g., exposure of children via toys, requiring a sub-population-specific assessment.

It is not always necessary to derive DNELs for all mentioned populations. Depending on the exposure pattern, only DNELs for the relevant populations will have to be derived. It has to be justified why the selected populations have been considered as relevant (and others as irrelevant). The DNELs should be named, e.g., *worker*-DNEL or *general population*-DNEL (see Table R. 8-1 for examples).

As already noted, APPENDIX R. 8-13 gives further guidance on how to address the situation where an occupational exposure limit (OEL) exists for workplace exposures.

R.8.1.2.4 Routes

In view of the anticipated exposure routes for the various populations, DNELs may have to be derived for oral exposure (consumers/human via the environment), inhalation exposure (workers/consumers/human via the environment), dermal exposure (workers/consumers, and potentially humans via the environment, e.g., via contaminated soil). As appropriate also combined exposure needs to be addressed (see Section E.3.5).

It is not always necessary to derive DNELs for all mentioned routes. Depending on the exposure pattern, only DNELs for the relevant routes of exposure will have to be derived. It has to be justified why the selected exposure routes have been considered as relevant (and others as irrelevant). In defining the DNEL, the route should be added last, e.g., worker-DNEL long-term for *dermal route* (Table R. 8-1).

R.8.1.2.5 Duration of exposure

Depending on the exposure scenario, the exposure duration can vary from a single event to an exposure for several days/weeks/months per year, or it might even be continuous (as is, e.g., the case for humans exposed via the environment). Since the duration of exposure will often have an impact on the effect(s) that may arise, DNELs may have to be derived for various exposure

durations, thereby matching as closely as possible the exposure duration in the toxicity study with the exposure duration in the exposure scenario.

Two main types of DNELs can be distinguished, $DNEL_{long\text{-}term}$ and $DNEL_{acute}$.

A $DNEL_{long\text{-}term}$, i.e., a DNEL for effects that occur upon repeated exposure, shall always be derived. Toxicity studies that give information on these possible 'long-term' effects of a substance are: repeated dose toxicity studies, reproductive toxicity studies (including developmental toxicity studies), and carcinogenicity studies. 'Long-term' is here used as a more general term, including, e.g., sub-chronic (usually 90 days) as well as chronic (usually 1.5 - 2 years) studies.

Given that often N(L)OAELs in toxicity studies decrease with increasing exposure duration, a DNEL based on a N(L)OAEL from a chronic toxicity study will generally be lower than a DNEL based on a N(L)OAEL from a sub-chronic, sub-acute or acute toxicity study. Thus, in general a DNEL established for chronic exposure will be the lowest DNEL, covering also shorter than chronic exposures. Therefore, for most substances and exposure scenarios the $DNEL_{long\text{-}term}$ will be sufficient for controlling risks. In defining the DNEL, the duration should be mentioned directly after DNEL, e.g., worker-$DNEL_{long\text{-}term}$ for dermal route.

Note that the repeated exposure resulting from a certain exposure scenario is to be expressed as the actual daily dose, bearing in mind that for workers a day is 8 hours, for human via the environment a day is 24 hours, and for consumers a day is 1-24 hours (depending on the scenario, e.g., type of consumer product). The actual daily dose is *independent* of the exposure frequency. This means that if for a certain scenario, worker or consumer exposure is for instance only for a number of days per year, the exposure value is the actual dose on the exposure days, and not the daily dose averaged out (and thus divided!) over the whole year.

The establishment of an acute toxicity DNEL set for effects occurring after a single exposure of a few minutes up to 24 hours is not only cumbersome (there is no established consensus methodology) and resource-intensive but probably unnecessary, as the long-term DNEL is normally sufficient to ensure that these effects do not occur. It is therefore proposed that if an acute toxicity hazard (leading to C&L) has been identified, a DNEL for acute toxicity is only established for the effects of peak exposures as these peaks can be significantly higher than the average daily exposure and the long-term DNEL (to be complied with *on average* over e.g. a working day) may be insufficient to limit them. Overall, therefore, a DNEL for acute toxicity should be derived if an acute toxicity hazard (leading to C&L) has been identified and there is a potential for high peak exposures, for instance when sampling or connecting/disconnecting vessels. This is most relevant for workers exposed to high peak concentrations of volatile and toxic substances, but may in some cases also be relevant for consumers. High peak exposures are usually assessed for the inhalation route only, so this guidance outlines how to set acute toxicity DNELs for the inhalation route.

Acute toxicity studies in animals generally concern single oral or dermal administration of substance, or inhalation exposure for four hours. A $DNEL_{acute}$ can generally be defined as a DNEL for effects that occur after exposure for a short period of time (from minutes to a few hours). The potential for short-term high level (i.e. peak) inhalation exposure is of most concern for workers, and hence, the occupational exposure assessment should always consider the possibility for such peak inhalation exposures, as these peaks could potentially be significantly above the typical (daily average) exposure level. If a DNEL for acute inhalation toxicity needs to be established (based on the toxicological profile of the substance concerned), this should be derived only for a specified fraction of the daily exposure duration (usually 15 minutes for workers) (see Sections R.7.4, R.8.2, and APPENDIX R. 8-8 for further information on setting DNELs for acute toxicity).

For peak inhalation exposure to volatiles, the available human experience (e.g., case studies) should be assessed. Information on effects occurring after peak exposure can thus be obtained from human experience and from acute toxicity studies in animals.

It is to be noted that 'acute' effects may manifest themselves immediately, but also considerable time after the exposure. Effects occurring early after onset of exposure in some target organ toxicity studies can also be useful, e.g. neurotoxicity, irritation and sensitisation studies, mutagenicity studies, and also repeated dose toxicity studies and reproductive toxicity studies (including developmental toxicity studies). For inhalation exposure for periods longer than 15 minutes, the long-term DNEL should be used.

For the dermal and oral exposure routes, 'short-term' exposures should normally be assessed using the long-term DNELs. However, for some substances, it may also be relevant to derive a $DNEL_{acute}$ for single dermal and/or oral exposures, in general following the principles outlined in APPENDIX R. 8-8.

R.8.1.2.6 Systemic and local effects

Depending on the substance, DNELs may have to be established for systemic effects, for local effects or for both.

- A *local effect* is an effect that is observed at the site of first contact, caused irrespective of whether a substance is systemically available.
- A *systemic effect* is defined as an effect that is normally observed distant from the site of first contact, i.e., after having passed through a physiological barrier (mucous membrane of the gastro-intestinal tract or of the respiratory tract, or the skin) and becomes systemically available.

It should be noted, however, that toxic effects on surface epithelia may reflect indirect effects as a consequence of systemic toxicity or secondary to systemic distribution of the substance or its active metabolite(s).

A DNEL should preferably cover both systemic and local effects. DNELs for systemic effects can in principle be based on all types of studies, unless low dose local effects prevent appropriately high systemic exposure to occur. For DNELs covering local inhalation and local dermal effects, however, route-specific data need to be available. If separate DNELs are set for local and systemic effects, the DNELs have to be specified by adding local or systemic to the DNEL (e.g., worker-DNEL long-term for dermal route-*systemic)* (Table R. 8-1).

R.8.1.2.7 Units

DNELs should generally be expressed as external values. Thus, for substances with inhalation as the single or major route of exposure, external values is preferred as they are more easily interpreted in compliance assessment of use conditions when mostly only external exposure estimates are available. Furthermore, for local effects, which per definition cannot be expressed in internal values, external values must be used. The units to be used are given in the footnotes of Table R. 8-1 (below).

However, the DNEL may also be expressed as internal biomarker values, but this only applies to the limited number of substances where internal values, i.e. biomonitoring data (e.g., biomarkers), are available and have been reliably associated with effects. In general, if having both biomonitoring

and monitoring data, and effects data corresponding to both types of exposure data, the most appropriate and/or reliable data/method should be used for setting the DNEL. When deriving an internal biomarker DNEL, it has to be clearly indicated that it is a biomarker value, e.g., by mentioning biomarker after DNEL ($DNEL_{biomarker}$).

The body weights to be used in the calculations are 60 and 70 kg for the general population and for workers, respectively.

Table R. 8-1 DN(M)ELs that may need to be derived, and examples on the nomenclature

Exposure pattern	DNEL/DMEL (appropriate unit)	
	Workers	**General population** [3]
Acute – inhalation, systemic effects[1]	worker-DNEL acute for inhalation route-systemic	General population-DNEL acute for inhalation route-systemic
Acute – dermal, local effects[2]	worker-DNEL acute for dermal route-local	General population-DNEL acute for dermal route-local
Acute – inhalation, local effects[2]	worker-DNEL acute for inhalation route-local	General population-DNEL acute for inhalation route-local
Long-term – dermal, systemic effects[1]	worker-DNEL long-term for dermal route-systemic	General population-DNEL long-term for dermal route-systemic
Long-term – inhalation, systemic effects[1]	worker-DNEL long-term for inhalation route-systemic	General population-DNEL long-term for inhalation route-systemic
Long-term – oral, systemic effects[1]	Not relevant	General population-DNEL long-term for oral route-systemic
Long-term – dermal, local effects[2]	worker-DNEL long-term for dermal route-local	General population-DNEL long-term for dermal route-local
Long-term – inhalation, local effects[2]	worker-DNEL long-term for inhalation route-local	General population-DNEL long-term for inhalation route-local

[1] Units for systemic exposure are mg/m^3 for inhalation, and mg/kg bw for oral and dermal exposure

[2] Units for local effects are mg/m^3 for inhalation; and for dermal exposure: mg/cm^2 skin, mg/person/day (e.g., calculated based on the deposited amount per cm^2 times the actually exposed body area), or a measure of concentration (% or ppm)

[3] General population includes consumers and humans via the environment. In rare cases it may also be relevant to derive a DNEL for specific subpolulations, such as children.

R.8.1.2.8 Human data as a source for the derivation of a DNEL and/or a DMEL

Since DNELs and DMELs are used in the assessment of risks to humans, human data are an appropriate basis also for the derivation of DNEL/DMEL. Human data are valuable as a source of hazard information because they apply directly to the human species, and the mode of action (MoA) is usually relevant. As a consequence, no inter-species assessment factor is needed when human data are used for derivation of DNEL/DMEL. Furthermore, human data have in most cases been obtained from relevant exposure conditions and are based on an adequate route of exposure. In addition, human data most often come from studies covering a more heterogeneous sample of the population than animal studies carried out on inbred strains. Nevertheless, the quality of the human data needs to be ensured. Under REACH human data, when available and relevant, are used in the

human hazard assessment and as part of the Chemical Safety Assessment as described in Annex I of the REACH Regulation. More specifically, according to Annex I the human health hazard assessment comprises four steps:

Step 1: Evaluation of non-human information.

Step 2: Evaluation of human information.

Step 3: Classification and labelling.

Step 4: Derivation of DNELs.

Furthermore, according to the provisions of Annex XI of the REACH Regulation "Historical Human Data"[6] can be used to adapt the standard testing requirements of Annexes VII to X provided that the quality of the data is properly assessed and found to be adequate.

Human data can come from analytical epidemiology studies, descriptive or correlation epidemiology studies, case reports, clinical studies, poison centre information, occupational disease registries or other occupational surveillance systems. When they are already available, well-conducted controlled human exposure studies in volunteers, including low exposure toxicokinetics studies, can also be used in risk assessment. However, few human experimental toxicity studies are available due to the practical and ethical considerations involved in deliberate exposure of individuals. Such studies, e.g. studies carried out for the authorisation of a medical product have to be conducted in line with the World Medical Association Declaration of Helsinki, which describes the general ethical principles for medical research involving human subjects (World Medical Association, 2000). It is emphasised that testing with human volunteers is strongly discouraged, but when there are good quality data already available they should be used as appropriate, in well justified cases (see Chapter R.4).

Human data differ from animal data in that they are mostly derived from observational (non experimental) studies in contrast to controlled experimental animal studies. This has profound consequences for the reviewing and handling of data. In experimental studies data quality is controlled a priori by the experimental study design while the relevance to humans needs to be assessed a posteriori. In observational human studies, data quality control is done in connection with data analysis (a posteriori) with focus on the validity of the data. Furthermore, animal studies are done in inbred strains, whereas epidemiological studies are done on heterogenic populations. This implies that the process to arrive at a dose descriptor from human studies is somewhat different from that of obtaining a dose descriptor from animal studies. Quality considerations in particular differ from those for experimental studies and the accuracy of the exposure information is an important issue. APPENDIX R.8-15 gives guidance for all the phases to arrive at a dose descriptor and to derive DNEL/DMEL from human data (see Figure R.8-6 in APPENDIX R.8-15). The process leading to the identification of the leading health effect and the associated levels of exposure is also described in APPENDIX R.8-15. There are specific uncertainties that deserve attention when using human data. These include the influence of bias, confounding from mixed exposures and other risk factors and accuracy of the exposure information. Therefore special expertise is needed when using epidemiological or other human data for obtaining DNELs/DMELs.

The term "dose descriptor" is used to designate the exposure level (dose or concentration) that corresponds to a quantified level of risk of a health effect in a specific study. In animal studies common dose descriptors for threshold effects are NOAEL (No Observed Adverse Effect Level) or

[6] Historical Human Data is a term used by REACH. It refers to already available human data.

LOAEL (Lowest Observed Adverse Effect Level), while examples of dose descriptors of non-threshold effects are T25 and BMD10. For epidemiological or other human data, typical dose descriptors for threshold effects are exposure levels for which health effects are not observed or are observed (NOAEL or LOAEL, or NOAEC etc.) as well. For non-threshold effects the dose descriptors are often expressed as levels of exposure that are associated with a Relative Risk (RR) or comparable relative risk metrix such as Odds Ratio (OR), Standardised Mortality Ratio (SMR) or Standardized Incidence Ratio (SIR). Nevertheless, levels of exposure associated with such relative risk metrics can also be used as dose descriptors for threshold effects.

Human data are in principle the most relevant source of information on human toxicity (see Chapter R.4). Since there may be limitations in reliability of human studies (e.g. problems in study design, analysis and reporting as well as limited coverage of the different target organs), they are normally considered together with animal and other data. For many chemicals, both human data and animal data are available (see Money 2007 for a summary on the use of human data under the Existing Substances Regulation). Therefore an integrated approach is required. This also applies to DNEL/DMEL derivation. APPENDIX R.8-15, Phase 9, provides advice on the integration of animal and human data. In this approach, the criteria for the selection of useful data are the *quality* and *relevance* of the data and the *level of the DNELs/DMELs* obtained from human versus animal data. It is assumed that before the integration phase is started, the relevant animal studies have been examined and the DNELs/DMELs have been derived from them according to the guidance given in Sections R.8.2 to R.8.6.

An important element of the CSR is the justification and documentation of the choices made in the DNEL/DMEL derivation, in particular, when choosing the assessment factors, dose descriptors and the leading health effect.

R.8.1.3 Overview of DNEL/DMEL-derivation, critical DNEL(s)/DMEL, other measures of potency

The process for deriving DNEL/DMEL and/or arriving at other measures for the potency can be derived as follows:

Step 1: Gather typical dose descriptors (e.g. N(L)OAEL, BMD, LD50, LC50, T25, BMD(L)10, OR, RR....) from all available and relevant studies on the different human health endpoints (see Section R.8.2) and/or other information of the potency when no dose descriptor is available.

Step 2: Decide on mode of action (threshold or non-threshold) and which next step(s) to choose (i.e., step 3-1, 3-2, and/or 3-3) (see Section R.8.3)

Step 3-1: Derive, where possible, DNEL(s) for threshold endpoints by
a) selection of relevant dose-descriptor(s) for the endpoint concerned
b) modification, when necessary, of relevant dose descriptor(s) per endpoint to the correct starting point (i.e., correct the unit of exposure)
c) application, when necessary, of assessment factors to the correct starting point to obtain endpoint-specific DNEL(s) for the relevant exposure pattern (duration, frequency, route and exposed human population) (see Section R.8.4)

Step 3-2: If possible, derive DMEL(s) for non-threshold endpoints by
a) selection of relevant dose-descriptor(s) for the endpoint concerned

b) modification, when necessary, of relevant dose descriptor(s) per endpoint to the correct starting point (i.e., correct the unit of exposure)
c) application, when necessary, of assessment factors/high to low dose risk extrapolation factor[7] to the correct starting point to obtain endpoint-specific DMEL(s) for the relevant exposure pattern (duration, frequency, route and exposed human population) (see Section R.8.5)

Step 3-3: Follow a more qualitative approach when no dose descriptor is available (see Section R.8.6)

Step 4: Select the leading health effect(s) and the corresponding DNEL, DMEL or other qualitative/semi-quantitative description (see Section R.8.7)

Please, note that the four step process described here is considered the most adequate for the derivation of DNEL/DMEL from animal studies. In case human data are available, the nine phase approach described in APPENDIX R.8-15 is recommended.

R.8.2 Step 1: Gather typical dose descriptors and/or other information on potency

Step 1: Gather typical dose descriptors (e.g. N(L)OAEL, BMD, LD50, LC50, T25, BMD(L)10, OR, RR....) from all available and relevant studies on the different human health endpoints and/or other information on potency when no dose descriptor is available

Dose response assessment - Derivation of a N(L)OAEL/BMD

It is generally agreed that many of the adverse effects of health caused by substances are not expressed until the substance, or an active metabolite, reaches a threshold concentration in the relevant organ. Whether or not this threshold concentration is reached is related to the level of exposure of the organism (human or test animal) to the substance: for a given route of exposure, there will be a threshold exposure level which must be attained before effects are induced. The threshold exposure dose or concentration may vary considerably for different routes of exposure, and for different species because of differences in toxicokinetics and possibly also in mechanisms of action. The observed threshold dose or effect level in a toxicity test will be influenced by the sensitivity of the test system and is a surrogate for the true so-called No Adverse Effect Level (NAEL).

The No Observed Adverse Effect level (NOAEL) identified in a particular test will be simply the highest dose level or concentration of the substance used in that test at which no statistically significant adverse effects were observed, i.e. it is an operational value derived from a limited test. For example if the dose levels of 200, 50, 10 and 5 $mg \cdot kg^{-1} \cdot day^{-1}$ of a substance have been used in a test and adverse effects were observed at 200 and 50 $mg \cdot kg^{-1} \cdot day^{-1}$ but not at 10 or 5 $mg \cdot kg^{-1} \cdot day^{-1}$, the derived NOAEL will be 10 $mg \cdot kg^{-1} \cdot day^{-1}$. Thus, the NOAEL and LOAEL (lowest observed adverse effect level) values for a given study will depend on the experimental study design, i.e. the selection of dose levels and the spacing between doses.

[7] The term assessment factor is used because of it being a neutral term. However, these factors can in the DMEL-approach also be viewed as 'correction factors' and 'uncertainty factors'.

If there are several studies addressing the same effects from which different NOAELs could be derived, normally the lowest relevant value should be used in DNEL derivation. When it is not possible to identify the NOAEL in a repeated dose study, the "lowest observed adverse effect level" (LOAEL) should be used in the risk characterisation. If a NOAEL becomes available subsequently, from another test, the risk characterisation should be re-addressed and revised, if necessary, in the light of the new information.

The sensitivity of a study, (which is related to the toxicological endpoint, the potency of the toxic substance, the exposure period and frequency, the variability within the species, the number of dose groups and the number of animals per dose group) may limit the extent to which it could be possible to derive a reliable NOAEL from a particular test. In these cases where it is impossible to derive a NOAEL, at least a LOAEL should be identified.

It is recognised that the NOAEL is not very accurate with respect to the degree to which it corresponds with the (unknown) true NAEL. Also, the data obtained at one dose (NOAEL) are used rather than the complete dose response data set (Woutersen et al., 1997). In case sufficient data are available, the shape of the dose response curve should be taken into account. In the case of a steep curve the derived NOAEL can be considered as more reliable (the greater the slope, the greater the reduction in response to reduced doses); in the case of a shallow curve, the uncertainty in the derived NOAEL may be higher and this has to be taken into account in the DNEL derivation. If a LOAEL has to be used, then this value can only be considered reliable in the case of a very steep curve. In response to the general call for consideration of the dose response curve as a whole rather than to use only the data obtained at one dose (NOAEL) for risk characterisation, alternatives for dose-response assessment have been proposed such as the benchmark dose (BMD) concept (Crump, 1984; Gaylor, 1988; US EPA, 1995; Slob and Pieters, 1998) and categorial regression (Hertzberg, 1989).

Advantages of this approach over the NOAEL are:

- the Benchmark dose is derived using all experimental data and reflects the dose-response pattern to a greater degree;
- the Benchmark dose is independent of predefined dose levels and spacing of dose levels;
- the Benchmark approach makes more reasonable use of sample size, with better designs resulting in higher Benchmark doses.

A disadvantage of this new method is the uncertainty with respect to the reliability of the approach in case results are obtained from toxicity studies performed according to the requirements defined in current guidelines (Annex V to the Directive 67/548/EEC methods[8], OECD guidelines). For the derivation of reliable dose-response relationships, the classical study design of three dose groups and a vehicle control group is far from ideal, especially if one considers the unfavourable possibility that in a particular experiment, adverse effects may be identified only at the highest dose level.

An improved benchmark model fit would be possible by increasing the number of dose groups without changing the total number of animals in the test. However, such a change in study design would generally no longer allow a proper derivation of a NOAEL. Thus, in practice, the NOAEL and the benchmark concepts appear to be incompatible.

The BMD can be used in parallel to derivation of a NOAEL or as an alternative when there is no reliable NOAEL. In addition, the Benchmark dose (BMD) approach is, when possible, preferred

[8] NB! This will in the future be repealed by a new Test Guidelines Regulation.

over the LOAEL-NAEL extrapolation (See also US EPA, 1995; Barnes et al., 1995; Slob, 1999; Vermeire et al., 1999, for further details on the BMD approach).

Unless a threshold mechanism of action is clearly demonstrated, it is generally considered prudent to assume that thresholds cannot be identified in relation to mutagenicity, genotoxicity, and genotoxic carcinogenicity, although a dose-response relationship may be shown under experimental conditions. Details on the derivation of different dose descriptors (T25, BMD(L)10) based on animal studies for non-threshold carcinogens are given in APPENDIX R.8-6.

It is possible that for a particular endpoint data from more than one study are available (e.g. in different species, with different durations), and that these studies are all relevant and appropriate (with respect to conduct, tested species relevant for humans, etc.). The dose descriptor can also be set based on experience of toxic effects in humans[9]. Since it is not possible to know beforehand which of these dose descriptors will turn out be critical for the endpoint-specific DNEL, it might sometimes be relevant to derive DN(M)EL for more than one study per endpoint. Particularly when there is exposure through several routes, there is need for a DNEL for each exposure route (see Section R.8.7). The choice of key studies and derivation of DN(M)ELs will depend on expert judgement, including the use of a weight of evidence approach. In any case the choice of one or more dose descriptors should be justified. Some special considerations on the identification of the typical dose descriptor for some of these endpoints are given below.

R.8.2.1 Dose descriptor for acute toxicity, irritation/corrosion, skin sensitisation, reproductive toxicity

As compared to the straight-forward derivation of DNEL for repeated dose toxicity, it may be more difficult for the endpoints acute toxicity, irritation/corrosion, skin sensitisation, and reproductive toxicity (see Chapter R.7). For instance, whereas the ideal starting point for the derivation of the acute toxicity DNEL should be the NOAEL or LOAEL for sub-lethal effects, such as local respiratory irritation caused by cytotoxicity or CNS depression, oftentimes only data from 'LD50-studies' are available. Likewise, there is usually no strict NOAEL or NOAEC identified in studies on irritation, corrosion, or sensitisation. Therefore, in many, or even most cases, the lack of NOAEL(C), dose-response or indication of potency will require that a more qualitative approach is followed (see Section R.8.6). However, for cases where good data (i.e., dose descriptors) are available, which has to be evaluated on a case-by-case basis, allowing setting DNELs for these end-points additional guidance on setting DNELs has been provided in APPENDIX R. 8-8 to APPENDIX R. 8-12. For these cases, the registrant needs to justify the approach within the context of the available data. For endpoints not mentioned in these appendices, the normal guidance should be followed.

Next step is to collect in a table (see Table R. 8-14 of APPENDIX R. 8-1) all available dose descriptors (or, if a dose descriptor can not be identified, other information potency) from the available data for the different human health endpoints, thereby making a distinction between local and systemic effects where applicable.

R.8.3 Step 2: Decide on mode of action (threshold or non-threshold) and which next step(s) to choose

[9] Further guidance on using human data for setting DNELs is under development.

Before actually deriving DNEL(s) or DMEL(s) on the basis of the derived dose descriptors, it is important to determine whether the substance exerts its effect by a non-threshold mode of action. In other words: is the substance a non-threshold mutagen or non-threshold carcinogen?

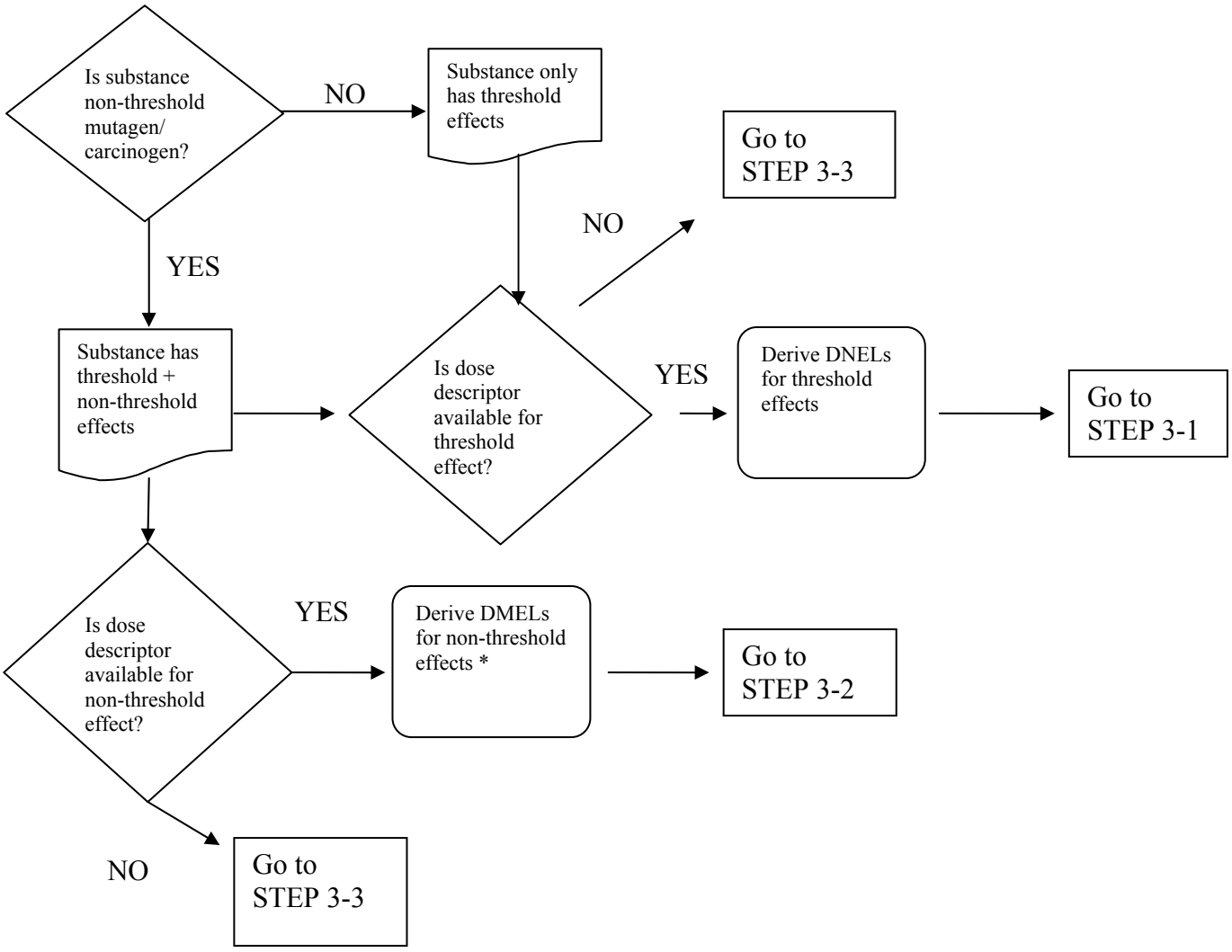
- If the answer is NO, the substance exerts its effect by a threshold mode of action. In principle, DNELs will have to be derived for the different threshold endpoints, based on the most relevant dose descriptors for these endpoints (step 3-1). When the available data do not allow to reliably identify the threshold, and thus no quantitative dose descriptor and DNEL can be derived, a more qualitative approach has to be taken (step 3-3).
- If the answer is YES, the substance exerts its effects entirely or partly by a non-threshold mode of action (for mutagenicity/carcinogenicity). In addition, the substance may also partly exert its effect by a threshold mode of action (for other human health endpoints). For the non-threshold effects, in principle any level of exposure carries a risk and thus no dose without effect can be established (see also Section R.8.1.1). Therefore, for these effects DMEL(s) should be derived (if data allow that), on the basis of the most relevant dose descriptors (step 3-2). For the threshold effects of the substance, DNELs will have to be derived, on the basis of the most relevant dose descriptors (step 3-1). When no dose descriptor is available and thus no DMEL/DNEL can be derived for a certain endpoint, a more qualitative approach has to be taken (step 3-3).

It is to be noted that the decision on a threshold and a non-threshold mode of action may not always be easy to make, especially when, although a biological threshold may be postulated, the data do not allow identification of it. If not clear, the assumption of a non-threshold mode of action would be the prudent choice.

For mutagens/carcinogens, it should be stressed that the Carcinogens and Mutagens Directive (2004/37/EC) requires that occupational exposures are avoided/minimised as far as technically feasible. As REACH does not overrule the Carcinogens and Mutagens Directive, the approach to controlling workplace exposure should therefore comply with this minimisation requirement.

Step 2 is exemplified in the following flow-chart (Figure R. 8-1).

Figure R. 8-1 Illustration of how to perform the assessment dependent on the type of mechanism of action (threshold and/or non-threshold)



* if relevant, apply Directive 2004/37/EC

DNELs should normally be developed in parallel to the DMEL. This can be particularly relevant when the DMEL and DNEL concerns different exposure pathways.

R.8.4 Step 3-1: Derive DNEL(s) for threshold endpoints

As indicated in Section R.8.1.3, DNELs for threshold endpoints are derived following a procedure consisting of the following steps:

a. selection of relevant dose-descriptor(s) for the endpoint concerned (see further Section R.8.4.1)
b. modification, when necessary, of relevant dose descriptor(s) per endpoint to the correct starting point (i.e., correct the unit of exposure) (see further Section R.8.4.2)
c. application, when necessary, of assessment factors to the correct starting point to obtain endpoint-specific DNEL(s) for the relevant exposure pattern (duration, frequency, route and exposed human population) (see further Section R.8.4.3)

R.8.4.1 a) Select the relevant dose-descriptor(s) for the endpoint concerned

For each human health threshold endpoint, one or more dose-descriptors from the available data have been compiled in step 1 (see Section R.8.2). When more than one dose descriptor is selected steps b and c have to be considered for each of these for that endpoint.

R.8.4.2 b) Modify, when necessary, the relevant dose descriptor(s) per endpoint to the correct starting point

In a few situations, the effects assessment is not directly comparable to the exposure assessment in terms of exposure route, units and/or dimensions. In these situations, it is necessary to convert the dose descriptor for the threshold effect (e.g. N(L)OAEL, benchmark dose, LD/LC50) into a correct starting point (i.e., correct the unit of exposure, e.g. corrected N(L)OAEL). This applies to the following situations:

1. If for a given human exposure route there is a dose descriptor for the same route in experimental animals but for that particular exposure route there is a difference in bioavailability between experimental animals and humans at the relevant level of exposure.
2. If for a given human exposure route there is not a dose descriptor for the same route (in experimental animals or humans).
3. Differences in human and experimental exposure conditions.
4. Differences in respiratory volumes between experimental animals (at rest) and humans (light activity).

It should be noted that modification is not appropriate in cases where human exposure is evaluated based on biological monitoring data. In such cases (availability of valid biomonitoring data), the calculation of DNEL values can be straightforward if studies in animals or humans are available which relate the effect directly or indirectly to the biomonitoring metric. Modification is generally also not needed when the dose descriptor is based on human data (e.g., case studies).

Ad 1.

The default situation, in the absence of information, is to assume the same bioavailability for experimental animals and humans for a particular exposure route. However, when available information indicates that at the relevant level of exposure humans absorb less (or more) than experimental animals, the dose descriptor needs to be corrected for this difference in bioavailability.

Ad 2.

If no adequate experimental effect data are available on the relevant route of exposure for the population under consideration, route-to-route extrapolation might be an alternative, however only for systemic effects, not for local effects (e.g. irritation of the lungs following inhalation of a substance).

Even for systemic effects route-to-route extrapolation is considered appropriate only under certain conditions (e.g. no first pass effects). Guidance on route-to-route extrapolation of toxicity data when assessing health risks of chemicals has for example been produced by IGHRC (2006). When route-to-route extrapolation is considered appropriate, corrections should be made for differences in kinetics and metabolism. In general, it is difficult to quantify differences in metabolism, excretion and distribution, so in practice only differences between the different routes as determined by the percentages of absorption into the systemic circulation can be accounted for.

It is to be noted that route-to-route extrapolation is associated with a high degree of uncertainty and should be conducted with caution relying on expert judgment (see APPENDIX R. 8-2). Additionally, relevant testing of the relevant exposure route should also be considered, or the use of PBPK modelling (see Section R.8.4.3.2).

Default absorption values have been proposed for the different routes of exposure (see Section R.7.12. on toxicokinetics), but substance-specific data on absorption via the different routes are to be preferred. Such information may for instance be generated based on considerations of the chemical structure.

In the absence of these data for both the starting route and the end route (the route to which the extrapolation is being made), worst case assumptions have to be made. Worst case in this context will be obtained assuming a limited absorption for the starting route, leading to a low (conservative) internal NOAEL. To secure a conservative external NOAEL a maximum absorption should there after be assumed for the end route, leading to a low external NOAEL. It is proposed, thus, in the absence of route-specific information on the starting route, to include a default factor of 2 (i.e. the absorption percentage for the starting route is half that of the end route) in the case of oral-to-inhalation extrapolation. The inclusion of this factor 2 means for example that 50% (instead of 100%) absorption is assumed for oral absorption, and 100% for inhalation. Note that if data on the starting route (oral) are available these should be used, but for the end route (inhalation), the worst case inhalation absorption should still be assumed (i.e. 100%). Note that this does not apply if there is a first pass effect, if there is non-resorption, or for bolus effects.

No default factor should be introduced (i.e. factor 1) in case of inhalation-to-oral extrapolation, because a two times higher oral compared to inhalation absorption appears on empirical grounds not justified.

On the assumption that, in general, dermal absorption will not be higher than oral absorption, no default factor (i.e. factor 1) should be introduced when performing oral-to-dermal extrapolation.

The other possible, but less usual, situations of route-to-route extrapolation (i.e. inhalation-to-dermal and *vice versa*) should be handled on a case-by-case basis.

Ad 3.

The exposure conditions for experimental animals in a toxicity study may differ from that of target populations. For example, in repeated dose inhalation studies exposure normally is 6 hours per day, which differs from that for workers (assumed 8 hours per day), human via the environment (assumed 24 hours per day), and consumers (assumed 1-24 hours per day, depending on exposure scenario). If the toxic effect is driven by the total (accumulated) dose, or depends on both total dose and the exposure concentration, concentration–time corrections (i.e. time scaling) have to be applied. Time scaling is not appropriate when the toxic effect is mainly driven by the exposure concentration (as for irritation). A useful tool for time scaling is the modified Haber's law ($C^n \times t = k$, where 'C' is the concentration, 'n' is a regression coefficient, 't' is the exposure time and 'k' is a constant) (see Section R.7.4 and APPENDIX R. 8-8 for further explanations). So, when for instance a NOAEC is available from a rat 6 h/d inhalation study, in most cases for workers this NOAEC needs to be corrected by a factor 0.75 (6/8), for human via the environment by a factor of 0.25 (6/24) (In these cases n=1 has been considered as the most appropriate value). However, it should also be considered that an exposure design based on 6 hours daily exposure also includes 18 hours daily recovery, whereas there is no recovery during continuous 24 hours exposure. The above correction (with a factor of factor 6/24) may therefore underestimate the risk for continuous exposure.

Ad 4.

In case the inhalation route is involved one should also keep the principle of allometric scaling in mind when using inhalation volumes for animals and humans. This implies that standard respiratory volumes (in l/min/kg bw) for rats and humans differ by a factor of 4 (see also Section R.8.4.3 and APPENDIX R. 8-2, part 1). The physiological default values are given in Table R. 8-2. For certain situations deviations from these assumptions might be necessary, e.g. during 8 hours light activity at work the respiratory rate becomes higher than standard. This deviation is consistent with the assumption of a total breathing volume of 10 m³ for an 8-hour shift and light activity at work. These differences need to be corrected for, in order to obtain the correct starting point. It is to be noted that within one endpoint, these corrections result in a corrected starting point that is not the same for workers as for the general population.

Table R. 8-2 Default physiological parameters under the allometric scaling principle

(see also corresponding Table in Section R.7.12)

Species/ Physiological parameters	Rat	Human
Body weight	250 g	70 kg
Respiratory volume (standard; sRV)	0.2 l/min/rat = 0.8 l/min/kg bw	allometric scaling [a] → 0.2 l/min/kg bw
for relevant duration: 6 h exposure 8 h exposure 24 h exposure	 0.29 m^3/kg bw 0.38 m^3/kg bw 1.15 m^3/kg bw	 5 m^3/person 6.7 m^3/person 20 m^3/person
Respiratory volume light activity for worker (wRV) 8 h exposure		 10 m^3/person

[a] Difference between metabolic rate scaling and body weight scaling for rats and humans: 4 (see also Table R. 8-3)

How to derive a correct starting point

In Figure R. 8-2 it is illustrated how the modification of the starting point works out in case a N(L)OAEC from a 6h/d inhalatory rat study is available and the human exposure conditions are different. In Figure R. 8-3 it is illustrated how the modification of the starting point works out for one of the most common situations, i.e. an oral N(L)OAEL from a rat study (in mg/kg bw/day) that is to be used to assess inhalatory exposure of humans (in mg/m^3). Detailed guidance on modification of the starting point for this and other situations is given in APPENDIX R. 8-2, part 1-2, and in Section R.8.4.3.1.

Figure R. 8-2 Modification of the starting point:

Conversion of an inhalatory rat N(L)OAEC into a corrected inhalatory N(L)OAEC in case of differences between experimental and human exposure conditions

$$\text{corrected N(L)OAEC} = \text{inhalatory N(L)OAEC} * \frac{\text{exp.cond.}_{rat}}{\text{exp.cond.}_{human}}$$

$$= \text{inhalatory N(L)OAEC} * \frac{6\ \text{h/d}}{8\ \text{h/d}} * \frac{6.7\ \text{m}^3\text{(8h)}}{10\ \text{m}^3\text{(8h)}}$$

(for workers, in case of 8h exposure/d)

$$= \text{inhalatory N(L)OAEC} * \frac{6\ \text{h/d}}{24\ \text{h/d}}$$

(for general population, in case of 24h exposure/d)

Figure R. 8-3 Modification of the starting point:

Conversion of an oral rat N(L)OAEL into a corrected inhalatory N(L)OAEC to assess human inhalatory exposure.

For general population (in case of 24h exposure/d):

$$\text{correctedinhalatoryN(L)OAEC} = \text{oral N(L)OAEL} * \frac{1}{\text{sRV}_{rat}} * \frac{\text{ABS}_{\text{oral-rat}}}{\text{ABS}_{\text{inh-rat}}} * \frac{\text{ABS}_{\text{inh-rat}}}{\text{ABS}_{\text{inh-human}}}$$

$$= \text{oral N(L)OAEL} * \frac{1}{1.15\ \text{m}^3/\text{kg/d}} * \frac{\text{ABS}_{\text{oral-rat}}}{\text{ABS}_{\text{inh-human}}}$$

For workers (in case of 8h exposure/d):

$$\text{corrected inhalatory N(L)OAEC} = \text{oral N(L)OAEL} * \frac{1}{\text{sRV}_{rat}} * \frac{\text{ABS}_{\text{oral-rat}}}{\text{ABS}_{\text{inh-human}}} * \frac{\text{sRV}_{human}}{\text{wRV}}$$

$$= \text{oral N(L)OAEL} * \frac{1}{0.38\ \text{m}^3/\text{kg/d}} * \frac{\text{ABS}_{\text{oral-rat}}}{\text{ABS}_{\text{inh-human}}} * \frac{6.7\ \text{m}^3\text{(8h)}}{10\ \text{m}^3\ \text{(8h)}}$$

ABS: Absorption; sRV: standard Respiratory Volume; wRV: worker Respiratory Volume

APPENDIX R. 8-8 to APPENDIX R. 8-12 give further guidance on endpoint-specific modification of dose-descriptors for acute toxicity, irritation/corrosion, sensitisation, and reproductive toxicity.

After modification, where necessary, of the relevant dose descriptor(s) for the various threshold endpoints, the corrected starting points should be collected in a table (see Table R. 8-15 of APPENDIX R. 8-1), one per exposed population.

R.8.4.3 c) apply, when necessary, assessment factors to the correct starting point

c) Apply, when necessary, assessment factors to the correct starting point to obtain endpoint-specific DNEL(s) for the relevant exposure pattern (duration, frequency, route and exposed human population)

The next step in the calculation of a DNEL is to address uncertainties in the extrapolation of experimental data to the real human exposure situation, taking into account variability and uncertainty. These uncertainties concern, e.g., differences between animals and humans in anticipated sensitivity towards the toxicity of the substance. All these uncertainties/differences are individually addressed by so-called assessment factors (AFs), that together result in an overall AF that is applied to the corrected dose descriptor to account for all these uncertainties. Preferably, the value for each individual assessment factor is based on substance-specific information. However, although sound in principle, in practice the approach has limitations (data are often scarce, especially toxicodynamic data, and human data) and, therefore, default assessment factors most often need to be used. Each step in the process, including any choice for an assessment factor value, whether substance-specific or default should be explained as transparently as possible, with a qualitative narrative in the chemical safety report (CSR).

The following sections give guidance on the main issues to include in derivation of the overall AF applied in the general assessment procedure for threshold endpoints. The individual factors contributing to the overall AF are described separately in Section R.8.4.3.1. In Section R.8.4.3.3, guidance is given on how to combine these into an ‘overall assessment factor’.

At the same time, the descriptions point to many issues to be considered in qualitative discussions of the applicability and reliability of the effects assessment database.

Assessment factors

Assessment factors are numerical values. They are used to address the differences between the experimental data and the human situation, taking into account the uncertainties in the extrapolation procedure and in the available data set. In principle, all data on a specific substance need to be reviewed thoroughly in order to use, as far as possible, substance-specific information for the establishment of appropriate values for the various assessment factors. When substance-specific information is not available, data on analogues, which act with the same mode of action as the chemical under consideration, should be taken into account. However, when the available data do not allow the derivation of substance-specific or analogue-specific assessment factors, default assessment factors should be applied. Although very often necessary to rely upon, the default assessment factors represent a fall back position rather than the starting point.

Several publications exist on the use and/or quantification of assessment factors in human health risk assessment. For illustration, a short overview of defaults proposed in some of these publications is given in APPENDIX R. 8-3. For more background information and further reading, the reader is referred to the original publications.

Defaults typically proposed for human health risk assessment are point estimates. Additionally, default distributions have been proposed for assessment factors, acknowledging that lognormal distributions best describe variability and uncertainty in these factors. Some of these distributions are based on NOAEL-ratios derived from comprehensive toxicological databases. Some risk assessors, however, doubt the thoroughness and validity of such derived distributions.

APPENDIX R. 8-3 reflects this situation by illustrating the wide variation in approaches. It is obvious, that a harmonised consensus becomes difficult although, interestingly, there are quite some similarities in individual and overall assessment factors obtained via the different approaches (see the table in APPENDIX R. 8-3). This, combined with the desire to recommend a harmonised set of default factors to be used in risk assessments and thereby securing transparency, has led to the default factors recommended in this guidance document.

It is to be stressed that any choice for an assessment factor, whether substance-specific or default, should be explained as transparently as possible in the chemical safety report. The concept of substance-specific data to replace part or all of default assessment factors for inter- and intra-species differences (as described below in Section R.8.4.3.1) is elaborated upon in a recent guidance document from IPCS (WHO/IPCS, 2005), exemplified with case studies illustrating the types of data most valuable and how they can be used.

R.8.4.3.1 Assessment factors relating to the extrapolation procedure

Several aspects are involved in the extrapolation of experimental data to the human situation, *inter alia*, from the variability in the experimental data and from intra- and inter-species variation, the nature and severity of the effect, and the sensitivity of the human (sub-)population (REACH Annex I, Section 1.4.1). These aspects will be discussed under the following headings;

- interspecies differences;
- intraspecies differences;
- differences in duration of exposure;
- issues related to dose-response;
- quality of whole database.

Interspecies differences

Data from animal studies are the typical starting points for risk characterisations and thus differences in sensitivity between experimental animals and humans need to be addressed, with the default assumption that humans are more sensitive than experimental animals. Where human data are used as the starting point for the risk characterisation, no extrapolation and no assessment factor is necessary for interspecies differences in sensitivity.

Interspecies differences result from variation in the sensitivity of species due to differences in toxicokinetics and toxicodynamics. Some of the toxicokinetic differences can be explained by differences in body size (and related differences in basal metabolic rate). Information on interspecies differences may be gathered from the toxicological database of a substance, or from the use of PBPK modelling (see Section R.8.4.3.2).

If no substance-specific data are available, the standard procedure for threshold effects would be, as a default, to correct for differences in metabolic rate (allometric scaling) and to apply an additional factor of 2.5 for other interspecies differences, i.e. toxicokinetic differences not related to metabolic rate (small part) and toxicodynamic differences (larger part). In case substance-specific information shows specific susceptibility differences between species, which are not related to differences in basal metabolic rate, the additional factor of 2.5 for 'remaining differences' should be modified accordingly.

What is allometric scaling? Allometric scaling extrapolates doses according to an overall assumption that equitoxic doses (when expressed in mg/kg bw/day) scale with body weight to the power of 0.75. This results in different default allometric scaling factors for the different animal species when compared with humans (see Table R. 8-3). APPENDIX R. 8-2 describes in more detail how allometric scaling is done.

Table R. 8-3 Allometric scaling factors for different species as compared to humans[a]

Species	Body weight (kg)	AS factor[b]
Rat	0.250	4
Mouse	0.03	7
Hamster	0.11	5
Guinea pig	0.8	3
Rabbit	2	2.4
Monkey	4	2
Dog	18	1.4

a) assuming the human body weight is 70 kg

b) not applicable when setting an inhalation DNEL based on an inhalation animal study (see APPENDIX R. 8-2)

The factors are derived according to the formula:

$$\frac{bw_{human}/bw_{animal}}{(bw_{human}/bw_{animal})^{0.75}} = (bw_{human}/bw_{animal})^{0.25}$$

Allometric scaling is based on the assumption which was originally predicted mathematically and subsequently substantiated by empirical investigations (Schneider et al., 2004) that the effects of toxicological relevance are driven by the basal metabolic rate as it affects physiological processes such as cardiac output, blood flow and perfusion of liver and kidneys which, in turn, affect the elimination/clearance of most chemicals.

Allometric scaling is an empirical approach for interspecies extrapolation of a significant number of kinetic processes related to toxicity which is generally applicable to substances that are essentially renally excreted, but not to compounds that are highly extracted by the liver and excreted in the bile. It appears that species differences in biliary excretion and glucuronidation are independent of caloric demand (Walton et al., 2001).

Allometric scaling according to caloric demand would apply most appropriately to those substances for which the unmetabolised parent or a stable metabolite is the relevant toxic species and clearance is according to first-order processes. Conversely, the applicability of allometric scaling when toxicity is a consequence of exposure to a very reactive parent compound (or metabolite) that is not removed from the site of formation, is less well supported (USEPA, 1992).

It is to be noted that allometric scaling should not be applied if the effects are not dependent on metabolic rate or systemic absorption, e.g. in the case of local effects. In general, as long as route-to-route extrapolation is not needed, allometric scaling should also not be applied in cases where doses in experimental animal studies are expressed as concentrations (e.g., in mg/m^3 in air, ppm in diet, or mg/l in the drinking water) as these are assumed to be already scaled according to the allometric principle, since ventilation rate and food intake directly depend on the basal metabolic rate. However, once the concentration (e.g., ppm in diet) has been converted into a dose (e.g., mg/kg/day), an allometric scaling factor has to be used. Thus, it is the dose unit (original or transformed), and not the (experimental) route of application, that triggers the necessity for a species-specific factor for allometric scaling.

Allometric scaling is also not appropriate for acute lethal effects as these effects, which are accomplished by an immediate and intolerable level of damage to some critical homeostatic processes, may be independent of caloric demand and related physiological processes which affect toxicity (USEPA, 2006).

For systemic effects in case that allometric scaling is not applicable, assessment factors established on the basis of substance specific information should be well justified and used in a case-by case manner.

Special care should be taken when route-to-route extrapolation has been performed. See detailed guidance on this in APPENDIX R. 8-2, and a short summary thereof in Table R. 8-4. Part 1 of APPENDIX R. 8-2 illustrates that the allometric scaling is partly dependent on how the route-to-route extrapolation is being performed, and that there are two possible approaches. The preferred approach involves route-to-route extrapolation within one species as the first step, and interspecies extrapolation within the same exposure route as the second step. In this detailed guidance it is also indicated for the preferred approach whether allometric scaling should be included in step c of the DNEL derivation or whether this is already implicitly done in step b Table R. 8-4 illustrates the application of allometric scaling in step c when following the preferred approach (right-hand side of Example R. 8-1 and Example R. 8-2 in APPENDIX R. 8-2, part 1).

Table R. 8-4 When to apply allometric scaling (AS) factor in step c

Human exposure route (unit)	Experimental animal effect parameter (unit)	Apply AS factor?[a]
Oral (mg/kg bw/day)	Oral (mg/kg bw/day)	Yes, see APPENDIX R. 8-2, part 2 examples A1/B2
	Dermal (mg/kg bw/day)	Yes, see APPENDIX R. 8-2, part 2 example B6
	Inhalatory (mg/m^3)	Yes, see APPENDIX R. 8-2, part 2 example B4
Dermal (mg/kg bw/day)	Oral (mg/kg bw/day)	Yes, see APPENDIX R. 8-2, part 2 example B5
	Dermal (mg/kg bw/day)	Yes, see APPENDIX R. 8-2, part 2 examples A1/B2
	Inhalatory (mg/m^3)	Yes, see APPENDIX R. 8-2, part 2 example B4
Inhalation (mg/m^3)	Oral (mg/kg bw/day)	No, see APPENDIX R. 8-2, part 2 example B3
	Dermal (mg/kg bw/day)	No, see APPENDIX R. 8-2, part 2 example B3
	Inhalatory (mg/m^3)	No, see APPENDIX R. 8-2, part 2 examples A2/B1

[a] It should be noted that if using the approach outlined in the left-hand side of Example R. 8-1 and Example R. 8-2 of APPENDIX R. 8-2, part 1, the answers in the third column would be different (with yes answers except the three cases involving inhalation exposure, i.e., lines 3,6, and 9, where no AS factor in needed).

Local effects

For local effects, i.e. effects at the portal of entry (on the skin, the eye, the respiratory tract or the gastro-intestinal tract), different factors should be taken into account when assessing interspecies differences. First of all, it should be noted that, since local effects are independent of the basal metabolic rate, allometric scaling should not be applied (allometric scaling factor of 1). For the remaining uncertainties in kinetic (at a smaller extent) and in dynamic (at a larger extent) interspecies differences, consideration of the mechanism of toxicity is crucial, e.g. if the effect is a simple destruction of membranes due to the physicochemical properties (e.g. pH) of the chemical concerned as opposed to a mechanism involving local metabolism.

First of all, as for systemic effects, if there is data informing on these remaining uncertainties, this should be used to define a chemical-specific or analogue-specific factor. If there is no data informing on these remaining interspecies uncertainties, a default factor needs to be applied.

Both in terms of kinetics and dynamics, a distinction needs to be made between local effects on the skin, eye or gastro-intestinal (GI) tract and local effects on the respiratory tract.

For effects on the skin, eye or GI tract, where the mechanism of effect is direct chemical/pH reactivity, no further kinetic considerations apply. Furthermore, in terms of dynamics, one might assume that animals and humans will respond to the insult in the same way. In this case, the default factor for remaining uncertainties of 2.5 could be reduced to 1. In contrast, where tissue metabolism is a factor, the same kinetic and dynamic considerations (i.e. a chemical specific remaining uncertainties factor or the default factor of 2.5) should apply, as would be the case for e.g. kidney or liver damage arising from systemic metabolism. If tissue metabolism is involved, which could lead to the formation of different metabolites at different rates in different species, interspecies dynamic differences on how these metabolites interact with specific targets (which will determine the ultimate toxic response) cannot be completely ruled out.

For effects on the respiratory tract, whether the mechanism indicates that the effect seen is a simple destruction of membranes due to the physico-chemical properties (e.g. pH) of the chemical concerned or whether a local metabolic process is involved, further kinetic and dynamic considerations still apply.

Given that there could be significant quantitative differences in deposition, airflow patterns, clearance rates and protective mechanisms between humans and animals and when there is no data to inform on this uncertainty, it is prudent to assume that humans would be more sensitive than animals to effects on the respiratory tract. In such a situation, a chemical-specific remaining uncertainties factor or the default factor of 2.5 should be applied, as would be the case for systemic effects.

Intraspecies differences

Humans differ in sensitivity to toxic insult due to a multitude of biological factors such as genetic polymorphism affecting e.g. toxicokinetics/metabolism, age, gender, health status and nutritional status. These differences can be the result of genetic and/or environmental influences. This intraspecies variation is greater in humans than in the more inbred experimental animal population.

If the dose descriptor (e.g. N(L)OAEL, benchmark dose, etc.) has been derived from an animal study, animal intraspecies variation/differences has already to some extent been accounted for in that dose descriptor. Ideally therefore, the intraspecies factor should reflect the *additional* interspecies variability, i.e. the difference between variability in the human population and variability in the animal population. The variability within the experimental animals is however assumed to be small and in addition, difficult to quantify. Therefore the intraspecies assessment factors suggested below are not corrected for animal variation.

It is recognised that in order to always cover the most sensitive person exposed to any chemical would require a very large default assessment factor. That is of course not workable and it is usually assumed that a default assessment factor of 10 is sufficient to protect the larger part of the population, including e.g. children and the elderly. For threshold effects, this factor of 10 is the standard procedure, as a default, when assessing exposure to the general population. It is recognised that there are differences between children and adults in toxicokinetics (especially babies in their first months) and toxicodynamics (especially at different stages of development). These differences may render children more or less susceptible to the toxic effects of a substance. A higher intraspecies assessment factor for children (US-EPA, 1996, recommends from 10 up to 100 when assessing pesticides in relation to food safety) should be considered when the following two criteria are both fulfilled:

- There are indications, obtained from, for example, experiments in adult animals, epidemiological studies, *in vitro* experiments and/or SARs (structure activity relationships), of effects on organ systems and functions that are especially vulnerable under development and maturation in early life (in particular the nervous, reproductive, endocrine and immune systems and also the metabolic pathways), and
- There are deficiencies in the database on such effects in young animals.

This line of reasoning and criteria of course also applies to the unborn child, i.e., to the pregnant woman.

For workers, as standard procedure for threshold effects a default assessment factor of 5 is to be used, based on the fact that this sub population does not cover the very young, the very old, and the very ill.

Local effects

It is considered that information on intraspecies variation for local (concentration-dependent) effects is very scarce and no attempt has therefore been made to refine the default intraspecies factors already used for systemic effects.

For local effects, the assessment factors to be used for intraspecies differences are therefore the same as those proposed above for systemic effects.

It is to be noted that, as is the case for interspecies assessment factors, relevant substance-specific information on intraspecies variations should always be used to adjust or substitute the default factors (see e.g. WHO/IPCS, 2005).

Differences in duration of exposure

A factor allowing for differences in the experimental exposure duration and the duration of exposure for the population and scenario under consideration needs to be considered taking into account that a) in general the experimental NOAEL will decrease with increasing exposure times and b) other and more serious adverse effects may appear with increasing exposure times. Consequently, to end up with the most conservative DNEL for repeated dose toxicity, chronic exposure is the ‘worst case’. So, if an adequate chronic toxicity study is available, this is the preferred starting point and no assessment factor for duration extrapolation is needed. If only a sub-acute or sub-chronic toxicity study is available, the following default assessment factors are to be applied, as a standard procedure (Table R. 8-5):

Table R. 8-5 Assessment factors for duration extrapolation

Duration	Default assessment factor
sub-chronic to chronic	2
subacute to chronic	6
subacute to sub-chronic	3

'sub-chronic' usually refers to a 90 day study
'sub-acute' usually refers to a 28 day study
'chronic' usually refers to a 1.5 - 2 year study (for rodents)

These default assessment factors should be used for systemic effects and, in case of toxicity testing by inhalation, for local tissue damage in the respiratory tract (see e.g. experimental evidence reported in Kalberlah et al. (2002)).

However, substance-specific information is preferred and, if available, should be used to modify the default values, upwards or downwards.

- A *lower* factor (minimum 1) may for instance be used if there is specific evidence that increasing exposure duration does not increase the incidence or severity of adverse effects. This applies to most local dermal effects. It is also relevant for certain local effects in the respiratory tract for which there is no substantial difference in N(L)OAECs following acute and subacute exposure by inhalation (the effects can thus be considered concentration- rather than dose-dependent.
- A *higher* factor may for instance be used if there are indications for potential severe chronic effects, which cannot possibly be detected in a short term study.

  E.g. in cases where there is *in vitro* or QSAR data suggestive of such effects.

- A *higher* factor may also be used if there are indications for potential accumulation.
  E.g. relevant for lipophilic substances, in which case the database needs to contain information on the rate of elimination to further explore the accumulation potential. If accumulation is likely, the toxicity studies need to be of sufficient length to cover the accumulation period (e.g. the time to reach a steady-state concentration). If there is limited information on these aspects, the assessor needs to consider whether the database may be inadequate, and to which extent this lack of information should affect the assessment factor. In relation to inhalation of particles of very low solubility these will accumulate in the lung tissue over time which may result in a further increase in toxicity following long term exposure (Morrow, 1992).

Dose-response relationship

For the dose-response relationship, consideration should be given to the uncertainties in the dose descriptor (NOAEL, benchmark dose…) as the surrogate for the true no-adverse-effect-level (NAEL), as well as to the extrapolation of the LOAEL to the NAEL (in cases where only a LOAEL is available or where a LOAEL is considered a more appropriate starting point).

The size of an assessment factor should take into account the dose spacing in the experiment (in recent study designs generally spacing of 2-4 fold), the shape and slope of the dose-response curve, and the extent and severity of the effect seen at the LOAEL.

When the starting point for the DNEL calculation is a LOAEL, it is suggested to use an assessment factor between 3 (as minimum/majority of cases) and 10 (as maximum/exceptional cases).

However, the benchmark dose (BMD) approach is, when possible, preferred over the LOAEL-NAEL extrapolation (see Section R.8.2).

A BMD calculated as the lower confidence limit of the dose that produces a response of 5% (BMD5) has, on average, been proposed to be comparable to a NOAEL (WHO, 2000). If other BMD indicators are used, e.g. a BMD10, it should be considered on a case-by-case basis whether an additional dose-response assessment factor is needed.

When the starting point for the DNEL calculation is a NOAEL, the default assessment factor, as a standard procedure, is 1. However, a larger assessment factor may be applied in specific cases such as the following:

- a shallow dose-response curve giving uncertainty about the statistical derivation of the NOAEL
- exceptional cases of serious effects (e.g. severe irreversible effects, major malformations, foetal or offspring lethality) at dose levels slightly higher than the NOAEL (i.e. at the LOAEL) – this corresponds to a very steep dose-response curve
- poor quality of study from which the NOAEL is derived (e.g. few animals and inconsistent spacing between doses) also give uncertainty about the statistically derived NOAEL
- other concerns related to the identified NOAEL; e.g. for sensitisation (how certain is the NOAEL identified?) and carcinogenicity (is the mode of action for a presumed threshold carcinogen well understood?)

It is difficult to give exact guidance for the magnitude of an assessment factor for such specific situations. They should be determined on a case-by-case basis. The registrant may also choose to discuss these issues qualitatively.

In some cases neither a NOAEL nor a LOAEL can be identified, which e.g. might be the case when assessing acute toxicity based on a LD50 or LC50. In such situations, where dose levels with lethal outcome are used for DNEL calculation a much higher extra assessment factor should applied, see APPENDIX R. 8-8.

Quality of whole database

An assessment factor on the quality of the whole database should, if justified, be applied to compensate for the potential remaining uncertainties in the derived DNEL.

Firstly, the evaluation of the total toxicological database should include an assessment whether the available information as a whole meets the tonnage driven data requirements necessary to fulfil the REACH requirements, or whether there are data gaps (completeness of the database).

To account for deficiencies in the available data set and in identifying its magnitude, the assessor should consider both the data lacking and the data available. When there are deficiencies in the toxicity studies considered crucial to provide useful information for establishing the starting point, extra caution should be taken to address this scientific uncertainty in deriving the DNEL. Further, in order to account for data gaps and deficiencies in the available data set and in identifying its magnitude, the assessor should consider the nature of the effect occurring in particular organ systems, endpoints as well as at different life stages.

Special consideration should also be given to alternative data, e.g. in vitro data, (Q)SAR, read across or chemical categories. The use of alternative data is stimulated under REACH and preferred above performing additional animal studies, if considered justified. However, using these data in a quantitative way (if at all possible) might be associated with some additional uncertainty in the dose descriptor derived (see Chapter R.7 and general guidance on (Q)SARs and grouping of chemicals (Chapter R.6)). This should be accounted for.

Secondly, the hazard data should be assessed for the reliability and consistency across different studies and endpoints and taking into account the quality of the testing method, size and power of the study design, biological plausibility, dose-response relationships and statistical association (adequacy of the database).

One aspect of adequacy is reliability. Especially for the quality of the dose descriptor used (e.g. N(L)OAEL, BMDL) it is recommended to consider issues such as the statistical power of the study to detect changes from the control values, and the slope of the dose-response curve. These aspects are to be dealt with in the AF accounting for dose-response relationship (see above).

The other aspect of adequacy is consistency. A weight-of-evidence approach should be used in assessing the level of consistency of the total data base and especially of the starting point (e.g. N(L)OAEL, BMDL) to be brought forward to the risk characterisation. This approach requires a critical evaluation of the entire body of available data for consistency and biological plausibility. Potentially relevant studies should be judged for quality and studies of high quality given more weight than those of lower quality. When both epidemiological and experimental data are available, occurrence of similarity of effects between humans and animals should be considered. If the mechanism or mode of action is well characterized, this information is used in the interpretation of observed effects in either human or animal studies.

The default assessment factor to be applied for good/standard quality of the database, taking into account completeness, consistency and the standard information requirements, is 1. A larger database AF should where relevant be applied and justified on a case-by-case basis.

Endpoint-specific issues on AF

APPENDIX R. 8-8 to APPENDIX R. 8-12 give further guidance on endpoint-specific application of assessment factors for acute toxicity, irritation/corrosion, sensitisation, and reproductive toxicity.

R.8.4.3.2 Use of PBPK modelling for deriving assessment factors

As indicated in the introduction of Section R.8.4.3, the magnitude of the different assessment factors should be based preferably on substance-specific information. One way to come to such substance-specific assessment factors is via the use of physiologically based pharmacokinetic (PBPK) modelling, which can also be of great value in route-to-route extrapolation.

A PBPK model is an independent structural mathematical model, comprising the tissues and organs of the body with each perfused by, and connected via, the blood circulatory system. The principle application of PBPK models is in the prediction of the *target tissue dose* of the parent chemical or its reactive metabolite. Use of the target tissue dose of the toxic moiety of a chemical in risk assessment calculations provides a better basis of relating to the observed toxic effects than the external or administered exposure concentration of the parent chemical.

Prediction of target tissue dose following different exposure scenarios, routes, doses and species can help reduce the uncertainty associated with conventional extrapolation approaches. The mechanistic and biological plausibility of the models is the basis for associating greater confidence to such extrapolations.

The complexity and data demands of PBPK models are such, however, that in practice this will be a viable option for only a few substances. Moreover, the use and interpretation of these models in risk assessment and for decision making purposes requires specific expertise. Guidance on how the different extrapolation processes involved in risk assessment can be performed by using PBPK modelling can be found in APPENDIX R. 8-4.

R.8.4.3.3 Overall assessment factor and its application to the correct starting point

The overall assessment factor is obtained by simple multiplication of individual assessment factors discussed in the previous paragraphs. Care should be taken to avoid double counting several aspects when multiplying the individual factors.

Table R. 8-6 presents an overview of the individual default assessment factors, which should be used in the absence of relevant substance-specific information.

Table R. 8-6 Default assessment factors

Assessment factor – accounting for differences in:		Default value systemic effects	Default value local effects
Interspecies	- correction for differences in metabolic rate per body weight	AS[a, b]	–
	- remaining differences	2.5	1[f] 2.5[g]
Intraspecies	- worker	5	5
	- general population	10[c]	10[c]
Exposure duration	- subacute to sub-chronic	3	3 [h]
	- sub-chronic to chronic	2	2 [h]
	- subacute to chronic	6	6 [h]
Dose-response	- issues related to reliability of the dose-response, incl. LOAEL/NAEL extrapolation and severity of effect	1[d]	1[d]
Quality of whole database	- issues related to completeness and consistency of the available data	1[d]	1[d]
	- issues related to reliability of the alternative data	1[e]	1[e]

[a] AS = factor for allometric scaling (see Table R. 8-3)
[b] Caution should be taken when the starting point is an inhalation or diet study
[c] Not always covering for very young children; see text for deviations from default
[d] See text for deviations from default
[e] Special consideration needed on a case-by-case basis
[f] for effects on skin, eye and GI tract via simple destruction of membranes
[g] for effects on skin, eye and GI tract via local metabolism; for effects on respiratory tract

[h] for effects on respiratory tract.

In order to derive endpoint-specific DNEL(s) for the relevant exposure pattern (duration, frequency, route and exposed human population), the overall AF is to be applied directly to the corrected (where necessary) dose descriptor(s) in the following manner (exemplified with NOAEL as the dose descriptor):

$$\text{Endpoint-specific DNEL} = \frac{\text{NOAEL}_{corr}}{AF_1 \times AF_2 \times \ldots \times AF_n} = \frac{\text{NOAEL}_{corr}}{\text{Overall AF}}$$

All derived endpoint-specific DNEL(s) should be collected in a table (see Table R. 8-16 ofAPPENDIX R. 8-1), one per exposed population, per route.

R.8.5 Step 3-2: If possible, derive DMEL(s) for non-threshold endpoints

As indicated in Section R.8.1.3, DMELs are derived following a procedure consisting of the following steps:

a. selection of relevant dose-descriptor(s) for the endpoint concerned
b. modification, when necessary, of relevant dose descriptor(s) to the correct starting point
c. application, when necessary, of assessment factors/high to low dose risk extrapolation factor[10] to the correct starting point to obtain endpoint-specific DMEL(s) for the relevant exposure pattern (duration, frequency, route and exposed human population). The linearised approach and the 'EFSA approach' are described below, but other approaches can be used, if justified.

R.8.5.1 Deriving a DMEL for a non-threshold carcinogen, with adequate human cancer data

The process for deriving a DMEL for a non-threshod carcinogen when adequate human data are available is described in APPENDIX R.8-15, Section B.

R.8.5.2 Deriving a DMEL for a non-threshold carcinogen, with adequate animal cancer data

Basically two semi-quantitative risk assessment formats can be followed: one format, the 'Linearised' approach, essentially results in DMEL values representing exposure levels where the likelihood that effects (as asses by the lifetime cancer risk) are avoided is appropriately high and considered to be of very low concern. The other format, called 'Large Assessment Factor' approach, is formally similar to the overall assessment factor approach applied for threshold effects in deriving DNELs, and results in DMEL values representing exposure levels where the likelihood

[10] The term assessment factor is used because of it being a neutral term. However, these factors can in the DMEL-approach also be viewed as 'correction factors' and 'uncertainty factors'.

that effects (cancer) are avoided is appropriately high and of low concern from a public health point of view (EFSA, 2005).

Both formats are based on similar principal elements of risk extrapolation and risk evaluation, using as dose-descriptor either T25, BMD10 or BMDL10. A description of these dose descriptors, their derivation and differences between them are described in APPENDIX R. 8-6.

Because of different perceptions of the uncertainties of risk assessment and risk evaluation and different approaches for risk communication, there may be preference for one of these formats as well as for one of the dose descriptors.

In APPENDIX R. 8-7 these two approaches and their outcome are illustrated side-by-side for a hypothetical example.

R.8.5.2.1 The 'Linearised' approach

This approach of deriving a DMEL basically is driven by the assumption of a linear dose response relationship between tumour formation and exposure, and which is incorporated in the *high to low dose extrapolation* assessment factor in step c). This cautious approach may however be replaced by a different dose response relationship (either supralinear or sublinear) in this high to low dose extrapolation step if there is sufficient information in support of this.

a) Select the relevant dose-descriptor(s), i.e. T25 and BMD(L)10

The T25 should be used as the default dose-descriptor in relation to linear extrapolation. The linear approach is used when there is an absence of sufficient information on modes of action or when mode of action information indicates that the dose-response curve at low dose is or is expected to be linear. The BMD10 i.e. the Benchmark-dose representing a 10% response should be used in certain cases in addition to the T25 when data are adequate for modelling purposes. Thus, based on the available data a decision is made on what dose descriptor to use. Both descriptors, their derivation, a comparison and their use are described in APPENDIX R. 8-6.

b) Modify, when necessary, the relevant dose descriptor(s) to the correct starting point

In a few situations, the effects assessment is not directly comparable to the exposure assessment in terms of exposure route, units and/or dimensions. In these situations, it is necessary to convert the dose descriptor for the non-threshold effect into a correct starting point (i.e. corrected T25, corrected BMD10 and corrected BMDL10). This applies to the following situations:

1. If for a given human exposure route there is a dose descriptor for the same route in experimental animals but for that particular exposure route there is a difference in bioavailability between experimental animals and humans at the relevant level of exposure.
2. If for a given human exposure route there is not a dose descriptor for the same route (in experimental animals or humans).
3. Differences in human and experimental exposure conditions.
4. Differences in respiratory volumes between experimental animals (at rest) and humans (light activity).
5. Differences between occupational and lifetime conditions of exposure.

The corrections for situations 1-4 are performed in the same way as described in Section R.8.4.2 for the derivation of a DNEL.

Ad 5.
For non-threshold carcinogens, lifetime risks for consumers and for humans exposed indirectly via the environment is associated with daily exposure of 24 hours (7 days a week) for 75 years. This exposure duration is considered equivalent to the life-time exposure in experimental studies of 1.5 to 2 years, dependent on species and strain used (see APPENDIX R. 8-2, part 3).

For workers, however, the exposure time is 8 hours per day, 5 days per week, 48 weeks per year for 40 years. This implies that for workers, a correction factor should be applied to the dose descriptor based on animal life-time exposure data. As a default, a value of 2.8 (7/5 x 52/48 x 75/40) is proposed for this correction factor for oral studies, whereas a factor of 1.5 (6/8 x 5/5 x 52/48 x 75/40) is proposed for inhalation studies (which normally involves 6 hours exposure per day, 5 days per week).

These corrections pertain to systemic tumours and to some extent also to local tumours (i.e., step 3-5).

After modification, where necessary, of all relevant dose descriptors for the non-threshold endpoints, the corrected starting points should be collected in a table (see Table R. 8-15 of APPENDIX R. 8-1), one per exposed population.

c)derive from this correct starting point a DMEL for each relevant exposure pattern essentially by linear high to low dose extrapolation, and by application of assessment factors (when necessary)

In the calculation of a DMEL the essential step in quantitative terms is by far the 'high to low dose' extrapolation (that will be dealt with below). First, however, differences between experimental effect assessment data and the real human exposure situation, taking into account variability and uncertainty will be addressed.

Assessment factors

For following assessment factors are considered (like in the derivation of a DNEL):

- interspecies differences
- intraspecies differences
- differences in duration of exposure
- issues related to dose-response
- quality of whole database

These assessment factors are to be considered in the same way as described for DNEL derivation in Section R.8.4.3.1, unless otherwise detailed below.

It is to be noted that any relevant substance-specific information on any assessment factor should always be used to adjust or substitute the default factors applied here (see e.g. WHO/IPCS, 2005).

Interspecies differences

For systemic non-threshold effects, only an assessment factor for differences in metabolic rate (allometric scaling) is to be applied. However, this assessment factor is not needed for non-threshold effects;

- that are induced locally at the ports of entry, or
- when a respiratory study is used as starting point for deriving a DMEL in air for humans.

It should be noted that it is the dose unit (original or transformed), and not the (experimental) route of application, that triggers the necessity for a species-specific factor for allometric scaling. By this follows, for instance, that an AS factor is needed also in chronic studies once the concentration (e.g., ppm in food) is transformed into a body burden or dose (mg/kg/day), which is then used in the risk assessment.

The above implies that, in contrast to threshold effects, as a default there will be no assessment factor for remaining uncertainty (i.e. in the absence of substance-specific information) for both systemic and local non-threshold effects. The reason for this approach is that the linear model used for high to low dose extrapolation (see part on *high to low dose extrapolation* below), which is over about four orders of magnitude, is considered sufficiently conservative to also cover these differences in interspecies sensitivity.

Intraspecies differences

In contrast to threshold effects, no assessment factor is to be applied for this extrapolation step for non-threshold effects. The reason for this approach is that the linear model used for high to low dose extrapolation (see part on *high to low dose extrapolation* below), which is over about four orders of magnitude, is considered sufficiently conservative to also cover these differences in intraspecies sensitivity.

Differences in duration of exposure

In contrast to threshold effects, no assessment factor is to be applied for this extrapolation step for non-threshold effects. The reason for this is that a correction for durations of exposure (and/or observation) is already performed in deriving the dose descriptors before use in step a): see APPENDIX R. 8-6. It is noted, though, that if human exposure is not for lifetime or far from continuous during lifetime, correction of the DMEL may be needed according to the correction described at ad 5. of step b) of this approach.

Issues related to dose-response

The dose descriptor for non-threshold effects is, by definition, a dose level representing an observable and significant response. This is different from the situation encountered by threshold effects, where dose descriptors representing a true no-effect level are to be established and which inherently has some specific uncertainty.

Uncertainties related to the observable region of dose response curve for non-threshold effects are described in APPENDIX R. 8-6 for genotoxic carcinogens. The dose descriptors T25, BMD10, and BMDL10 have in increasing order incorporated uncertainty in their estimate. As indicated, preference is given to the T25, unless dose response curves have an exceptional supra- or sublinear shape. There is no separate assessment factor to account for this.

Another related issue concerning the dose response that is relevant specifically for non-threshold effects is high to low dose extrapolation. This is separately dealt with below.

Quality of whole database

An assessment factor on the quality of the whole database should, if justified, be applied to compensate for the potential remaining uncertainties in the derived DMEL.

Special consideration should be given to the situation that alternative data are used, e.g. use of (Q)SAR, read across or chemical categories or the use of subchronic studies for deriving some surrogate dose descriptor (see Section R.8.5.3). The situation of absence of substance-specific carcinogenicity data will quite frequently be encountered, also because the use of alternative data is stimulated under REACH and preferred above performing additional animal studies.

However, using these data in a semi-quantitative way (in cases where this is considered possible) might be associated with some additional uncertainty in the dose descriptor derived. Though this should be accounted for, there is no standard recipe for this, and expert judgement is critically demanded here.

The default assessment factor to be applied for good/standard quality of the database, taking into account completeness and consistency, is 1. A larger database AF should be justified on a case-by-case basis when data do not meet the mentioned qualification.

High to low dose risk extrapolation factor

The preceding steps (correction of the starting point, and application of assessment factors) have resulted in relevant (i.e. with regard to route and absorption) human equivalent lifetime daily doses HT25 ('Human T25'), and occasionally HBMD10 ('Human BMD10'), assumed to represent human daily exposures associated with tumour incidences of 25%, and 10%, respectively. This *high to low dose* extrapolation step is to arrive at the DMEL, i.e. an exposure level that is considered to represent a risk level where the likelihood that effects (cancer) are avoided is appropriately high and of very low concern, acknowledging the fact that for non-threshold carcinogens a dose level without any residual cancer risk cannot be identified.

This risk level of very low concern has to be decided on a policy level. Although there is no EU legislation setting the 'tolerable' risk level for carcinogens in the society, cancer risk levels have been set and used in different contexts (See APPENDIX R. 8-14 for various values previously applied within and outside the EU). Based on these experiences, cancer risk levels of 10^{-5} and 10^{-6} could be seen as indicative tolerable risks levels when setting DMELs for workers and the general population, respectively.

How is this DMEL derived?

How should one extrapolate from these generally high dose levels associated with high cancer risks [i.e. 25 or 10%] - to the low dose levels of human exposure associated with this risk level of very low concern?

The assessment of dose levels associated with these low risks, i.e. in any quantitative terms, is not possible, as they cannot be verified either experimentally or by epidemiological studies. Specific mathematical models have been developed for this purpose, i.e. for translating risks observed at high doses generally used in animal carcinogenicity tests (or observed in human occupational exposure situations) to risks associated with substantially lower exposure levels usually encountered in human situations. However, because of the small number of doses tested experimentally, i.e. usually only 2 or 3, almost all data sets equally well fit these various mathematical models, while their low dose risk predictions may vary by orders of magnitude because of their different theoretical assumptions.

This high to low dose response assessment today is generally performed in two steps; assessment of the dose response in the observable range for the tumour type under consideration to derive a dose descriptor, that subsequently serves as a starting point to extrapolation to lower dose levels.

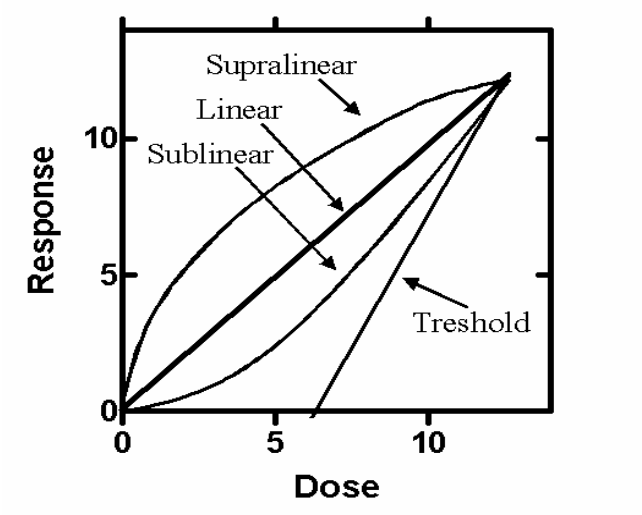
The default approach by several regulatory bodies is to extend a straight line, as a precautionary approach, to some prefixed exposure, representing the preferred risk level of very low concern or to actual human exposures for estimating associated risks[11].

Deviation from default linear approach

In general there is no need for an adjustment of this step in the effect assessment. If, however, the available data for the chosen tumour strongly suggest that linear extrapolation from the dose-descriptor value to some (very) low dose is not accurate and in fact indicate that the calculated risks are clearly under- or overestimating actual risks (i.e. the data indicate, respectively a supralinear or sublinear dose-response relationship for this part of the response curve; see Figure R. 8-4), some quantitative or qualitative judgement can be made or e.g. biologically based models or other non-linear models can be used, provided this is sufficiently supported by the available data. This should all be on a case-by-case basis, and clearly needs expert judgement.

11 Three different methods have been used or proposed by regulatory authorities in Europe and USA. The "Linearised Multistage Model (LMS) has been previously used extensively by US EPA (1986). The "LED10 method" has later been proposed by US EPA (1996) and the "T25 method" has been used in Europe (Sanner et al., 2001). The results obtained with these extrapolation methods are in most cases nearly indistinguishable (Sanner et al., 2001) and the differences are much smaller than generally found when different tumours or experiments are considered. It is recognised, though, that linear extrapolation may in some cases result in overestimation of risks at low exposures, but this may be acceptable from a precautionary principle standpoint. If the available data do indicate a deviation from linearity, these data should be taken into account resulting in a modification of the default approach. It should be noted that, in cases where high-quality epidemiology and animal carcinogenicity studies are available, a good agreement was found between hazard assessment based on epidemiology and hazard assessment based on animal studies using the T25 method (Sanner and Dybing, 2005).

Figure R. 8-4 Illustration of the supra- and sublinear dose response shapes referred to in the text, and the distinction with a thresholded dose response curve.



Derivation of DMEL

The DMEL is derived by applying the above assessment factors, and the high to low dose risk extrapolation factor to the correct starting point. Table R. 8-7 presents an overview of the individual default assessment factors that are to be used in the absence of relevant substance-specific information for this approach.

Table R. 8-7 Factors in the 'Linearised' approach to derive a DMEL

Assessment factor (AF_n)		Default value systemic tumours
Interspecies	- correction for differences in metabolic rate per body weight	AS [a,b]
	- remaining differences	1
Intraspecies	- general population	1
	- workers	1
Differences in duration of exposure	- lifetime exposure	1 [c]
Quality of database	- substance-specific data	1
	- non-testing data	> 1
	- other	case-by-case
High to low dose risk extrapolation factor (HtLF)		**Default value systemic tumours**
High-to-low-dose extrapolation	In case of e.g.:	For T25 ; for BMD10
	- 10^{-5} risk	25.000 ; 10.000
	- 10^{-6} risk	250.000 ; 100.000

[a] AS = factor for allometric scaling (see Table R. 8-3)

[b] Caution should be taken when the starting point is an inhalation or diet study

[c] Already accounted for in step b (Ad 5.).

There are two sets of High to low dose risk extrapolation factors (HtLF), depending on whether the starting point represents a 10 % or 25 % cancer risk. In addition, the HtLF depends on what cancer risk one wants to calculate, with a higher HtLF giving a lower risk of cancer (a ref to the HtLF to be added).

Derivation of the DMEL (based on a T25 as a starting point) for e.g. a risk for cancer of one per 100.000 exposed is arrived at in the following way:

$$\text{DMEL representing a } 10^{-5} \text{ risk} = \frac{T25_{corr}}{AF_1 \times \ldots \times HtLF} = \frac{T25_{corr}}{AS \times 25.000}$$

A DMEL for this risk level from a $BMD10_{corr}$ is derived in the same way, but the HtLF is then 10.000. Using the 'Linearised' approach, different DMEL values can be calculated, representing different lifetime cancer risks, e.g., a risk for cancer in 1 per 100.000 exposed individuals (10^{-5}).

Although there is no EU legislation setting the 'tolerable' risk level for carcinogens in the society, cancer risk levels have been set and used in different contexts (See APPENDIX R. 8-14 for various values previously applied within and outside the EU). Based on these experiences, cancer risk levels of 10^{-5} and 10^{-6} could be seen as indicative tolerable risks levels when setting DMELs for workers and the general population, respectively.

For workers, the requirements of the Carcinogens and Mutagens Directive (2004/37/EC) shall be complied with. This requires compliance with objectives to prevent exposure, substitution of dangerous chemicals by less dangerous chemicals and, where this is not technically possible, by minimisation of exposure.

All DMEL(s) per exposed population and per route should be collected in a table (see Table R. 8-16 of APPENDIX R. 8-1).

R.8.5.2.2 The 'Large Assessment Factor' approach ("EFSA" approach)

In the previous Section R.8.5.2.1, the 'linearised' approach was described, i.e. the derivation of a DMEL value that included a high to low dose extrapolation under step c) that by default is taken as linear. Another approach that might be used in order to characterise and evaluate carcinogenic risks was recently forwarded by Scientific Committee of the European Food Safety Authority (EFSA SC) when providing guidance for managing risks posed by contaminants in food (EFSA, 2005). Basically the same steps that apply to the 'linearised' approach apply here as well.

a) Select the relevant dose-descriptor(s), i.e. BMD(L)10 and T25

This procedure uses the BMDL10 as preferential dose descriptor, because the BMD approach is the preferred approach by EFSA SC, and this value is the lowest statistically significant increased incidence that can be measured in most studies, and would normally require little or no extrapolation outside the observed experimental data.

In cases where the dose-response data are inadequate for deriving an estimate of the BMD10 and BMDL10, EFSA SC recommends the use of the T25 as the reference point; as it can be easily applied and it is already in use in the European Union.

In case the BMDL10 is deviating more than one order of magnitude from the corresponding BMD10, the T25 should be used as dose descriptor. These descriptors, their derivation, a comparison and their use are described in APPENDIX R. 8-6.

b) Modify, when necessary, the relevant dose descriptor(s) to the correct starting point

The same modifications, when applicable, are to be applied here as described for step b) for deriving a DMEL via the 'linearised' approach.

After modification, where necessary, of all most relevant dose descriptors for the non-threshold endpoints, the corrected starting points should be collected in a table (see Table R. 8-15 of APPENDIX R. 8-1), one per exposed population.

c) Apply assessment factors to the correct starting point to obtain DMEL(s) for the relevant exposure pattern (route and exposed human population)

In this procedure the following assessment factors are taken into account to arrive at an exposure level viewed as "a low priority for risk management" (EFSA, 2005):

- interspecies differences & intraspecies differences
- the nature of the carcinogenic process (inter-individual human variability in cell cycle control and DNA repair)
- the reference point on the animal dose-response curve is not a NOAEL

Interspecies differences & Intraspecies differences

The usual default factor of 100 for non-genotoxic substances represents the product of two 10-fold factors, one to allow for possible interspecies differences, and one to allow for human variability (WHO, 1987 and 1994). These 10-fold factors allow for physiological and metabolic differences and these would also be relevant for substances which are both genotoxic and carcinogenic. These default factors of 10 could be reduced or increased when appropriate chemical specific data are available as described for instance by IPCS (WHO/IPCS, 2005 and IPCS website http://www.who.int/ipcs/en/).

The impact of polymorphisms of drug metabolism on cancer susceptibility has been widely investigated. Genetic polymorphism in a pathway of metabolism can lead to a more than 10-fold difference in the internal dose of the substance, but this is a rare situation and only occurs if it is a functional polymorphism in the major route of elimination (Dorne and Renwick, 2005). The overall conclusion drawn from a number of laboratory and epidemiology case-control studies is that genetic variation in xenobiotic-metabolising enzymes has in general a modest effect on the individual cancer risk associated with low-level environmental exposure (Hirvonen *et al.*, 1999; Taningher *et al.*, 1999; Pavanello and Clonfero, 2000). This is substantiated by a meta-analysis of cancer risk estimates from case-control studies, which showed odds ratios lower than 2 for variant genotype population groups (D'Errico *et al.*, 1999).

EFSA SC considers that the same physiological and metabolic differences apply also for substances that are both genotoxic and carcinogenic, consequently a difference between the reference point and human intakes of at least 100 would be sufficient to allow for these inter- and intraspecies differences.

It is to be noted that any relevant substance-specific or analogue-specific information on these assessment factors should be used to adjust or replace the default factors applied here.

The nature of the carcinogenic process (inter-individual human variability in cell cycle control and DNA repair)

The mode of action for substances that are both genotoxic and carcinogenic includes irreversible steps, such as the fixation of DNA lesions into permanent and inheritable mutations. The consequences of irreversible steps are amplified by clonal expansion of a single mutated cell, accumulation of genetic changes and progression of the mutated cells into cancer.

Genetic factors modulate the individual risk of cancer associated with environmental exposures (Shield and Harris, 2000). The probability of genetic alterations at critical targets following exposure to exogenous or endogenous genotoxic substances may be dependent on the efficiency of repair of DNA damage and cell cycle control. Candidate genes which may influence individual cancer risk by counteracting fixation of DNA-lesions into mutations include DNA repair genes, immune function genes, and genes controlling cell-cycle and apoptosis (Brennan, 2002).

Attention has focused in recent years on the possible association between DNA repair and cancer risk (Mohrenweiser and Jones, 1998; Hu *et al.*, 2002). Mutagen sensitivity varies little between identical twins compared to dizygotic twins and siblings, indicating a genetic basis in the individual susceptibility to DNA damage (Cloos *et al.*, 1999; Tedeschi *et al.*, 2004).

The majority of investigations on variations in DNA repair in humans involve a comparison between cancer patients with cancer free individuals. Such differences may be due to intrinsic differences in DNA repair within the human population but could also arise as a consequence of tumour development. As a conservative approach it is assumed that reported individual differences in DNA repair can occur within a cancer free population. Mohrenweiser (2004) recently reviewed studies that compared measures of DNA-repair capacity between cancer case subjects and healthy control subjects. The conclusion was that reductions of 20 to 35% in DNA-repair capacity were associated with elevations in cancer risk in the majority of studies, usually with odds ratios in the range of 3 to 6.

Data from molecular epidemiology studies also are consistent with an association between some variant alleles of DNA repair genes and increased risk of lung, breast and prostate cancers (Goode *et al.*, 2002).

Most genes preventing genome instability and the genes regulating cell proliferation are polymorphic in the human population, with common variants with low penetrance which may affect cancer susceptibility. In particular, polymorphisms of the gene TP53 (producing the protein p53), *p21* and cyclin D1 have been associated with increased susceptibility/poor prognosis of breast cancer (Powell *et al.*, 2002), cancer of the urinary bladder (Wang *et al.*, 2002) and lung cancer (Qiuling *et al.*, 2003), all with odds ratios of 2 to 3.

After *in vitro* treatment of blood cells from healthy subjects with genotoxic agents a variation in response in a range of around an order of magnitude has been reported (Gu *et al.*, 1999), but the contributions of individual variant alleles of DNA repair genes is modest, less than two-fold although the impact of low penetrance polymorphisms may theoretically be barely detectable (Mohrenweiser *et al.*, 2003). In addition, nutritional and lifestyle factors may be superimposed on the genetic diversity, modulating the level of DNA damage and contributing to the individual DNA repair phenotype (Collins, 2003; Palli *et al.*, 2003; Wei *et al.*, 2003). EFSA noted that most of these studies have been performed *in vitro*, and that their relevance to *in vivo* situations remains uncertain.

The EFSA SC considered that a default factor of 10 would cover this area of uncertainty. However, it is to be noted that any relevant substance-specific or analogue-specific information on this assessment factor should be used to adjust or replace the default factor applied here.

The reference point on the animal dose-response curve is not a NOAEL

As discussed above EFSA SC considered that a BMDL10 would be the most appropriate reference point. This reference point on the animal dose-response curve relates to a small but measurable response and so cannot be regarded as a surrogate for a threshold in the case of a substance that is both genotoxic and carcinogenic.

In addition the dose effect relationship below the reference point, and the dose level below which cancer incidence is not increased are unknown, representing additional uncertainties.

The EFSA SC considered that a default factor of 10 would cover this area of uncertainty. However, it is to be noted that any relevant substance-specific or analogue-specific information on this assessment factor should be used to adjust or replace the default factor applied here.

Since the T25 is less conservative than the BMDL10, this would also need to be taken into account when establishing an exposure level viewed as "a low priority for risk management": an additional factor of 2.5 has to be applied in these cases.

These different AF can be tabled as follows (Table R. 8-8):

Table R. 8-8 Default assessment factors in the 'Large assessment factor' approach

Assessment factor	Default value systemic tumours
Interspecies	10
Intraspecies	10 5[a]
Nature of the carcinogenic process	10
The point of comparison ('BMD/T25 is not a NOAEL')	10

[a] Not addressed by EFSA; a value of 5 is suggested for workers

Derivation of the DMEL for the general population via this procedure is arrived at from a $BMDL10_{corr}$ in the following way:

$$DMEL = \frac{BMDL10_{corr}}{AF_1 \times AF_2 \times \ldots \times AF_n} = \frac{BMDL10_{corr}}{10.000}$$

A DMEL based on a T25 is derived in the same way, but the overall AF is then 10,000 x 2.5. It is noted that the value of a BMDL10, and $BMDL10_{corr}$ may differ substantially from a BMD10, i.e. may be substantially lower, dependent on the confidence of the available data. For this reason it is unclear, how the two DMEL values derived by the 'Linearised' and 'Large assessment factor' approaches, respectively, will compare.

All DMEL(s) per exposed population, and per route should be collected in a table (see Table R. 8-16 of APPENDIX R. 8-1).

R.8.5.2.3 Alternatives to the conventional extrapolation procedures

PBPK modelling is an alternative to the approach described above in Section R.8.5.2.1 and R.8.5.2.2, and can be used in the derivation of DMELs, if such information of adequate quality is available. For details on how this can be done see Section R.8.4.3.2 and APPENDIX R. 8-4.

R.8.5.3 Deriving a DMEL for a non-threshold carcinogen/mutagen, without adequate cancer data

In some cases a risk characterisation as outlined in the previous sections is not possible, i.e.: in the absence of carcinogenicity data:

1. The testing strategy opens the possibility for regarding substances which are positive for mutagenicity in somatic cells as potential genotoxic carcinogens without carrying out a carcinogenicity bioassay.
2. In some cases, groups of substances are classified as a single entry. While there is often good reason for including a group of substances as a single entry, the actual lifetime cancer risk of the individual substances within the group may vary. In these cases a risk characterisation may be difficult. or when available carcinogenicity data are inadequate or difficult to interpret:
3. Available data may be of insufficient quality.
4. The available data may be from well-performed assays with transgenic animals that are known for their potential higher sensitivity towards carcinogenic agents.

In such a case the following possibilities may be explored to derive a DMEL:

- read across
- use of subchronic studies
- the threshold of toxicological concern (TTC) concept

Read across

Carcinogenicity data may be available for substances that have clearly close structural similarity to the substance under investigation. These data may be evaluated for deriving some surrogate dose descriptor value for the substance under investigation by so-called read across. Basically, there are two options here: i.e. carcinogenicity data are available

- only for one other member (analogue approach),
- for several other members (category approach) [12]

The most ideal situation would be the last option of several members with available data. In this case the surrogate dose descriptor value for the substance of interest may be obtained via interpolation from dose descriptors (for carcinogenicity) for members with data or, if this appears not feasible, by taking some reasonable worst case estimate, e.g. the lower 95-percentile value of a distribution of the available dose descriptors values.

If data are available for only one member, a clear reasoning should be given on whether and how from this data a surrogate dose descriptor value for the substance of interest can be derived. The way the category is formed and structured is critical here.

[12] Note that elsewhere guidance is provided how to establish a chemical category for a specific chemical structure, and how information from other members in this category should be applied for both C&L, and dose response evaluation (see Section R.6.2).

In case substances with close structural similarity are available, but for none of these carcinogenicity data are available, one could potentially explore the possible use of the other options described below, although there is not yet any agreed guidance for these approaches.

Use of subchronic studies

In this approach, an estimate of the DMEL may alternatively be obtained by use of the available data from animal toxicity studies: i.e. by identifying the minimal toxic dose in sub-chronic studies (if available, as some surrogate value for the dose descriptor) and by applying a large assessment factor; see for further guidance Gold *et al*. (2003). It is stressed that expert judgement is definitively needed here, and further development and stakeholder agreement on the use of this approach is needed before guidance can be given.

Threshold of Toxicological Concern

The threshold of toxicological concern (TTC) is a principle which refers to the possibility of establishing a human exposure threshold value, below which there is no appreciable risk to human health for each of 3 structural classes by considering extensive databases of toxicity data (by the oral route) generated in the past. Thus, these TTC values would only be applicable to the oral route of exposure.

Currently, the TTC concept is used for regulatory purposes in the risk assessment of flavourings and food additive substances[13]. A more extended description of the TTC concept is presented in Appendix R.7-1).

Clearly, any approaches involving the TTC concept or use of subchronic studies need further (permanent) development and stakeholder agreement before any guidance can be given.

Transgenic animals

In case only transgenic animal data are available one may, of course, apply the above described approaches, when considered appropriate. Alternatively, the dose descriptor from this transgenic assay can be used to derive some surrogate DMEL value (by applying the same approach as with a dose descriptor from non-transgenic animals). Deviation from this latter approach, e.g. because of anticipated higher sensitivity of the animals used, should be clearly documented and justified.

R.8.6 Step 3-3: Follow a more qualitative approach when no dose descriptor is available for an endpoint

When no reliable dose descriptor can be set for a given endpoint, a more qualitative approach has to be chosen. This may apply for acute toxicity, irritation/corrosion, sensitisation, and mutagenicity/carcinogenicity. A brief description of when this may be relevant for the different end-points are given below, whereas more detailed guidance on qualitative risk characterisation can be found in APPENDIX R. 8-8 to APPENDIX R. 8-11 (guidance related to some specific endpoints)

[13] The TTC concept forms the scientific basis of the US Food and Drug Administration (FDA) 'Threshold of Regulation' for food contact materials (Federal Register, 1993), and has also been adopted by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) in its evaluation of flavouring substances (WHO, 1996). The FDA came in their early analysis (for dietary exposure) to the conclusion that the established 'Threshold of Regulation' also applied to carcinogens (Federal Register, 1993).

and in Section E.3.4. It should be noted that DNEL/DMELs may still need to be set for other endpoints or routes of exposure. This is further explained in Section R.8.7 and Part E.

A qualitative risk characterisation of **acute toxicity** is recommended for substances showing a very high acute toxicity (i.e., labelled T or T+ in the present C&L system) through any route of exposure in particular when based on the available data on acute toxicity no sufficiently robust DNEL can be set. This may, for example, apply when the lethality data has been obtained by a different route of exposure than the relevant route of human exposure (see also APPENDIX R. 8-8 and Section E.3.4). Very strict risk management measures (RMMs) and operational conditions (OCs) will apply for these substances (e.g., closed systems, etc) in order to ensure control. Basically, the RMMs/OCs should ensure that peak concentrations exceeding the long-term DNEL will not occur (see also APPENDIX R. 8-8 and Section E.3.4).

When no DNEL can be derived for **irritation/corrosion**, a more qualitative approach to assessing and controlling these risks is appropriate. This can be the case when only the following types of data are available: pH, in vitro data on skin and eye irritation/corrosion, in vivo data with no information on dose-response, or QSAR/read-across. From these types of data, only qualitative information (yes/no) and sometimes the potency of the irritation and corrosion can be obtained (APPENDIX R. 8-9 and Section E.3.4).

In case of **skin sensitisation**, the first step should always be a qualitative approach to assessing and controlling the risks and setting a DNEL (if possible) could be used to judge the remaining/residual likelihood of risks. If available, information on the potency can be used in qualitative risk characterisation and for recommendation of appropriate RMMs and OCs. Proposals for potency categorization of sensitizers based on the LLNA, Guinea pig maximisation test (GPMT) and the Buehler test are available (see APPENDIX R. 8-10). In case of the LLNA, the potency is categorised based on the EC3 value, and in case of the GPMT and the Buehler test is based on the percentage of positive animals in relation to the concentration tested (see APPENDIX R. 8-11 and Section E.3.4).

Since there are currently no available methods to determine the thresholds and to establish DNEL for **respiratory hypersensitivity**, only qualitative risk assessment for this endpoint can be performed. There is evidence from both human and animal studies which indicate that effective sensitisation of the respiratory tract can result from dermal contact with a chemical respiratory allergen. Thus there is growing view, that the effective prevention of respiratory sensitisation requires appropriate protection of both the respiratory tract and skin. The generic advice is that appropriate strategies to minimise the risk of sensitisation to chemical allergens will require consideration of providing protection of all relevant routes of exposure (see APPENDIX R. 8-11 and Section E.3.4).

When no DNEL/DMEL can be set for carcinogens and/or mutagens, e.g., either because of lack of in vivo data or lack of quantitative dose descriptors in vivo studies, a more qualitative assessment needs to be done (see Section E.3.4).

The outcome of this step should be a qualitative description of the severity and potency of the endpoint, including the classification and labelling.

R.8.7 Step 4: Select the leading health effect(s)

Step 4: Select the leading health effect(s) and the corresponding DNEL/DMEL and/or other qualitative/semi-quantitative description

R.8.7.1 Selection of the critical DN(M)EL

Following the derivation of endpoint-specific DN(M)EL(s) as per steps 3-1 to 3-2, as appropriate, the leading health effect and the corresponding critical DN(M)EL should be selected for the relevant exposure patterns (i.e. combinations of population/route/exposure). These critical DN(M)EL(s), used for the risk characterisation, should be the lowest DN(M)EL obtained for each exposure pattern. This can be deduced from Table R. 8-16 of APPENDIX R. 8-1, in which all available endpoint-specific DNEL(s)/DMEL(s) (whether on the basis of substance-specific data, read across from one or more structural analogues or other alternative data) have been collected. Note that assessments covering simultaneous exposure via several routes of exposure will require that relevant DNELs for each exposure route are set (See further details in Section E.3.5).

Thus, in principle step 4 should be easy and straightforward when endpoint-specific DN(M)EL values for the different identified hazards have been derived. The lowest DNEL or DMEL value can then be selected. Note that, depending on the exposure patterns, there may be more than one critical DN(M)EL (see Section R.8.7.2). For most substances and exposure scenarios, the critical DN(M)ELs will be representing repeated exposure (i.e. a DNELlong-term; see Section R.8.1.2) rather than representing exposure for a short period of time (i.e. a DNELacute; see Section R.8.1.2).

In case, however, peak exposure cannot be ruled out upfront for a certain Exposure Scenario, the assessment should **also** include specific assessment of 'acute' exposure, e.g., 15 minutes peak exposures. The human peak values should then specifically be compared with the relevant $DNEL_{acute}$, ensuring that the peak exposure complies with the $DNEL_{acute}$ for the exposure route under consideration (normally inhalation), even if this $DNEL_{acute}$ is less critical than the $DNEL_{long-term}$. Systemic effects after acute oral and dermal exposure should in a first tier be assessed using the corresponding long-tem DNELs. However, on a case by case basis, acute DNELs for single oral and/or dermal exposure may need to be set. Table R. 8-9 illustrates the DN(M)ELs that normally may need to be derived.

For exposure to dust, it should be considered whether a derived DNEL for inhalation may have to be lowered. The general dust limits of 10 mg/m^3 for the inhalable airborne fraction and 3 mg/m^3 for the respirable airborne fraction used in setting Occupational Exposure Limits in many countries should be considered in combination with nature of the dust. The following should be considered:

- For non-soluble inert dusts if the derived DNEL for inhalation is above these dust limits, the general dust limits should apply for exposure scenarios with exposure to dust
- For significantly soluble dusts, if the derived DNEL for inhalation is above, the general dust limit might apply. Where it is not to be used, the rationale for any deviation from the general dust limits should be justified.

Note that DNELs derived based on substance specific data can never be adjusted upwards based on the general dust limits and that the dust limits can not be used as a surrogate DNEL when there is no data to set a substance-specific DNEL.

Table R. 8-9 DN(M)ELs that normally may need to be derived

Exposure pattern	DNEL/DMEL (appropriate unit)	
	Workers	General population [3]
Acute – inhalation, systemic effects[1]		
Acute – dermal, local effects[2] (e.g., for irritation, corrosion, & sensitisation,, if DNELs can be set)		
Acute – inhalation, local effects[2] (e.g., for irritation, corrosion, & sensitisation,, if DNELs can be set)		
Long-term – dermal, systemic effects[1]		
Long-term – inhalation, systemic effects[1]		
Long-term – oral, systemic effects[1]	Not relevant	
Long-term – dermal, local effects[2]		
Long-term – inhalation, local effects[2]		

[1] Units are mg/m^3 for inhalation, and mg/kg bw for oral and dermal exposure

[2] Units are mg/m^3 for inhalation, and mg/cm^2 or ppm for dermal exposure

[3] General population includes consumers and humans via the environment. In rare cases it may also be relevant to derive a DNEL for specific subpolulations, such as children.

R.8.7.2 Endpoints for which no DNEL/DMEL can be derived

However, step 4 is not so straightforward when, for some of the endpoint, no endpoint-specific DN(M)EL values could be derived. This may be the case for

a. (a) non-threshold mutagens with no cancer data,
b. (b) non-threshold carcinogens with no suitable quantitative data,
c. (c) respiratory sensitizers,
d. (d) skin sensitizers,
e. (e) skin and eye irritants and/or
f. (f) other groups of substances determined on a case-by-case basis for which the experimental data do not allow the establishment of a threshold.

For these endpoints, the qualitative description of potency from Step 3.3 (see Section R.8.6) should be taken forward to the Risk Characterisation. Part E outlines how to conduct the Risk Characterisation when endpoints with no DNELs as well a critical DNEL (or DMEL) are available for a given exposure pattern.

R.8.7.3 Using DN(M)EL for human exposure patterns

The DNELs or DMELs are then used for the appropriate human exposure patterns, as illustrated in the tables below (Table R. 8-10 to Table R. 8-13). The tables are, thus, valid also for DMELs, even though DMELs are not, for simplicity, mentioned in the tables.

The lowest $DN(M)EL_{long\text{-}term}$ is usually the starting point for the RC, and it is normally set based on data from repeated dose toxicity studies. Such studies include the 28 and 90 days repeated dose toxicity studies, reproductive toxicity studies (including developmental toxicity studies), and chronic/carcinogenicity studies.

For **systemic, long-term effects**, DNELs are generally needed for worker dermal and inhalation exposure. Thus, in a first tier these two worker DNELs (Table R. 8-10) usually need to be set and used to assess the occupational exposure.

Table R. 8-10 Worker long-term DN(M)ELs generally needed

DNEL	Duration and routes of exposure to humans corresponding to the DNEL
Worker-DNEL long-term dermal	Repeated worker dermal exposure for a day or more (this exposure is generally modelled as a dermal daily deposition expressed in mg substance/cm^2 skin)
Worker-DNEL long-term inhalation	Repeated worker inhalation exposure for a day or more (exposure is modelled or measured as a daily air concentration in mg substance/m^3)

Additionally, long-term DNELs may need to be set for the general population if the substance is present in consumer–available products or is released to the environment and present as an environmental contaminant (Table R. 8-11). The DNELs are used in the following scenarios.

Table R. 8-11 Long-term DNELs that may be needed for the general population

DNEL	Duration and routes of exposure to humans corresponding to the DNEL
General Population-DNEL long-term oral	Repeated exposure oral of the general population (consumers, humans via the environment, expressed as mg/kg/day)
General Population-DNEL long-term dermal	Repeated dermal exposure of the general population (consumers)(generally modelled as a dermal daily exposure expressed in mg substance/cm^2 skin)
General Population-DNEL long-term inhalation	Repeated inhalation exposure of the general population (consumers or humans via the environment)(modelled or measured as a daily air concentration in mg substance/m^3)

For some toxic substances, for which there may be peak exposures, a $DNEL_{acute}$ need to be set and assessed in relation to the human peak exposure levels (see further APPENDIX R. 8-8).

The $DNEL_{acute}$ is set based on studies involving exposure for very short periods (for inhalation normally 15 minutes' peak exposures). However, even though an effect may occur later, after the 'acute' exposure episode, the effect is highly relevant and should be considered as basis for the NOAEL/DNEL. The acute toxicity studies are the most relevant studies. Also human data, such as

from case studies, need to be assessed. In the absence of experimental data, the acute DNEL can by default be set as 1-5 times the long-term DNEL.

Thus, as a rule of thumb, if actual peak exposure levels to toxic substances exceed the long-term DNEL by several-fold, a detailed acute risk assessment clearly has to be made.

For **systemic, acute** effects, two DNELs (Table R. 8-12) are normally relevant to compare with peak exposures, although occupational inhalation exposure is most often the most important one.

Table R. 8-12 Acute DNELs that may be needed

DNEL	Duration and routes of exposure to humans corresponding to the DNEL
Worker-DNEL acute inhalation	Worker inhalation peak exposure
General Population - DNEL acute inhalation	Occasional inhalation exposure (minutes-hours) of the general population (consumers, humans via the environment)

Rarely, and on a case-by-case basis, the other routes may need to be assessed (potentially constituting three different DNELs). That includes a systemic DNEL acute dermal for workers and the general population, and a systemic DNEL acute oral for the general population, in both cases representing single exposure. However, in a first tier, these exposures should be compared against the corresponding long-term DNELs.

For both acute and long-term **local effects**; four of the above scenarios may apply for substances causing irritation, corrosion and/or sensitisation, assuming that the data allow setting a DNEL (Table R. 8-13). Acute dermal and inhalation, and long-term dermal and inhalation DNELs (oral is not relevant) may be needed for workers and the general population. The corresponding human exposure situations are as above, with a comparison of external exposure levels with external DNELs.

Table R. 8-13 Acute and long-term DNELs that may be set for local effects, e.g., irritation, corrosion, sensitisation.

Note that corresponding DNELs may need to be set for the general population.

DNEL	Duration and routes of exposure to humans corresponding to the DNEL
worker-DNEL acute dermal local	Worker dermal single exposure
worker-DNEL acute inhalation local	Worker inhalation peak exposure
worker-DNEL long-term dermal local	Repeated worker dermal exposure
worker-DNEL long-term inhalation local	Repeated worker inhalation exposure

Part E outlines how to conduct the risk characterisation.

REFERENCES

Brennan, P. (2002) Gene-environment interactions and aethiology of cancer: what does it mean and how can we measure it? Carcinogenesis 23, 381-387.

Collins, A.R.; Harrington, V.; Drew, J.; Melvin, R. (2003) Nutritional modulation of DNA repair in a human intervention study. Carcinogenesis 24, 511-515.

Cloos, J.; Nierwenhuis, E.J.; Boomsma, D.I.; Kuik, D.J.;van der Sterre, M.L.; Arwert, F.; Snow, G.B.; Braakhuis, B.J. (1999) Inherited susceptibility to bleomycin-induced chromatid breaks in cultured peripheral blood lymphocytes. J. Natl. Cancer Instit. 91, 1125-1130.

D'Errico, A.; Malats, N.; Vineis, P.; Boffetta, P. (1999) Review of studies of selected polymorphisms and cancer. In: W. Ryder (ed.) Metabolic polymorphisms and susceptibility to cancer. IARC Scientific Publication n. 148, Lyon, International Agency for Research on Cancer 1999, pp. 323-393.

Dorne, J.L.C.M., and Renwick, A.G. (2005) Refinement of uncertainty/safety factors in risk assessment by the incorporation of data on toxicokinetic variability in humans. Toxicol. Sci. 86, 20-26.

EFSA (2005) Opinion of the Scientific Committee on a request from EFSA related to A Harmonized Approach for Risk Assessment of Substances Which are both Genotoxic and Carcinogenic. The EFSA Journal 282,1-30.

Federal Register (1993). Food additives; threshold of regulation for substances used in food-contact articles. Federal Register 58 (195), 52719-52727

Gold LS, Gaylor DW, and Slone TH (2003) Comparison of cancer risk estimates based on a variety of risk assessment methodologies. Regulatory Toxicology and Pharmacology 37, 45-53.

Goode, E.L.; Ulrich, C.L.; Potter, J.D. (2002) Polymorphisms in DNA repair genes and association with cancer risk. Cancer Epidem Biomar 11, 1513-1530.

Gu, J.; Bondy, M. L.; Sigurdson, A., Spitz; M. R.; Hsu, T. C.; Wu, X. (1999) Three Measures of Mutagen Sensitivity in a Cancer-Free Population. Cancer Genet. Cytogenet. 110, 65-69.

Hirvonen, A. (1999) Polymorphisms of xenobiotic-metabolizing enzymes and susceptibility to cancer. Environ. Health Perspect. 107 (suppl. 1), 37-47.

Hu, J.J.; Mohrenweiser, H.W.; Bell, D.A.; Leadon, S.A.; Miller, S. (2002) Symposium overview: genetic polymorphisms in DNA repair and cancer risk. Toxicol. Appl. Pharmacol. 185, 64-73

IGHRC (2006). Guidelines on route-to-route extrapolation of toxicity data when assessing health risks of chemicals. The Interdepartmental Group on Health Risks from Chemicals, http://www.silsoe.cranfield.ac.uk /ieh/ighrc/ighrc.html

Kalberlah, Först, Schneider (2002). Time extrapolation and interspecies extrapolation for locally acting substances in case of limited toxicological data. Ann Occup Hyg 46(2), 175-185.

Mohrenweiser, H.W. and Jones, I.M. (1998) Variation in DNA repair is a factor in cancer susceptibility: a paradigm for the premise and perils of individual and population risk estimation? Mutat. Res. 400, 15-24.

Mohrenweiser, H.W.; Wilson III, D.M.; Jones, I.M. (2003) Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive

genetic variation in human DNA repair genes. Mutat. Res. 526, 93-125.

Mohrenweiser, H.W. (2004) Genetic variation and exposure related risk estimation: will toxicology enter a new era? DNA repair and cancer as a paradigm. Toxicol. Pathol. 32, 136-145.

Morrow PE (1992). Contemporary issues in toxicology – dust overloading of the lungs: update and appraisal. Toxicol. and Appl. Pharmacol. 113, 1-12.

Palli, D.; Masala, G.; Vineis, P.; Garte, S.; Saieva, C.; Krogh, V.; Panico, S.; Tumino, R.;

Munnia, A.; Riboli, E.; Peluso, M. (2003) Biomarkers of dietary intake of micronutrients modulate DNA adduct levels in healthy adults. Carcinogenesis 24, 739-746.

Pavanello S. and Clonfero, E. (2000) Biological indicators of genotoxic risk and metabolic polymorphisms. Mutat. Res. 463(3):285-308.

Powell, B.L.; Iris, L.; van Staveren, I.L.; Rooske, P.; Grieu, F.; Berns, E.M.J.J.; Iacopetta,

B. (2002) Associations between common polymorphisms in TP53 and p21WAF1/Cip1 and phenotypic features of breast cancer. Carcinogenesis 23, 311-315.

Qiuling, S.; Yuxin, Z.; Suhua, Z.; Cheng, X.; Shuguang, L.; Fengsheng, H. (2003) Cyclin D1 gene polymorphism and susceptibility to lung cancer in a Chinese population. Carcinogenesis 24, 1499-1503.

Sanner T et al. (2001). A simple method for quantitative risk assessment of non-threshold carcinogens based on the dose descriptor T25. Pharmacol. Toxicol. 88, 331.

Sanner T and Dybing E (2005) Comparison of carcinogen hazard characterisation based on animal studies and epidemiology. Basic Clinical Pharmacol. Toxicol. 96, 66-70.

Schneider K, Oltmanns J and Hassauer M (2004). Allometric principles for interspecies extrapolation in toxicological risk assessment – empirical investigations. Regulatory Tox and Pharm, 39(3), 334-347.

Shield, P.G. and Harris, C.C. (2000) Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J. Clinical Oncol. 18, 2309-2315.

Taningher, M.; Malacarne, D.; Izzotti, A.; Ugolini, D; Parodi, S. (1999) Drug metabolism polymorphisms as modulators of cancer susceptibility. Mutat. Res. 436(3); 27-61.

Tedeschi, B.; Cicchetti, R.; Argentin, G.; Caporossi, D.; Pittaluga, M.; Parisi, P.; Vernole, P. (2004) Aphidicolin and bleomycin induced chromosome damage as biomarker of mutagen sensitivity: a twin study. Mutat. Res. 546(1-2); 55-64.

U.S. EPA (1986) Guidelines for carcinogen risk assessment. Federal Register 51(185):33992–34003.

US EPA (1992). Draft Report: A cross-species scaling factor for carcinogen risk assessment based on equivalence of mg/kg3/4/day; Notice. Federal Register 57(109):24152-2424173.

US EPA (1996). U.S. Public Law 104-170. 104th Congress. Food Quality Protection Act (FQPA) of 1996.

US EPA (2005) Guidelines for Carcinogen Risk Assessment , Risk Assessment Forum U.S. Environmental Protection Agency (EPA/630/P-03/001F), Washington DC, March 2005.

US EPA (2006). External Review Draft: Harmonisation in interspecies extrapolation – use of BW3/4 as default method in derivation of the oral RfD. Risk Assessment Forum Technical Panel. EPA/630/R-06/001.

Walton K, Dorne JL, and Renwick AG (2001). Uncertainty factors for chemical risk assessment: interspecies differences in glucuronidation. Food Chem Toxicol 39, 1175-1190.

Wang, L.; Habuchi, T.; Takahashi, T.; Mitsumori, K.; Kamoto, T.; Kakehi, Y.; Kakinuma, H.; Sato, K.; Nakamura, A.; Ogawa, O.; Kato, T. (2002) Cyclin D1 gene polymorphism is associated with an increased risk of urinary bladder cancer. Carcinogenesis 23, 257-264.

Wei, Q.; Shen H.; Wang, L-E.; Duphorne, C.M.; Pillow, P.C.; Guo, Z.; Qiao, Y.; Spitz, M.R. (2003) Association between low dietary folate intake and suboptimal cellular DNA repair capacity. Cancer Epidemiology, Biomarkers and Prevention 12, 963-969.

WHO (1987) Principles for the assessment of risk to human health from exposure to chemicals. Environmental Health Criteria, 70, World Health Organization, Geneva.

http://www.inchem.org/documents/ehc/ehc/ehc70.htm

WHO (1994) International Programme on Chemical Safety (IPCS): Assessing human health risks of chemicals: Derivation of Guidance values for health-based exposure limits. Environmental Health Criteria, 170, 73 pp., World Health Organisation, Geneva.

http://www.inchem.org/documents/ehc/ehc/ehc170.htm

WHO: World Health Organisation (1996). Toxicological evaluations of certain food additives and contaminations in food. WHO Food Additives Series 35.

WHO (2000) Air Quality Guidelines (second edition).

WHO/IPCS (2005). Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration–response assessment. IPCS Harmonization Project Document No. 2, http://whqlibdoc.who.int/publications/2005/9241546786_eng.pdf

APPENDIX R. 8-1 Summary tables for dose-response information and DNELs/DMELs.

Table R. 8-14 Available dose-descriptor(s) per endpoint as a result of hazard assessment or, if no dose descriptor can be identified, other information on potency

Endpoint	Quantitative dose descriptor[1] (appropriate unit) or other information on potency		Associated relevant effect[2]	Remarks on study[3]
	Local effect[4]	Systemic effect[5]		
Acute toxicity [6] oral dermal inhalation				
Irritation/Corrosivity skin eye resp. tract		NA [7] NA NA		
Sensitisation skin resp. tract		NA NA		
Repeated dose toxicity [8] sub-acute/ sub-chronic/ chronic oral dermal inhalation				
Mutagenicity in vitro in vivo				
Carcinogenicity oral dermal inhalation				
Reproductive toxicity [8] fertility impairment oral dermal inhalation developmental tox oral dermal inhalation				

[1] NOAEL (NOAEC), LOAEL , T25, BMD(L)10 or any other dose descriptor; indicate whether this concerns a no or lowest observed effect level etc.

[2] In this column the relevant effect for which the dose descriptor is determined is provided

[3] This column is for indicating whether data were available, whether the substance is classified for this endpoint, for shortly describing specifics of the study (e.g. 28-d gavage rat, 5 d/wk or 2-gen diet rat, 7 d/wk), and for indicating (additional) uncertainty in available data

[4] Units are mg/m^3 for inhalation, and mg/cm^2 or ppm for dermal exposure

[5] Units are mg/m^3 for inhalation, and mg/kg bw/day for oral and dermal exposure

[6] In general, sublethal toxicity is a more rational starting point for acute toxicity than mortality data; information on acute toxicity may also be derived from e.g. repeated dose toxicity studies or reproductive toxicity studies

[7] Not applicable

[8] These repeated exposure studies may also show relevant acute effects of the test substance; these should be accounted for under the endpoint acute toxicity

Table R. 8-15 Corrected dose descriptor(s)

Per endpoint for the relevant exposure pattern for workers/consumers/human via the environment[1]

Endpoint	Most relevant quantitative dose descriptor[2] (appropriate unit)		Corrected dose descriptor (appropriate unit)	
	Local[3]	Systemic[4]	Local[3]	Systemic[4]
Acute toxicity - oral - dermal - inhalation				
Irritation/Corrosivity - skin - eye - resp. tract		NA [5] NA NA		
Sensitisation - skin - resp. tract		NA NA		
Repeated dose toxicity sub-acute/ sub-chronic/ chronic - oral - dermal - inhalation				
Mutagenicity - in vitro - in vivo				
Carcinogenicity - oral - dermal - inhalation				
Reproductive toxicity fertility impairment - oral - dermal - inhalation developmental tox - oral - dermal - inhalation				

[1] Select the relevant population

[2] NOAEL (NOAEC), LOAEL , T25, BMD10 etc or any other dose descriptor; indicate whether this concerns a no or lowest observed effect level etc.

[3] Units are mg/m^3 for inhalation, and mg/cm^2 or ppm for dermal exposure

[4] Units are mg/m^3 for inhalation, and mg/kg bw/day for oral and dermal exposure

[5] Not applicable

Table R. 8-16 Endpoint-specific DNEL(s)/DMEL(s)

For the relevant exposure pattern for workers/consumers/human via the environment[1]

Endpoint	Corrected dose descriptor (appropriate unit)		Overall AF applied	Endpoint-specific DNEL/DMEL (appropriate unit)	
	Local[2]	Systemic[3]		Local[2]	Systemic[3]
Acute toxicity - oral - dermal - inhalation					
Irritation/Corrosivity - skin - eye - resp. tract		NA[4] NA NA			NA[4] NA NA
Sensitisation - skin - resp. tract		NA NA			NA NA
Repeated dose toxicity sub-acute/ sub-chronic/ chronic - oral - dermal - inhalation					
Mutagenicity - in vitro - in vivo					
Carcinogenicity - oral - dermal - inhalation					
Reproductive toxicity fertility impairment - oral - dermal - inhalation developmental tox - oral - dermal - inhalation					

[1] Select the relevant population

[2] Units are mg/m^3 for inhalation, and mg/cm^2 pr ppm for dermal exposure

[3] Units are mg/m^3 for inhalation, and mg/kg bw/day for oral and dermal exposure

APPENDIX R. 8-2 Bioavailability, route-to-route extrapolation and allometric scaling

Examples to illustrate how to obtain consistent results

When transferring study results from animals to humans, care has to be taken to use a meaningful physiological parameter as a reference value for scaling issues. This Appendix illustrates in part 1 some issues related to this. Part 2 gives specific guidance on how to deal with differences in bioavailability (in practice as determined by differences in absorption) and how to conduct route-to-route extrapolation in the situations identified in Section R.8.4.2. In Part 3 default parameters for lifetime cancer studies are summarised which are relevant to derive consistent dose descriptors. More information on these issues can be found in the references in the end of this appendix.

Part 1- Scaling issues

Where inhalative data are concerned, air concentrations for animal and human exposure are generally compared directly. Using this approach implies standardisation of inhalative data with reference to the respiratory rates. Since respiratory rates depend directly on caloric demand this means, that inhalative study results are (implicitly) extrapolated to humans on the basis of metabolic rate scaling (also termed allometric scaling).

Oral data usually are expressed in dose per kg bodyweight. Comparing oral data directly would mean, to use body weight as a reference for scaling purposes. If, however, allometric scaling shall be used for standardisation, it has to be taken into account that metabolic rate does not correlate directly with body weight but with the body weight modified by the exponent 0.75 (metabolic rate ≈ body weight$^{0.75}$). On that background data from different species expressed as dose per kg bodyweight need to be adjusted to caloric demand before they can be compared based on metabolic rate. According to the different average bodyweights of the animal species, when comparing oral and dermal data with humans, specific allometric scaling factors are needed for each species (see Table R. 8-4 in Section R.8.4.3.1).

If oral data are used to evaluate inhalative exposure situations and the oral data are scaled on the basis of body weight, risk assessors need to be aware of the aspect outlined above. Usually respiratory rates for animals and humans are used for dose adjustment. For consistent results care has to be taken that the respiratory rates used in combination with the respective bodyweights match the allometric equation. In addition a special situation occurs for workers. Compared to a standardised situation with basal caloric demand, workers usually are in a status of elevated activity with higher respiratory rates. This has to be compensated for as well.

The following examples shall outline the procedure. The physiological values used in these examples are taken from Table R. 8-2 in Section R.8.4.2.

In Example R. 8-1 and Example R. 8-2 on the next pages, oral data from the rat are used to decide on a corresponding air concentration for humans. For simplicity 100% absorption for the oral and the inhalative route for animals and humans is assumed. The air concentration is calculated in two different ways.

On the right side of the examples, which is the preferred pathway and the pathway illustrated in Table R. 8-4 in Section R.8.4.3.1, the oral dose for the rat is converted to the corresponding air concentration using a standard breathing volume for the rat (1.15 m^3/kg for 24 hours exposure of general public, 0.38 m^3/kg for 8 hours exposure of workers, see Table R. 8-2 in Section R.8.4.2). For workers the resulting air concentration needs to be additionally corrected for the difference between basal caloric demand and caloric demand under light activity.

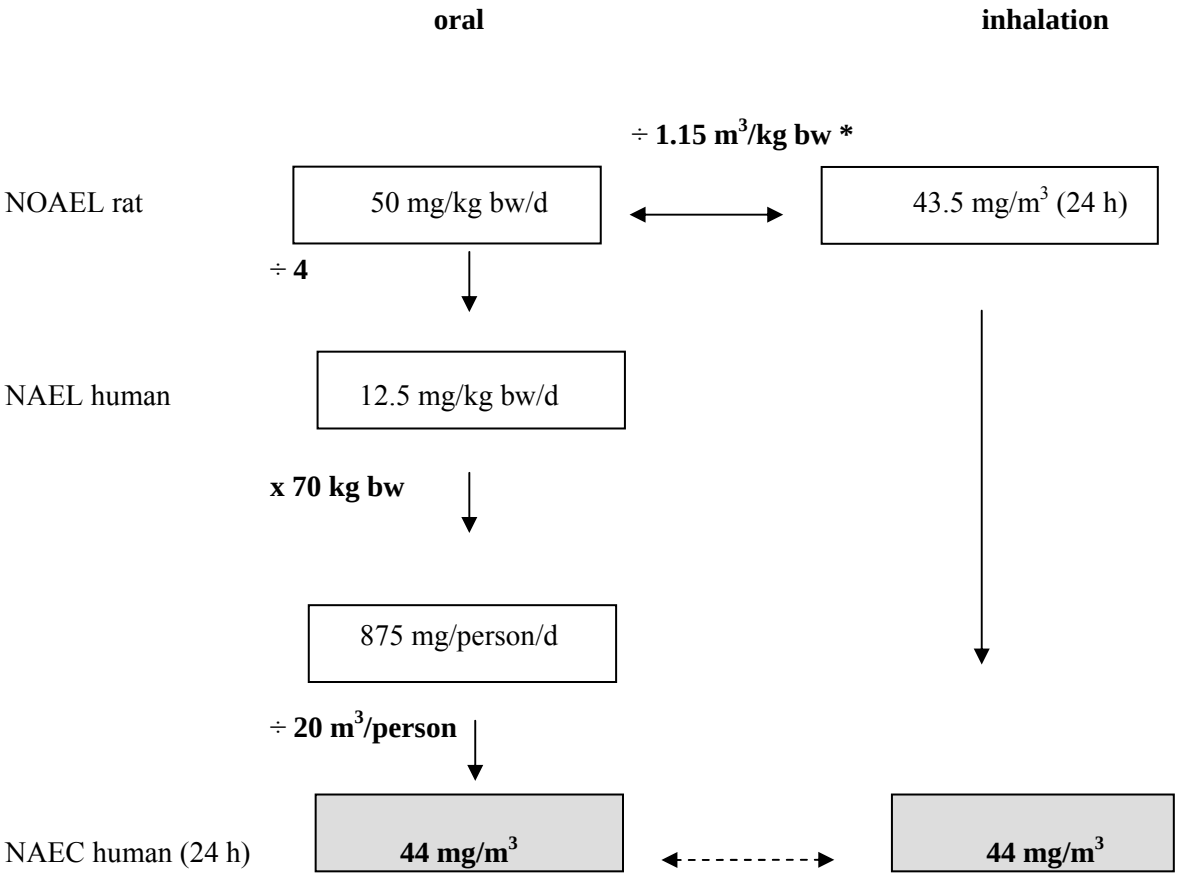
This correction factor derives from the inhalative volumes in 8 hours under the respective conditions (6.7 m^3 for base level, 10 m^3 for light activity).

On the left side of the examples, which is not the preferred way and which therefore is not illustrated in Table R. 8-4 in Section R.8.4.3.1, the oral NOAEL for the rat in a first step is transferred to humans with a factor of 4 for allometric scaling. With help of a standard human body weight (70 kg) and a default human breathing volume referring to the specific conditions of the respective population (20 m^3 for general public in 24 h hours and basal caloric demand, 10 m^3 for workers in 8h and light activity), this dose is then translated into an air concentration.

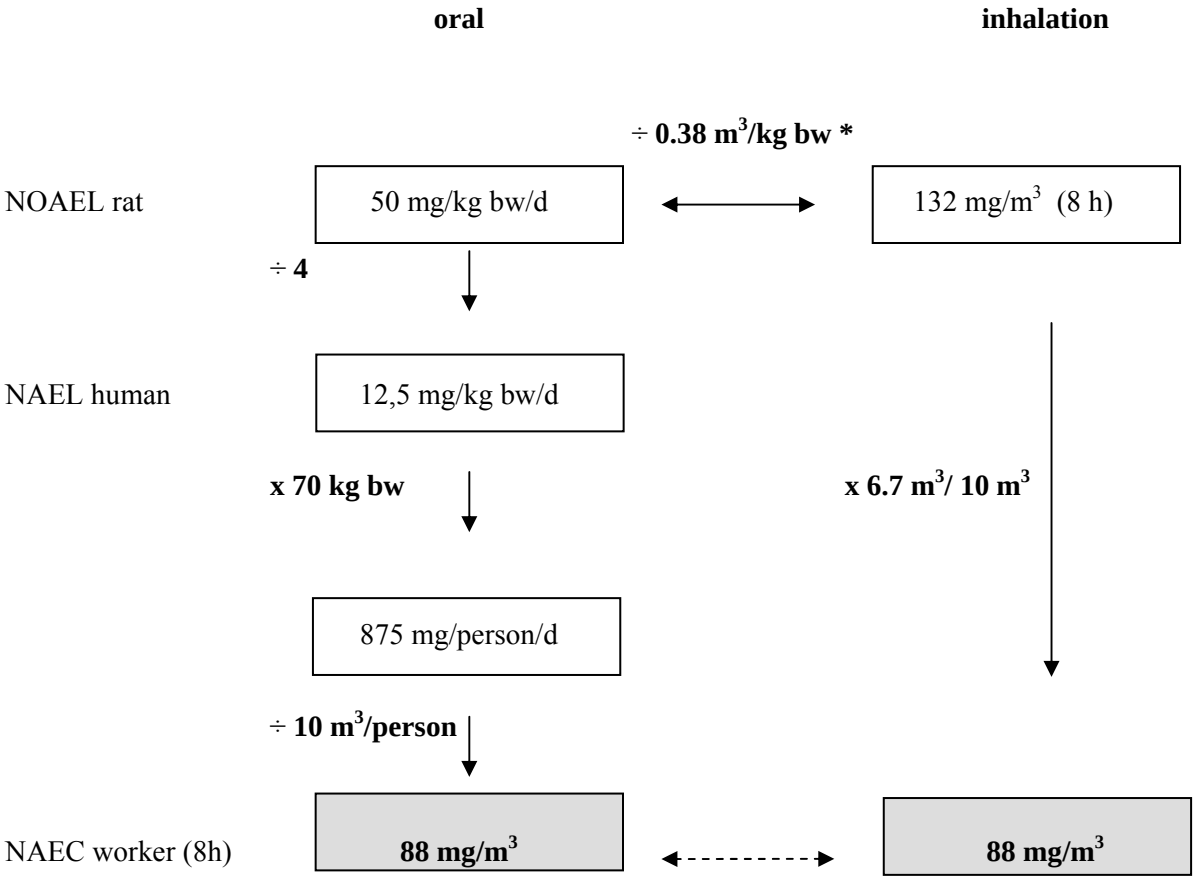
As can be seen from the results, the two different ways of calculation lead to the same results.

Examples (assuming 100 % absorption for both routes in both species)

Example R. 8-1 General public



Example R. 8-2 Workers



* See *Table R. 8-2* in Section R.8.4.2 for explanation of this factor

Part 2 - Guidance on modification of starting point

This part specifically outlines the procedure taking into account bioavailability issues. The examples especially concern the conversion of the N(L)OAEL/C into an adequate starting point for DNEL derivation, but also apply to other dose descriptors for e.g. non-threshold effects. *Please, note that the examples below only illustrates extrapolations conducted according to the procedure outlined in the right-hand side of examples I and II above, i.e., when the route-to-route extrapolation is performed within one species as the first step.* The examples presented also indicate whether allometric scaling should be included in step c of the DNEL/DMEL derivation or whether this is already implicitly done at this point (step b) (see Section R.8.4 for explanation of step b and c). In most cases substance-specific information relating to differences in bioavailability will not be available. Section R.8.4.2 suggests default factors for some of these situations.

It is to be noted that in all cases where the starting point is an inhalatory N(L)OAEC (examples A2, B1 and B4), it must be considered whether an additional concentration–time correction is needed when the experimental exposure conditions (e.g. 6 h/d) do not equal the human exposure conditions (e.g. for workers 8 h/d) (see Section R.8.4.2).

A. If for a given human exposure route there is an effect parameter for the same route (in experimental animals or humans) and for that particular exposure route there is no difference in absorption between experimental animals and humans at the relevant level of exposure, then in principle no modification of starting point is necessary [Example A. 1].

However, if the exposure route is via inhalation, then for workers a correction is necessary for the differences in respiratory rates under standard conditions and under conditions of light activity [Example A. 2].

Example A. 1 Oral[#] exposure; oral absorption rat = oral absorption human

For step b: no modification necessary
For step c: as to interspecies differences, apply factor for allometric scaling (4 for rat).

[#]similar situation for the dermal route

Example A. 2 Inhalatory exposure; inhalation absorption rat = inhalation absorption human

For step b: **I**. for workers, the inhalatory N(L)OAEC rat needs to be corrected for the difference between respiratory rates under standard conditions and under conditions of light activity (sRV_{human} versus wRV; see Table R. 8-2 in Section R.8.4.2):

For workers:

$$\text{corrected N(L)OAEC} = \text{inhalatory N(L)OAEC} \times \frac{sRV_{human}}{wRV}$$

$$= \text{inhalatory N(L)OAEC} \times \frac{6{,}7\ m^3}{10\ m^3}$$

II. consider when the inhalatory N(L)OAEC rat needs to be corrected for differences in the experimental and human exposure conditions (all populations).

For step c: as to interspecies differences, do not apply factor for allometric scaling.

B. Modification of starting point is necessary

- If for a given human exposure route there is an effect parameter for the same route (in experimental animals or humans) but for that particular exposure route there is a difference in absorption between experimental animals and humans at the relevant level of exposure [Example B. 1 and Example B. 2].
- If for a given human exposure route there is not an effect parameter for the same route (in experimental animals or humans) [Example B3 to
- Example B. 6].

Example B. 1 Inhalatory exposure; inhalation absorption rat ≠ inhalation absorption human

For step b: **I**. correct inhalatory N(L)OAEC rat (in mg/m^3) for differences in inhalation absorption between rats and humans. Additionally, for workers a correction is needed for the difference between respiratory rates under standard conditions and under conditions of light activity (sRV_{human} versus wRV; see Table R. 8-2 in Section R.8.4.2).

$$\text{corrected N(L)OAEC} = \text{inhalatory N(L)OAEC} \times \frac{ABS_{inh\text{-}rat}}{ABS_{inh\text{-}human}}$$

For workers:

$$\text{corrected N(L)OAEC} = \text{inhalatory N(L)OAEC} \times \frac{ABS_{inh\text{-}rat}}{ABS_{inh\text{-}human}} \times \frac{sRV_{human}}{wRV}$$

II. consider when the inhalatory N(L)OAEC rat needs to be corrected for differences in the experimental and human exposure conditions.

For step c: as to interspecies differences, do not apply factor for allometric scaling.

Example B. 2 Oral[#] exposure; oral absorption rat ≠ oral absorption human

For step b: correct oral N(L)OAEL rat (in mg/kg bw/day) for differences in oral absorption between rats and humans as follows:

$$\text{corrected N(L)OAEL} = \text{oral N(L)OAEL} \times \frac{ABS_{oral\text{-}rat}}{ABS_{oral\text{-}human}}$$

For step c: as to interspecies differences, apply factor for allometric scaling (4 for rat).

[#] similar situation for the dermal route

Example B. 3 Inhalatory exposure; oral[#] N(L)OAEL rat

For step b: convert oral N(L)OAEL rat (in mg/kg bw/day) into inhalatory N(L)OAEC rat (in mg/m^3) by using a default respiratory volume for the rat corresponding to the daily duration of human exposure (sRV_{rat}; see Table R. 8-2 in Section R.8.4.2), followed by a correction for differences in absorption between routes (if the case), and a correction for differences in inhalation absorption between rats and humans (if the case). For workers an additional correction is needed for the difference between respiratory rates under standard conditions and under conditions of light activity (sRV_{human} versus wRV; see Table R. 8-2 in Section R.8.4.2).

$$\text{corrected inhalatory N(L)OAEC} = \text{oral N(L)OAEL} \times \frac{1}{sRV_{rat}} \times \frac{ABS_{oral\text{-}rat}}{ABS_{inh\text{-}rat}} \times \frac{ABS_{inh\text{-}rat}}{ABS_{inh\text{-}human}}$$

$$= \text{oral N(L)OAEL} \times \frac{1}{sRV_{rat}} \times \frac{ABS_{oral\text{-}rat}}{ABS_{inh\text{-}human}}$$

For workers:

$$\text{corrected inhalatory N(L)OAEC} = \text{oral N(L)OAEL} \times \frac{1}{sRV_{rat}} \times \frac{ABS_{oral\text{-}rat}}{ABS_{inh\text{-}human}} \times \frac{sRV_{human}}{wRV}$$

For step c: as to interspecies differences, do not apply factor for allometric scaling.

[#] similar situation with dermal N(L)OAEL

Example B. 4 Oral[#] exposure; inhalatory N(L)OAEC rat

For step b: **I**. convert inhalatory N(L)OAEC rat (in mg/m^3) into oral N(L)OAEL rat (in mg/kg bw/day) by using a default respiratory volume for the rat corresponding to the daily duration of human exposure (sRV_{rat}; see Table R. 8-2 in Section R.8.4.2), followed by a correction for differences in absorption between routes (if the case), and a correction for differences in oral absorption between rats and humans (if the case).

$$\text{corrected oral N(L)OAEL} = \text{inhalatory N(L)OAEC} \times sRV_{rat} \times \frac{ABS_{inh\text{-}rat}}{ABS_{oral\text{-}rat}} \times \frac{ABS_{oral\text{-}rat}}{ABS_{oral\text{-}human}}$$

$$= \text{inhalatory N(L)OAEC} \times sRV_{rat} \times \frac{ABS_{inh\text{-}rat}}{ABS_{oral\text{-}human}}$$

II. Consider when the inhalatory N(L)OAEC rat needs to be corrected for differences in the experimental and human exposure conditions.

For step c: as to interspecies differences, apply factor for allometric scaling (4 for rat).

similar situation with dermal exposure

Example B. 5 Dermal exposure; oral N(L)OAEL rat

For step b: convert oral N(L)OAEL rat (in mg/kg bw/day) into dermal N(L)OAEL rat (in mg/kg bw/day) by correcting for differences in absorption between routes (if the case) as well as for differences in dermal absorption between rats and humans (if the case):

$$\text{corrected dermal N(L)OAEL} = \text{oral N(L)OAEL} \times \frac{ABS_{oral\text{-}rat}}{ABS_{derm\text{-}rat}} \times \frac{ABS_{derm\text{-}rat}}{ABS_{derm\text{-}human}}$$

$$= \text{oral N(L)OAEL} \times \frac{ABS_{oral\text{-}rat}}{ABS_{derm\text{-}human}}$$

For step c: as to interspecies differences, apply factor for allometric scaling (4 for rat).

Example B. 6 Oral exposure; dermal N(L)OAEL rat

For step b: convert dermal N(L)OAEL rat (in mg/kg bw/day) into oral N(L)OAEL rat (in mg/kg bw/day) by correcting for differences in absorption between routes (if the case) as well as for differences in oral absorption between rats and humans (if the case):

$$\text{corrected oral N(L)OAEL} = \text{dermal N(L)OAEL} \times \frac{ABS_{derm\text{-}rat}}{ABS_{oral\text{-}rat}} \times \frac{ABS_{oral\text{-}rat}}{ABS_{oral\text{-}human}}$$

$$= \text{dermal N(L)OAEL} \times \frac{ABS_{derm\text{-}rat}}{ABS_{oral\text{-}human}}$$

For step c: as to interspecies differences, apply factor for allometric scaling (4 for rat).

Part 3 - Dose calculations in lifetime studies

Table R. 8-17 Default values for dose calculations i.e. standard lifespan, body weights, food and water intake and inhalation volume (based on Gold et al., 1984 and Paulussen et al., 1998)

Experimental animal	Sex	Standard lifespan[a] (years)	Body weight[c] (kg)	Food per day[b] (g)	Water per day[b] (ml)	Inhalation volume (l/ hr)
Mouse	Male	1.5 - 2	0.03	3.6 (120)	5 (167)	2.5
	Female	1.5 - 2	0.025	3.25 (130)	5 (200)	2.2
Rat	Male	2	0.5	20 (40)	25 (50)	20.5
	Female	2	0.35	17.5 (50)	20 (57)	15.7
Hamster	Male	2	0.125	11.5 (92)	15 (120)	7.2
	Female	2	0.110	11.5 (105)	15 (136)	7.2

a) Note: for certain strains of mice documented lower lifespan values of minimally 1.5 years are acceptable (OECD TG 451);

b) In brackets the daily food or water consumption is given in g or ml per kg body weight per day, as appropriate.

c) These are typical values used for lifetime studies.

Table R. 8-18 Standard values for dose calculations for humans exposed in workplaces as consumers and via the environment (taken from Gold et al., 1984 and ICRP, 1975).

Parameter		DEFAULT Value	
Consumer, Humans-via-the-environment		Worker	
Lifespan (year)	75	Worklife (year)	40
Body weight (kg)	70	Length of workday (hour)	8
Food intake (kg/day)	1.4	Working days/week	5
Water intake (l/day)	2.0	Working weeks/year	48
Inhalation volume (m^3/24 hours)	20	Body weight, male and female (kg)	70
		Inhalation volume (m^3/8 hours) light work	10

REFERENCES

Gold, L.S.; Sawyer, C.B.; Magaw, R.; Backman, G.M.; de Veciana M.;, Levinson, R.; Hooper, N.K.; Havender, W.R.; Bernstein, L.; Peto, R.; Pike, M.C.; Ames, B.N. (1984) A carcinogenic potency database of the standardized results of animal bioassays. Environ. Health Perspect. 58, 9-319.

Paulussen JJC, Mahieu CM and Bos PMJ (1998). Default Values in Occupational Risk Assessment. Organisation for Applied Scientific Research (TNO), TNO Nutrition and Food Research Institute, TNO Report V98.390, Zeist, The Netherlands.

ICRP (1975). Report of the task group on Reference Man. International Commission on Radiological Protection (ICRP), Report No. 23, Pergamon Press, New York, NY. Quoted in Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP (1997). Physiological parameter values for physiologically based pharmacokinetic models, Toxicol. Ind. Health **13(4),** 407-484.

IGHRC (2006). Guidelines on route-to-route extrapolation of toxicity data when assessing health risks of chemicals. The Interdepartmental Group on Health Risks from Chemicals, http://www.silsoe.cranfield.ac.uk /ieh/ighrc/ighrc.html

Kalberlah, Först, Schneider (2002). Time extrapolation and interspecies extrapolation for locally acting substances in case of limited toxicological data. Ann Occup Hyg 46(2), 175-185.

WHO/IPCS (2005). Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration–response assessment. IPCS Harmonization Project Document No. 2, http://whqlibdoc.who.int/publications/2005/9241546786_eng.pdf

Schneider K, Oltmanns J and Hassauer M (2004). Allometric principles for interspecies extrapolation in toxicological risk assessment – empirical investigations. Regulatory Tox and Pharm, 39(3), 334-347.

APPENDIX R. 8-3 Assessment factors suggested from different research groups and regulatory bodies

In the DNEL/DMEL-approach, effect assessment uncertainties are dealt with by means of assessment factors, which should preferably be substance-specific, otherwise default (see Section R.8.4.3). This appendix presents a short overview (including a summary table, Table R. 8-19) with defaults suggested for various assessment factors by different research groups and regulatory bodies. The overview is not meant to be exhaustive, and for more background information and further reading, the reader is referred to the original publications.

As can be seen from Table R. 8-19, defaults typically proposed for human health risk assessment are point estimates. A more recent development is the suggestion for probabilistic distributions as defaults for assessment factors: as lognormal distributions are thought to best describe variability and uncertainty in assessment factors, these distributions have been derived based on NOAEL-ratios from comprehensive toxicological databases. Although promising, up to now these probabilistic distributions have not been widely used in risk assessment, a.o. because it requires decisions on the percentile of the population one wants to protect (e.g. 50^{th} percentile (= geometric mean of distribution) or 90^{th}, 95^{th} or 99^{th} percentile (= P_{90}, P_{95} or P_{99} of distribution).

Explanation

Assessment factors for interspecies differences

Interspecies differences result from variation in the sensitivity of species due to differences in toxicokinetics and toxicodynamics. Where human data are used as the starting point for the risk characterisation, no extrapolation is necessary and hence no assessment factor is normally suggested for interspecies differences in sensitivity.

Where data from animal studies are the typical starting point for risk characterisation, the default assumption in general is that humans are more sensitive than experimental animals. As can be seen from Table R. 8-19, the traditional default suggested for interspecies extrapolation is 10, which sometimes is subdivided in a default of 4 ($10^{0.6}$) for toxicokinetic differences and a default of 2.5 ($10^{0.4}$) for toxicodynamic differences.

Since some of the toxicokinetic differences can be explained by differences in body size (and related differences in basal metabolic rate), others have suggested as a default to, where appropriate, correct for differences in metabolic rate (allometric scaling; see Section R.8.4.3), followed by the application of a default factor for other toxicokinetic and toxicodynamic differences. The size of the latter varies from 1 to 3 (see also footnotes to Table R. 8-19). Next to these point estimates, also default lognormal distributions have been established for this additional factor.

Assessment factors for intraspecies differences

Humans differ in sensitivity due to a number of biological factors (such as age, gender, genetic composition and nutritional status). The intraspecies variation in humans is greater than in the more homogeneous experimental animal population.

Although other values have been proposed, defaults typically suggested for the general population (representing all age groups, including children and elderly) are a factor of 10, sometimes equally subdivided in defaults of 3.16 ($10^{0.5}$) for both toxicokinetic and toxicodynamic differences. A lower default factor is generally suggested for the worker population, because the very young and very old are not part of this population.

For the intraspecies assessment factor also probabilistic distributions have been proposed. It is to be noted that the ones proposed by Vermeire et al. (1999, 2001) are not database-derived distributions, but theoretical distributions.

Assessment factors for differences in duration of exposure

Taking into account that a) in general the experimental NOAEL will decrease with increasing exposure times and b) other and more serious adverse effects may appear with increasing exposure times, a factor allowing for differences in the experimental exposure duration and the duration of exposure for the population and scenario under consideration is normally applied in risk assessment.

As can be seen in Table R. 8-19 different factors have been suggested for exposure duration extrapolation, depending on the type of extrapolation (subacute to subchronic, subchronic to chronic, subacute to chronic) and the kind of effect (systemic or local). Probabilistic distributions have also been suggested.

Assessment factor for uncertainty in route-to-route extrapolation

Given the uncertain nature of route-to-route extrapolation and the fact that it can only be applied in specific cases, no defaults have been typically proposed for this factor, necessary in case no adequate data are available on the relevant route of exposure for the population and exposure scenario under consideration. Note that the present guidance addresses this as part of Section R.8.4.2 on modification of the starting point.

Assessment factor for dose-response relationship

For the dose-response relationship, consideration should be given to the uncertainties in the NOAEL as the surrogate for the true no-adverse-effect-level (NAEL), as well as to the extrapolation of the LOAEL to the NAEL (in cases where only a LOAEL is available or where a LOAEL is considered a more appropriate starting point). Taking into account the dose spacing in the experiment, the shape and slope of the dose-response curve (and in some approaches the extent and severity of the effect seen at the LOAEL), defaults typically suggested for this assessment factor range from 1–10 (see Table R. 8-19). The Benchmark dose has also been suggested as acceptable alternative to the LOAEL-NAEL extrapolation, or even a probabilistically derived benchmark dose distribution.

Other aspects relating to the dataset

Next to extrapolation, other important aspects of risk characterisation are the adequacy of and confidence in the available dataset and the nature of the effect. Most often these aspects are dealt with in a qualitative way. When dealt with in a quantitative way, default values of 1-10 have been proposed (see Table R. 8-19), but there is no agreed basis for these values. The US-EPA uses the term modifying factor to cover uncertainties other than the ‘extrapolation’ assessment factors.

Overall assessment factor

Typically, the overall assessment factor is the product of the individual assessment factors, by assuming independency of the factors. It is to be realised that this multiplication is in general very conservative: when each individual assessment factor by itself is regarded as conservative, multiplication will lead to a piling up of conservatism. Hence, the more extrapolation steps are taken into account, the higher the level of conservatism.

Although not widely used up to now, a more recent development in risk assessment is the use of probability distributions and Monte Carlo simulation to obtain the overall assessment factor. By

acknowledging that each assessment factor is uncertain and is best described by a lognormal distribution, propagation of the uncertainty can be evaluated by Monte Carlo simulation yielding a lognormal overall distribution for the combined assessment factor. This offers the possibility for a quantitative estimate of the probability that an adverse effect will occur in a certain population at the estimated exposure level. Moreover, the distribution of the overall assessment factor can be probabilistically combined with the distribution of the Benchmark dose, as also the effect parameter is uncertain and is best described by a lognormal distribution.

REFERENCES

Danish EPA (2006). Vejledning fra Miljøstyrelsen, 5/2006. Metoder til fastsættelse af kvalitetskriterier for kemiske stoffer i jord, luft og drikkevand med henblik på at beskytte sundheden. (Guidance from the Danish EPA no5/2006. Methods for the derivation of health based limit values for chemical substances in soil, air, and drinking water), link: http://www2.mst.dk/common/Udgivramme/Frame.asp?pg=http://www2.mst.dk/Udgiv/publikationer/2006/87-7052-182-4/html/default.htm

ECETOC (2003) Derivation of assessment factors for human health risk assessment. Technical Report No. 86, (European Centre for Ecotoxicology and Toxicology of Chemicals) Brussels.

Kalberlah F and Schneider K (1998). Quantification of extrapolation factors. Final report of the research project No. 116 06 113 of the Federal Environmental Agency. Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Fb 797, Dortmund/Berlin.

Kalberlah F, Schneider K, Schuhmacher US, Voss J-U, Ioannidis I and Oltmanns J (1999). Zeitextrapolation und Interspezies-extraplation bei lokal wirksamen Stoffen mit begrenzter Datenlage. Schriftenreihe der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Fb 862, Dortmund/Berlin. [Also published in: Annals of Occupational Hygiene, vol. 46, p.175-185, 2002]

Schneider K, Hassauer M, Oltmanns, Schuhmacher-Wolz U, Elmshäuser E and Mosbach-Schulz O (2005). Uncertainty analysis in workplace effect assessment, Project F1824/F1825/F1826, english version of the final report, Federal Institute for Occupational Safety and Health, Dortmund, http://www.baua.de/english/fors/f1824-26e.htm

TRGS 901, Ergänzungen (1998) Kriterien für die Ableitung von gesundheitsbasierten Luftgrenzwerten bei limitierter Datenlage. Bundesarbeitsblatt 10: 74-76. [Toxicology Advisory Group of the Dangerous Substances Committee - Criteria for the derivation of health-based airborne concentration limits from limited data. 26 June 1998. Doc. No 5722/98 EN]

US-EPA (1993) IRIS - Reference dose (RfD): description and use in health risk assessments. Background document 1A. U.S. Environmental Protection Agency, http://www.epa.gov/iris/rfd.htm.

Vermeire TG, Stevenson H, Pieters MN, Rennen M, Slob W and Hakkert BC (1999). Assessment factors for human health risk assessment: a discussion paper. Crit. Reviews Toxicol. **29** (5), 439-490.

Vermeire T, Pieters M, Rennen M and Bos P (2001). Probabilistic assessment factors for human health risk assessment – a practical guide. RIVM report 601516 005/TNO report V3489. National Institute of Public Health and the Environment (RIVM) (in cooperation with TNO Nutrition and Food Research), Bilthoven, The Netherlands.

WHO/IPCS (1987) Principles for the safety assessment of food additives and contaminants in food. Environmental Health Criteria 70. International Programme on Chemical Safety, WHO/FAO/UNEP/ILO, World Health Organization, Geneva.

WHO/IPCS (1990) Principles for the toxicological assessment of pesticide residues in food. Environmental Health Criteria 104. International Programme on Chemical Safety, WHO/UNEP/ILO, Geneva.

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

WHO/IPCS (1999) Principles for the assessment of risks to human health from exposure to chemicals. Environmental Health Criteria 210. International Programme on Chemical Safety, WHO/UNEP/ILO, World Health Organization, Geneva.

Table R. 8-19 Summary of default assessment factors used in human health risk assessment.

Assessment factors	WHO/ IPCS (1987, 1990, 1994, 1999)	US-EPA (1993)	ECETOC (2003)	BAUA (D) (TRGS, 1998)	Danish EPA 2001	Kalberlah & Schneider (1998); Kalberlah et al. (1999)	Schneider et al. (2005)	TNO/RIVM (NL) (Vermeire et al., 1999, 2001)	
							Probabilistic[m]	deterministic	Probabilistic[m]
Interspecies	10	10	AS[a]		10	AS[e,f]	AS[e] x (GM 0.97; GSD 3.24) [4.35 (P_{90}); 6.67 (P_{95}); 14.9 (P_{99})]		AS[i] x (GM 1; GSD 4.5) [7 (P_{90}); 12 (P_{95}); 33 (P_{99})]
Non-occupational - toxicokinetics - toxicodynamics	 4.0 2.5							10	
Occupational				AS[e]				AS[i] x 3	
Intraspecies	10	10			10				
Non-occupational - toxicokinetics - toxicodynamics	 3.16 3.16		5[b]			25 8 3		10	GM 1+3; GSD 1.6 [6.6 (P_{90}); 7.6 (P_{95}); 10 (P_{99})]
Occupational			3	5[g]		5[g]	for P_{90} of individuals: [n] GM 1+2.31; GSD 3.57 [12.8 (P_{90}); 19.8 (P_{95}); 45.7 (P_{99})] for P_{95} of individuals: [n] GM 1+3.82; GSD 4.34 [26.1 (P_{90}); 43.8 (P_{95}); 117 (P_{99})] for P_{99} of individuals: [n] GM 1+8.96; GSD 6.45 [98.7 (P_{90}); 193 (P_{95}); 687 (P_{99})]	3	GM 1+1.4; GSD 1.2 [2.7 (P_{90}); 2.85 (P_{95}); 3 (P_{99})]

Table 1 continued overleafAssessment factors	WHO/ IPCS (1987, 1990, 1994, 1999)	US-EPA (1993)	ECETOC (2003)	BAUA (D) (TRGS, 1998)	Danish EPA 2001	Kalberlah & Schneider (1998); Kalberlah et al. (1999)	Schneider et al. (2005)	TNO/RIVM (NL) (Vermeire et al., 1999, 2001)	
							Probabilistic[m]	deterministic	Probabilistic[m]
Duration of exposure *System. eff./Local inhal.*		10			1-100[p]				

eff.			2/NN[c]	2/4		2-3 (GM)	GM 4.39; GSD 1.82	10	GM 2; GSD 3.5
- sub-chronic to							[9.45 (P_{90}); 11.8 (P_{95}); 17.6 (P_{99})]		[10 (P_{90}); 16 (P_{95}); 37 (P_{99})]
chronic				2/4		2-3 (GM)	GM 3.95; GSD 2.14	10	GM 2; GSD 4
- subacute to sub-							[10.5 (P_{90}); 13.8 (P_{95}); 23.2 (P_{99})]		[12 (P_{90}); 20 (P_{95}); 50 (P_{99})]
chronic			6/NN[c]	6/12		6-7 (GM)	GM 4.14; GSD 2.03	50-100	GM 5; GSD 3.5
- subacute to chronic							[10.3 (P_{90}); 13.3 (P_{95}); 21.6 (P_{99})]		[25 (P_{90}); 39 (P_{95}); 92 (P_{99})]
			NN[c]					NN[j]	
Local dermal effects									
Route-to-route[#]			ND					ND[k]	ND[k]
Oral to inhalation				1[h]					
Oral to dermal				1[h]					
Type of leading effect	1-10							1	
Dose-response curve									
Appropriate NOAEL			NN						
LOAEL to NAEL	3-10	10	3[d]	3		10		1-10	
Alternative	BMD	BMD	BMD	BMD		BMD	BMD distribution	BMD	BMD (or BMD distribution)
Confidence in database/ database adequacy	1-10							1	
Modifying factor		>0-10							
Overall factor	mult.	mult.	mult.	mult. [o]	mult.	mult.	Prob. [q]	mult.	prob. + mult. (for point estimates) [l]

Abbreviations:

AS = allometric scaling ($bw^{0.75}$); BMD = benchmark dose; GM = geometric mean; GSD = geometric standard deviation; mult. = multiplication of the different assessment factors; ND = no default proposed; NN = no (additional) assessment factor needed; P = percentile; prob. = probabilistic combination of distributions for the different assessment factors

Notes:

Only for systemic effects (under certain conditions), not for local effects.

a mouse 7, rat 4, monkey 2, dog 2

AS not to be applied for inhalation route and for local effects; Although AS does not completely account for interspecies differences, no additional assessment factor for 'residual' interspecies variability because that is largely accounted for in the assessment factor for intraspecies variability.

b No additional assessment factor for children needed, but attention should be given to effects on developing organ systems, such as reproductive development in pre-puberty.

c For local effects below the threshold of cytotoxicity.

d May need to be adjusted depending on dose spacing, shape and slope of dose-response curve and extent and severity of effect seen at LOAEL.

e mouse 7, rat 4, dog 2, monkey (marmoset) 4, monkey (rhesus) 2 (rounded figures)

AS only to be applied for systemic effects, with doses in mg/kg bw (not for doses in mg/m^3 or mg/kg feed); not for local effects.

f Additionally to the AS a factor for possible additional toxicokinetic or toxicodynamic variability is applied, depending on percentile of population one wants to protect: 1 (50^{th} perc.) or 2-3 (95^{th} perc.).

g After allometric scaling this factor of 5 should be applied as combined assessment factor for intra- and interspecies extrapolation.

h Similar absorption by all routes is assumed (not necessarily 100%).

i mouse (25 g) 7, rat (250 g) 4, guinea pig (750 g) 3, rabbit (2 kg) 2.4, monkey (5 kg) 2, dog (15 kg) 1.4

AS only to be applied for systemic effects following oral and dermal route of administration (with doses in mg/kg bw), not for inhalation route and for local effects.

j For local skin effects it is assumed that exposure duration can influence the severity of the effects but will not influence the height of the NOAEL.

k Depends on substance-specific data on absorption for starting route and end route. In case no substance-specific data are available for both routes, a default factor of 2 is used, i.e. as a worst case assumption the absorption percentage for the starting route is half that of the end route.

l Based on the individual distributions that have been established for some assessment factors, Vermeire et al. (2001) have proposed default distributions of the overall assessment factors for the general population (including consumers) and for workers. If additionally point estimates are involved (e.g. for allometric scaling) these overall distributions have been multiplied accordingly. Vermeire et al. (2001) also give guidance on how to apply the probabilistic default distributions in human health risk assessment and on how to quantitatively interpret the results. For e.g. inter x intra x sub-chronic/chronic extrapolation (based on sub-chronic rat study) this works out as follows:

	Default distribution of combined AFs (inter x intra x sub-chronic/chronic)				AS (rat)	Default distribution of overall AF			
	GM	GSD	P90	P95		GM	P90	P95	P (of default)
General population	8	7.5	101	206	4	32	404	824	99 (10x10x10)
Workers	4.8	7.1	60	121	4	19	240	484	93 (3x4x3x10)

In this table the confidence limits (GM and P-values) are indicative of the percentile of the population that one wants to protect (e.g. GM: 50%, P95: 95%).

m Lognormal distributions with parameters geometric mean (GM); geometric standard deviation (GSD) and shift, if not zero.

n The three distributions cover the difference between the median sensitive and the sensitive individual at the workplace. Sensitive individuals are defined as being equal or more sensitive than 90%, 95%, or 99% of the working population (P_{90}, P_{95}, P_{99}-percentile of the interindividual distribution for a specific substance), respectively, for distributions 1, 2, and 3. Distributions (lognormal with shift 1) describe variation over substances and case studies evaluated in regard to inter-individual sensitivity

o By estimating the different parameters as typical values with central tendency, the product of these parameters reveals a central tendency estimate of the combined assessment factors. For evaluation of existing chemicals this approach is modified as follows: an additional factor is used to account for the uncertainty of the assessment and the confidence in the database. By multiplication with this factor the initial estimate is modified in terms of precaution. The resulting value represents the overall assessment factor.

p An overall assessment factor concerning "quality and relevance" of the data covering uncertainties in relation to e.g. LOAEL to NOAEL extrapolation; duration of exposure; route-to-route extrapolation; severity of effects; lack of data etc.

q The individual distributions that have been established, including the BMD distribution, are combined by probabilistic modelling. As result a target distribution is obtained which represents a substance-specific probabilistic estimate of the health-based reference value and its uncertainty for a certain quantile of the human population (e.g. P95 of individuals). If another quantile is of interest, a new calculation using the according distribution for intraspecies differences (e.g. P99 of individuals) is performed.

APPENDIX R. 8-4 PBPK Modelling and the derivation of DNELs/DMELs

This document describes how PBPK modelling can be used in the derivation of DNELs/DMELs. The focus of the document is on how the different extrapolation processes involved in risk assessment can be performed by using PBPK modelling. However, it is important to emphasise that when PBPK modelling is used for risk assessment purposes, the whole process of PBPK modelling (i.e. the generation of the model, application of the model, validation of the model, confidence in the model, etc...) should be as transparent as possible. Confidence in the model should be as high as possible. Furthermore, risk assessors, who are using these models, should be able to adequately interpret these models and their output.

PBPK Modelling

A physiologically based pharmacokinetic (PBPK) model is an independent structural mathematical model, comprising the tissues and organs of the body with each perfused by, and connected via, the blood circulatory system. The principle application of PBPK models is in the prediction of the *target tissue dose* of the parent chemical or its reactive metabolite. Use of the target tissue dose of the toxic moiety of a chemical in risk assessment calculations provides a better basis of relating to the observed toxic effects than the external or administered exposure concentration of the parent chemical. Prediction of target tissue dose following different exposure scenarios, routes, doses and species can help reduce the uncertainty associated with conventional extrapolation approaches. The mechanistic and biological plausibility of the models is the basis for associating greater confidence to such extrapolations.

Parameters in PBPK modelling

PBPK models comprise four main types of parameter:

1. Physiological
2. Anatomical
3. Biochemical
4. Physicochemical

Physiological and anatomical parameters include tissue masses and blood perfusion rates, estimates of cardiac output and alveolar ventilation rates. Biochemical parameters include e.g. enzyme metabolic rates and polymorphisms, enzyme synthesis and inactivation rates, receptor and protein binding constants. Physicochemical parameters refer to e.g. partition coefficients, vapour pressures, solubilities in different media. A partition coefficient is a ratio of the solubility of a chemical in a biological medium, usually blood-air and tissue-blood.

The ability to *quantify* the parameters that comprise models affords the ability to conduct the various extrapolations discussed below. Anatomical and physiological parameters are readily available and many have been obtained by measurement. Biochemical and physicochemical parameters are compound specific. When such parameters are measured and used to construct an *a priori* model that qualitatively describes a dataset, then confidence in such a model should be high. In the absence of measured data, such as partition coefficients, these may be estimated using tissue-composition based algorithms (Poulin and Krishnan 1995, 1996; Theil *et al.* 2003). Metabolic rate constants may be fitted using a PBPK model, although this practice should only be undertaken if there are no other alternatives.

The importance of any single or set of parameters within a model should be determined by applying sensitivity analysis. Sensitivity analysis is a means of evaluating how sensitive the model output is to any perturbation of any single or set of parameters. Therefore, confidence in a model that contains estimated parameters may still be high, if those parameters do not significantly influence model output. Conversely, influential estimated parameters would reduce confidence in a model.

Confidence in PBPK modelling

When used for risk assessment purposes, confidence in PBPK models should be high (Barton et al. (2007)). Therefore, their predictive capacity should be carefully evaluated with respect to the following aspects (US EPA, 2006):

- Model verification (i.e. biological plausibility of the model structure and parameter and correctness of the mathematical equations);
- Model validation (i.e. ability of the model to predict the kinetic behaviour of a compound);
- Model documentation;
- Sensitivity, variability and uncertainty analyses.

Risk assessors need to be able to understand these processes well enough when they are used in risk assessment and for decision making purposes.

Interspecies Extrapolation

Interspecies extrapolation of the pharmacokinetic behaviour of a chemical requires quantitative estimates of species differences in the values of these parameters. Tissue-blood partition coefficients of chemicals appear to be relatively constant across species, while blood-air partition coefficients show some species-dependent variability (Gargas *et al.* 1989). Physiological and anatomical parameters generally vary coherently across species. These parameters are readily available in the literature (Brown *et al.* 1997; ICRP 1975) and can therefore be used in a PBPK model where quantitative differences can be evaluated. The kinetic constants for metabolizing enzymes do not necessarily follow any type of readily predictable pattern (Dedrick and Bischoff 1980). The approach adopted in the past, and one that is still often used, is to apply the "metabolic rate scaling" (Section R.8.4.3.1 and APPENDIX R. 8-2). Therefore, a metabolic rate constant, such as V_{max}, obtained in a rodent would be multiplied by body weight of the human raised to the 3/4 power to obtain the human equivalent. This scaling factor is generally justified on the basis of the studies by Schneider *et al* (Schneider *et al.* 2004), who examined the interspecies differences in toxicity of a variety of chemicals.

For a chemical that demonstrates significant interspecies variation in toxicity in animal experiments, the most susceptible species is generally used as the reference for this extrapolation. Uncertainty factors of 10 to 1,000 or more have been applied in recognition of the uncertainty involved. Whereas a metabolic rate constant estimated in this way may be used in a PBPK model, it is preferable, where possible, to determine such parameters *in vitro* using tissue subcellular fractions or estimate them by fitting a PBPK model to an appropriate dataset. Furthermore, if a PBPK model is used to extrapolate from animals to humans, the proposed model should be validated by data from humans if these are available, and extrapolations from the model should be within or close to the range of experimental measurements used to validate the model. If there is no validation of the model by data from humans, PBPK models may be used to support an interpretation of toxicodynamic data or toxicological findings rather than as a basis for the derivation of a DN(M)EL (ECB, 2002).

Intraspecies Variability

Differences in sensitivity to exposure to chemicals within the same species occur as a result of variation in anatomical, physiological and biochemical parameters with age, gender, genetic predisposition and health status. These may be further confounded by nutritional and other lifestyle and environmental factors. The quantification of these parameters using PBPK models to determine the differences in tissue dose in intraspecies variability is analogous to the quantification of interspecies variability. For example, age-specific parameters would be required to estimate the tissue doses in adults and young children. Such data are increasingly available. The propagation of uncertainty and variability from model parameters to model output can be quantified using probabilistic techniques such as Monte Carlo sampling. A PBPK model is run with parameter values sampled from distributions that reflect the observed variation in each pharmacokinetic parameter in the human population. Each time the model is run with a sampled set of parameter values, effectively representing a single hypothetical human being, the appropriate dose metric for the toxicity of interest is output. The process is repeated a large number of times to generate a distribution of the dose metric for a simulated population. It is important to note that human physiological data have a range of values. Therefore, modelling should be preferentially performed with ranges of values leading to distributions of outcome

High-Dose-Low-Dose Extrapolation

The non-linear kinetic behaviour of chemicals in a biological organism is the result of a number of mechanisms e.g., saturable metabolism, enzyme induction, enzyme inactivation and depletion of glutathione and other cofactor reserves. High-dose-low-dose extrapolation of tissue dose is accomplished with PBPK modelling by accounting for such mechanisms (Clewell III and Andersen 1987).

Route-to-Route Extrapolation

Route-to-route extrapolations can be conducted quite readily with PBPK models. For example, the procedure would involve describing a model for the inhalation route. Ideally the model would be validated against an appropriate dataset. Equations describing other routes of administration, such as, dermal and oral may be added later and again, ideally, the model should be validated against a different, but appropriate dataset. In the case of oral uptake, first-pass metabolism and enterohepatic circulation may also be included if significant elimination of parent chemical occurs due to these mechanisms (Clewell III and Andersen 1987).

PBPK Modelling and the Development of Assessment Factors

A simple but fictional example of the development of an assessment factor for *interspecies* differences using PBPK modelling is presented. A fictional chemical, compound A, is a low molecular weight, volatile solvent, with potential central nervous system (CNS) depressant properties. Evidence for the latter comes from a number of controlled human volunteer studies where a battery of neurobehavioural tests were conducted during, and after, exposure by inhalation to compound A. Due to a number of inconsistencies in the type of tests performed in the different studies, a clear, robust NOAEL could not be identified from these human data.

Compound A is metabolised *in vitro* by the phase 1, mixed-function oxidase enzyme, cytochrome P450 2E1 (CYP2E1) by both rat and human hepatic microsomes. There are also some *in vivo* data in rats exposed by inhalation to compound A, with and without pre-treatment with diallyl sulphide, an inhibitor of CYP2E1, that are consistent with metabolism of compound A by this enzyme.

PBPK models for the rat and standard human male or female for exposure by inhalation to compound A are built. The rat model was validated by simulating experimentally determined decreases in chamber concentrations of compound A following exposure of rats to a range of initial concentrations in a closed-recirculated atmosphere exposure chamber. The removal of chamber concentration of compound A over time is due to uptake by the rat and elimination, primarily by metabolism. The human PBPK model was validated by simulating experimentally determined venous blood concentrations of compound A in male and female volunteers exposed by inhalation to a constant concentration of compound A in a controlled-atmosphere exposure chamber.

It is assumed that the following have been identified for the chemical: 1) the active moiety of the chemical, and 2) the relevant dose-metric (i.e., the appropriate form of the active moiety e.g., peak plasma concentration (C_{max}), area-under-the-curve of parent chemical in venous blood (AUC_B), average amount metabolised in target tissue per 24 hours, peak rate of hepatic metabolism, etc). In this case, it is hypothesised that the peak plasma concentration C_{max} of compound A is the most likely *surrogate* dose metric for CNS concentrations of compound A thought to cause a reversible CNS depressant effect. However, C_{max}, is dependent upon the peak rate of hepatic metabolism. Therefore, the validated rat and human PBPK models were run to simulate the exposure time and concentrations of the rat studies at which no CNS depressant effects were observed. The dose metric, peak rate of hepatic metabolism$_t$ for the rat would be divided by the peak rate of hepatic metabolism for the human. This ratio would represent the magnitude of the difference between a specified rat strain and average human male or female. This value may then replace the default interspecies kinetic value since it is based on chemical-specific data. Therefore, the derivation of an appropriate 'assessment factor' in setting a DNEL can be justified more readily using quantitative and mechanistic data.

PBPK Modelling in Risk Assessment

PBPK models will not remove all of the uncertainty from the risk assessment process. The rationale for using PBPK models in risk assessment is that they provide a documentable, scientifically defensible means of bridging the gap between animal bioassays and human risk estimates. In particular, they shift the risk assessment from the administered dose to a dose more closely associated with the toxic effect by explicitly describing their relationships as a function of dose, species, route and exposure scenario. The increased complexity and data demands of PBPK models must be counter-balanced by the increased accuracy, biological plausibility and scientific justifiability of any risk assessment using them. It follows from this that PBPK models are more likely to be used for chemicals of high concern. A guidance document on "Good Practice in PBPK modelling" is under preparation by WHO/IPCS and should be taken into account when PBPK modelling is used in risk assessment.

REFERENCES

Barton, H.A., Chiu, W.A., Setzer, R.W., Andersen, M.E., et al. (2007). Characterizing uncertainty and variability in physiologically-based pharmacokinetic (PBPK) models: state of the science and needs for research and implementation, Toxicol Sci., (in press). [Toxicol Sci., Advance Access published May 4, 2007 doi: 10.1093/toxsci/kfm100].

Brown, R. P., Delp, M. D., Lindstedt, S. L., Rhomberg, L. R., and Beliles, R. P. (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and Industrial Health* **13**, 407-484.

Clewell III, H. J., and Andersen, M. E. (1987). Dose, species and route extrapolation using physiologically-based pharmacokinetic modeling. *Drinking Water and Health* **8**, 159-184.

Dedrick, R. L., and Bischoff, K. B. (1980). Species similarities in pharmacokinetics. *Fed Proc* **39**, 54-9.

ECB (2002): Technical Guidance Document on Risk Assessment

Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H., and Andersen, M. E. (1989). Partition coefficients of low-molecular weight volatile chemicals in various liquids and tissues. *Toxicology and Applied Pharmacology* **98**, 87-99.

ICRP (1975). *Report of the Task Group on Reference Man*. Pergamon Press, New York.

Poulin, P., and Krishnan, K. (1995). A biologically-based algorithm for predicting human tissue:blood partition coefficients of organic chemicals. *Hum. Exp. Toxicol.* **14**, 273-280.

Poulin, P., and Krishnan, K. (1996). A mechanistic algorithm for predicting blood:air partition coefficients of organic chemicals with the consideration of reversible binding in hemoglobin. *Toxicol. Appl. Pharmacol.* **136**, 131-137.

Schneider K, Oltmanns J and Hassaner M (2004) Allometric principles for interspecies extrapolation in toxicological risk assessment - empirical investigations. *Regulatory Tox Pharm* **39(3)**, 334-347.

Theil, F. P., Guentert, T. W., Haddad, S., and Poulin, P. (2003). Utility of physiologically based pharmacokinetic models to drug development and rational drug discovery candidate selection. *Toxicology Letters* **138**, 29-49.

US EPA (2006): Approaches for the Application of Physiologically-Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessment. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=157668.

APPENDIX R. 8-5 Derivation of DNELs using biomonitoring data

It is generally recognised that systemic effects of substances are related to the concentration of the substance itself or one of its active metabolite(s) in the target organ. Biomarkers of exposure are the concentrations of these substances or their metabolites measured in one of the bodily fluids (e.g. blood and urine) and are inherently more closely related to the actual dose in the target organ than external dose metrics. Hence, they should be considered most relevant for human risk assessment. In addition, biomarkers of exposure provide an integrated measure of total exposure, irrespective of the route of exposure.

Whenever exposure of workers, consumers or humans-via-the-environment can be assessed through biomonitoring, the use of a DNEL derived from biomonitoring data can be straightforward and preferable to avoid additional uncertainty by the need of using assessment and conversion factors. Moreover, it has been shown that biomarker data have less variance than air measurements, particularly in environmental settings [1].

Biomarker DNELs can be derived from biomonitoring data using the same methodologies being used to derive health-based biological limit values. In principle, two different situations may exist in which establishing a biomarker DNEL may be preferable.

A clear dose-response correlation exists between biomarker levels and the effect(s) in humans or in animals

In this case, the simplest situation occurs if a dose-effect correlation between the biomarker and the effect is obtained from human data since in that case the NOAEL expressed as the biomarker level can directly serve as the basis for the DNEL. Although the assessment factor(s) for interspecies extrapolation and route-to-route extrapolation are of course not applicable, other assessment factors, such as that for intraspecies variation, should be applied, as appropriate. The most straightforward examples of this situation comprise the heavy metals (e.g. lead, cadmium) and a number of organic solvents (e.g. acetone, cyclohexane).

If the biomarker dose and effects were measured in animal studies, again, the NOAEL expressed as the biomarker dose level can directly serve as the basis for the DNEL using the assessment factors for inter- and intraspecies variation, but without correcting for route-to-route extrapolation as they are superfluous for biomarker dose-based DNELs.

***Example 1*:**

Exposure to acrylamide may occur by various routes and is best assessed by determination of its N-terminal valine adduct in haemoglobin (N-2-carbamoylethylvaline). Using this parameter, a NOAEL for the peripheral nervous system effects of acrylamide was found at 0.51 nmol/g globin in 200 workers occupationally exposed to acrylamide [2]. This value can directly serve as the basis for the DNEL for this effect. Since the NOAEL was obtained in a relatively homogeneous group of individuals, an assessment factor might be applied to account for intraspecies variation to derive a DNEL for the general population.

A clear relationship exists between an external exposure metric, which is linked to the effect(s), and the biomarker in the same species.

Although in this situation a biomarker DNEL cannot directly be derived from the biomarker data, it may nevertheless be worthwhile to derive a DNEL using the biomarker data. In particular, if exposure monitoring is routinely done through biomonitoring, which is often the case for substances with low and/or intermittent exposures and for substances with more than one significant exposure route, then biomonitoring is more reliable than air monitoring [1, 3].

Expressing DNELs in biomarker dose units is feasible in cases where (1) a relationship has been established between an external exposure metric (often inhalation exposure measurements) and the biomarker in the same species in which the external exposure metric has been linked to the effect, or (2) a validated (pharmacologically-based) toxicokinetic (PBPK) model establishing correlations between internal and external dose metrics is available for the same species in which the external dose metric has been linked to the effect (e.g. 2-butoxyethanol, [5, 6]). Using these relationships, the biomarker can be linked to the external marker or vice versa. Subsequently, a DNEL can be established using (human or animal) data in the same way as is done for the more conventional external exposure metric using assessment factors to compensate for inter- (in case of animal data only) and intra-species differences (in case of both animal and/or human data).

Example 2:

The relationship between airborne concentrations of 2-butoxyethanol (2-BE) and the urinary excretion of its major urinary metabolite 2-butoxyacetic acid (BAA) has been established in several studies in a variety of circumstances in humans [7-9]. A linear correlation was found over the range from below detection level up to 25 ppm of 2-BE in air. The most conservative relationship gives a slope of 16 mg BAA/g creatinine per ppm (8-h TWA) of airborne 2-butoxyethanol [8]. This relationship can be used to express BAA biomarker data directly as airborne equivalent concentrations and vice versa. In addition, the toxicokinetics of 2-BE have been well described and a validated PBPK model and data from human volunteer studies are available [5, 6, 10]. This allows linking a NOAEL obtained in animals, expressed as the concentration of a biomarker, to the external dose metric needed to reach the same concentration of the same biomarker in humans. A DNEL can then be simply derived by application of the same assessment factors used to derive a DNEL for an external dose metric based on animal data.

In addition, the PBPK model would allow to disentangle the relative contributions of dermal exposure if biomonitoring data are available in combination with personal air monitoring data, which is essential for effective risk management. Ideally, this is done on an individual basis, i.e. with the internal and external exposure metric determined in the same person. If individual data are not available, group averages may be used to get a realistic estimate of the average contributions of the different exposure routes. The use of modelling, in particular PBPK modelling, may reduce uncertainty due to route-to-route extrapolation and species differences.

Example 3:

The DNEL value for acrylamide (see previous example) can be linked to airborne values to compare the human data with data obtained in animal experiments. For humans, the concentration of acrylamide, expressed in $\mu g/m^3$, can be calculated from the concentration of Hb adducts in blood (Cblood), expressed in pmol N-2-carbamoylethylvaline per gram of globin, by the following equation: $Cair = (0.197 \times Cblood) + 5.10$ [4].

Hence, it can be calculated that the background values of N-2-carbamoylethylvaline measured in the study by Hagmar et al. [2] in the workers without occupational exposure to acrylamide, which are probably due to oral exposure to acrylamide from food, correspond to airborne levels of acrylamide up to 9 $\mu g/m^3$.

Similarly, the highly exposed workers had experienced equivalent airborne exposures up to 3.5 mg/m^3 of acrylamide. The human NOAEL for peripheral nervous system effects of 0.51 nmol/g globin, which can serve as the basis for a DNEL, corresponds to a calculated airborne exposure to acrylamide of 95 $\mu g/m^3$ (8-h TWA). This value of 95 $\mu g/m^3$, which can only be derived from the human biomarker data, can subsequently be compared to animal data.

[Note: applying the standard conversion factors (70 kg individual inhaling 10 m^3 per day), the value of 95 $\mu g/m^3$ (8-h TWA) corresponds to an exposure to acrylamide of 0.014 mg/kg body weight, suggesting that in humans peripheral neuropathy is a more critical endpoint than cancer.]

In the human situation often only a single dose-effect relationship is available. If this is not the leading health effect, extrapolation based on animal data or (toxicokinetic) models may be necessary.

REFERENCES

Lin YS, Kupper LL, Rappaport SM. Air samples versus biomarkers for epidemiology. Occ Env Med 2005;62:750-760

Hagmar L, Törnqvist M, Nordaner C, Rosen I, Bruze M, Kautiainen A, Magnussen A-L, Malmberg B, Aprea P, Granath F, Axmon A. Health effects of occupational exposure to acrylamide using haemoglobin adducts as biomakers of internal dose. Scand J Work Env Health 2001;27:219-226

Boogaard PJ. Use of haemoglobin adducts in exposure monitoring and risk assessment. J Chrom B 2002;778:311-324

Jones K, Garfitt S, Emms V, Warren N, Cocker J, Farmer P. Correlation of haemoglobin-acrylamide adducts with airborne exposure: an occupational survey. Toxicol Lett 2006;162:174-180

Corley RA, Bormett GA, Ghanayem BI. Physiologically based pharmacokinetics of 2-butoxyethanol and its major metabolite, 2-butoxyacetic acid, in rats and humans. Toxicol Apl Pharmacol 1994;129:61-79

Corley RA, Markham DA, Banks C, Delorme P, Masterman A, Houle JM. Physiologically based pharmacokinetics and the dermal absorption of 2-butoxyethanol vapour by humans. Fund Appl Toxicol 1997;39:120-130

Angerer J, Gündel J. Butoxyacetic acid, in (Angerer J, Schaller KH, eds): Analysis of hazardous substances in biological materials, DFG, Volume 4, VCH Verlag, Weinheim, Germany, pp. 131-145, 1994

NIOSH. Criteria for a Recommended Standard. Occupational exposure to ethylene glycol monobutyl ether and ethylene glycol monobutyl ether acetate. US Dept. of Health and Human Services, Public Health Service, Centers for Disease Control, Washington, DC, USA, 1990

Laitinen J. Correspondence between occupational exposure limit and biological action level values for alkoxyethanols and their acetates. Int Arch Occup Environ Health 1998;71:117-124

Jakasa I, Mohammadi N, Krüse J, Kežić S. Percutaneous absorption of neat and aqueous solutions of 2-butoxyethanol in volunteers. Int Arch Occup Environ Health 2004;77:79-84

APPENDIX R. 8-6 Animal dose descriptors for non-threshold carcinogenic responses

Derivation of the dose-descriptor

Available animal carcinogenicity studies are compared and evaluated with regard to their suitability for analysis of carcinogenic activity. The data for calculating dose-descriptors should preferentially be from lifetime oral studies (feed, drinking water or gavage) or inhalation studies according to Annex V of Directive 67/548/EEC or other accepted guidelines (e.g. OECD guidelines). Occasionally, skin painting studies may be used in the quantitative risk assessment. The use of other studies can be considered on a case-by-case basis.

The standard lifespan is considered to be 2 years for rats and mice. For certain strains of mice the standard lifespan may be between 18 months to 2 years.

In case a cancer study does not have full compliance to an actual testing guideline with regard to duration of exposure and/or observation, the following corrections should be made. If an experiment is terminated before the standard lifespan, the number of tumours found is assumed to be an underestimate of the number that would have been present after lifetime administration. If, for example, dosing is terminated at w weeks (w<104 weeks) before the standard lifespan of 104 weeks and the animals are observed until termination of the experiment at 104 weeks, the lifetime daily dose d giving the observed tumour incidence is corrected by w/104. If dosing is terminated at w1 and observation is until w2 weeks, the lifetime daily dose giving the observed tumour incidence is corrected by (w1/104) (w2/104).

Thus, for an experiment lasting for 18 months in rats with the standard lifespan of 24 months the dose descriptor will then be corrected by $(18/24)^2$ x applied dose.

Along the same line of argument: if animals are dosed 5 days per week, the daily dose giving the observed tumour incidence over lifetime will be derived by simply correcting the applied dose as follows: daily dose = (5/7) x applied dose.

For a substance considered genotoxic any tumour type observed in an animal bioassay is taken as relevant to humans and as starting point for a dose-descriptor determination, unless evidence to deviate from this approach is considered sufficiently convincing. This also concerns specific tumour types for which evidence for an alternative non-genotoxic mode of action is available.

Though presumed to be already considered in the effects assessment part of the evaluation, it is stressed here once again that the relevance of a potential carcinogenic response should primarily be based on biological criteria rather than on statistics alone. For example, the occurrence of 3 brain tumours in a population of 50 animals, although not necessarily statistically significant can be a good indication of a carcinogenic response if the (historical) control population very rarely shows brain tumours. In contrast a 10% increase of liver tumours in e.g. B6C3F1 mice may be statistically significant but can very well be within the historical control range. Further biological evidence for the plausibility of the carcinogenic response, e.g. pre-neoplastic changes, target organ toxicity, DNA adducts, etc., should be taken into account.

For a substance inducing more than one type of tumours, the determination of a dose-descriptor value is from each relevant tumour type rather than from the number of tumour bearing animals. If several relevant data sets on tumour-incidences are available, dose-descriptors values should be derived for all these. Malignant tumours as well as benign tumours that are suspected of possibly progressing to malignant tumours are taken into account in obtaining the dose-descriptor values.

Further details about how to calculate the dose-descriptors T25 and BMD10, may be found later in this appendix (for T25 also Annex 3 of EC (1998) and Dybing et al. (1997)).

Dose-descriptors

T25

The dose-descriptor T25 is defined as the chronic dose rate that will give 25% of the animals' tumours at a specific tissue site after correction for spontaneous incidence, within the standard life time of that species. It is a value calculated from a single observed dose-response and based upon the assumption of a linear dose-response relationship over the entire dose-range (Dybing et al., 1997).

The lowest tumorigenic doses showing a significant response (on statistical or biological basis) are generally used for obtaining the dose-descriptor value. If based on a statistical analysis, a statistical significance of $p<0.05$ is used as a border for rare tumour types, or tumour types having a relatively 'normal' background incidence, i.e. at most 10%. For tumours that have a higher spontaneous incidence a significant response is obtained at a $p<0.01$ level, if no specific reasons for deviating from this rule are available. Also, positive trend tests and historical incidences should be discussed when considered relevant with respect to the determination of a dose-descriptor.

The dose-descriptor value T25 is calculated from the tumour incidence at the selected tumorigenic dose (determined above), using linear intrapolation or extrapolation (e.g. in case of a net 15% incidence, multiply by 25/15). If the tumour incidence at higher dose levels results in a lower T25, this latter value is used unless the higher tumour incidence is likely to be associated with increased general toxicity or local toxicity at the tumour site, which may have interfered with tumour formation.

The T25 values are not corrected for intercurrent mortality. This should not be a problem in situations where the dose levels do not materially affect non-neoplastic causes of death. In situations where there is premature death due to toxicity, the use of data from the lowest dose giving a response considered statistically and/or biologically significant should in most instances avoid the problem of intercurrent mortality or at least reduce it to an acceptable degree. If this is not the case, however, dose-response modelling with Kaplan-Meier survival adjustment should be performed (see below).

In order to create a common scaling basis for all carcinogens, the dose-descriptor value should be expressed in mg/kg bw/day. To enable a conversion of feed, drinking water or air concentrations of carcinogens to this dose descriptor, physiological parameters should be used, which normally are provided by the study itself, for further support typical default values for lifetime studies are given in Table R. 8-14 1 of APPENDIX R. 8-2. In cases where only inhalation is relevant the unit mg/m^3 may be used directly without the need for conversion to mg/kg bw/day.

BMD10

The BMD10 is defined as the Benchmark-dose representing a 10% tumour response upon lifetime exposure. This response value is derived by fitting quantitative information available from all dose levels using a multistage model, a widely accepted model for the dose-response relationship in carcinogenicity.

The choice of this 10% response is influenced by the discriminating power of the animal experiment. If the control shows a response of 0 in 50 animals, a statistically significant ($p< 5\%$) difference based on the Fisher exact test (one sided) starts at 5 in 50 animals, i.e. an incidence of 10%. However, as in most cases several dose groups are used, a trend test can detect a significant difference at lower incidences, but only by considering simultaneously the other dose groups. The net 10% incidence chosen is thus almost always within the observed range.

However, in order to use the advantages that the BMD10 offers, it is necessary that a good set of quantitative data are available, i.e. a control population and at least three dose levels. In case the model shows a good fit it is recommended that the point estimate of the BMD 10% is calculated. A quality check of the data set may be obtained by comparing the point estimate and the lower confidence limit of the BMD 10% as suggested by Murrell et al. (1998) (see later in this appendix).

The justification for using a BMD (rather than the BMDL) is that (1) the point estimate is the best estimate of the response, (2) a sufficient amount of conservative assumptions are included in a linear extrapolation to the origin and (3) it is in line with the procedure used for the T25 methods which also uses the point estimate as the starting point.

The selection of the appropriate tumour endpoints and calculation of lifetime dose follows the same route as described for the T25 method.

BMDL10

The BMDL10 is defined as the lower 95% confidence dose of a Benchmark-dose representing a 10% tumour response upon lifetime exposure, i.e. the lower 95% confidence dose of a BMD10. The use of the BMDL10 rather than the BMD10 is recommended if one wants to reflect the uncertainties and statistical errors in the available cancer dose-response data.
For further details see BDM10 description.

Use of dose descriptors

Within the EU context the T25 dose-descriptor is in use for inclusion of potency considerations in setting specific concentration limits for carcinogens in Annex I of Directive 67/548/EEC (EC, 1998). Additionally, a T25 has been used by some Member States for risk characterisation within the framework of existing substances (EC Regulation 1488/94; EC, 1994) and for risk characterisation of chemicals in general. Moreover, the Scientific Committee on Consumer Products (previously: Scientific Committee on Cosmetics and Non-Food Products) uses the T25 method for safety evaluation in relation to non-threshold carcinogens SCCNFP/00690/03.

The BMD10 has regulatory use as its lowest confidence value (i.e. BMDL10) for the assessment of risks posed by food ingredients by EFSA (EFSA, 2005). A comparable dose descriptor, ED10, has regulatory use as its lowest confidence value (i.e. LED10) in cancer risk assessment practice by US EPA (US EPA, 2005).

Comparison of the dose-descriptor

An extensive comparison by Crump et al. (2000) showed that in case of linear or close to linear dose response relationships the results of the two procedures (i.e. the T25 and BMD05 procedures) are virtually identical, and estimates by the T25 method were at maximum two-fold higher or lower for sub- or supralinear dose response relationships, respectively.

In agreement with this it has been found that when risk assessments based on the same data sets were carried out using the T25 method and the linearised multistage model or the LED10 method, the doses representing a certain low risk calculated with the T25 method were on the average only 20 – 25% higher than with the two US methods. Only in very few cases, it was found that the doses calculated with the T25 method were more than double or less than half that calculated with the US methods (Sanner et al., 2001).

In case of a linear response a BMD10 will produce the same value as the T25 method, however, with clearly sublinear or supralinear dose response curves it will do more justice to the available dose-response relationship from the bioassay than the T25 method.

The BMD10 dose descriptor should not be applied, however, if some of the dose groups are considered to be inappropriate for this purpose; e.g. in situations of clear (target) organ and/or general toxicity. This may be the case for dose groups higher than the one already inducing a significant tumour response. Because of the potential involvement of other mechanisms at these experimental exposure levels not relevant for low dose risk characterisation, the incorporation of these dose-response data in deriving the dose descriptor appears inappropriate. Usually, dose selection based on the guidance in the OECD Test Guideline on carcinogenicity testing, TG451 or on the equivalent EU method (B32) from Annex V of Directive 67/548/EEC will avoid the inclusion of such unsuitable dose groups in the study design.

Presentation of the dose descriptors

It is recommended that, in case of multiple tumours of interest or a single tumour type found in several studies, all quantitative results are summarised in a table.

A graphical presentation of T25 and BMD10 in relation to the experimental data may also be helpful in deciding on which dose-descriptor should be used for lifetime risk or MOE calculations.

Kaplan-Meier Adjustment

In chronic studies with the goal to identify the carcinogenic potency of a substance several groups of experimental animals are exposed to different levels of the substance. Unfortunately, a carcinogenic substance may induce not only specific tumours, but also non-neoplastic effects. Neoplastic and non-neoplastic effects may have an impact on the survival. The survival may decrease with increasing dose level. A decreased survival with increasing dose level diminishes the incidence of tumours over the life span in the higher dose groups and troubles the recognition of the true dose-response relationship. In such particular cases, the calculation of the Kaplan Meier tumour probability may help to provide a more correct dose-response relationship. The advantage of the Kaplan Meier tumour probability is that the specific tumour incidence is calculated as if it is the only serious event in the study without interference of any other death cause. The shape of the dose-response (Kaplan-Meier tumour probability) can be explored as follows:

- Dose-response analysis at a specific observation period (response = Kaplan Meier tumour probability).

Further, one of the characteristics of a true carcinogen is that at higher dose levels tumours tend to occur earlier than at lower dose levels. A true carcinogen diminishes the time to tumour occurrence with increasing dose level. It is quite a difference, if a tumour occurs at day 350 or at day 700 of a chronic study.

The current guideline practise of overall counting of animals with tumours per dose group at the end of the study does not take this aspect into account. In those rare cases where there is enough information on the exact time to tumour incidence, a further analysis may be performed:

- Dose-time-response analysis, in which the response is related to both observation period and dose level (response = Kaplan Meier tumour probability).

The estimation of the Kaplan Meier tumour probability (P_{K-M}) is explained as follows:

Let us assume that a number of events (observation of animals with a specific tumour at death) n_i occurs at time t_i and that at time t_i there are Y_i survivors. The period of observation starts at t_1 and ends at t_C. The Kaplan-Meier tumour probability of having a tumour at death is represented by:

$$P_{K-M}(t_i) = 1 - \prod_{t \geq t_i}^{i=1 \text{ to } C} \left[1 - \frac{n_i}{Y_i}\right] \quad if \quad t \geq t_i \quad t < t_C$$

$$n_i = number\ of\ events\ at\ time\ t_i$$

$$Y_i = number\ of\ survivors\ at\ time\ t_i$$

The Kaplan-Meier tumour probability can be fitted simply to the dose in case of a specific observation time, identical for all dose groups, according to equation 1 below.

The value of the regression coefficient B_2 controls the shape of the dose-response

$$P_{K-M}(D) = 1 - \exp(-B_0 - B_1 * D^{B_2}) \qquad eq.1$$

$$P_{K-M}(D,T) = 1 - \exp\left[(-B_0 - B_1 * D^{B_2}) * T^{B_3}\right] \qquad eq.2$$

$$B_0, B_1, B_2, B_3 = regression\ coefficients$$

$$D = dose\ level \quad T = observation\ time$$

$$B_2 < 1 \rightarrow shape\ supralinear$$

$$B_2 = 1 \rightarrow shape\ linear$$

$$B_2 > 1 \rightarrow shape\ sublinear$$

EXAMPLES TO ILLUSTRATE WHEN THE DOSE DESCRIPTOR BMD05 SHOULD BE USED IN ADDITION TO THE DOSE DESCRIPTOR T25.

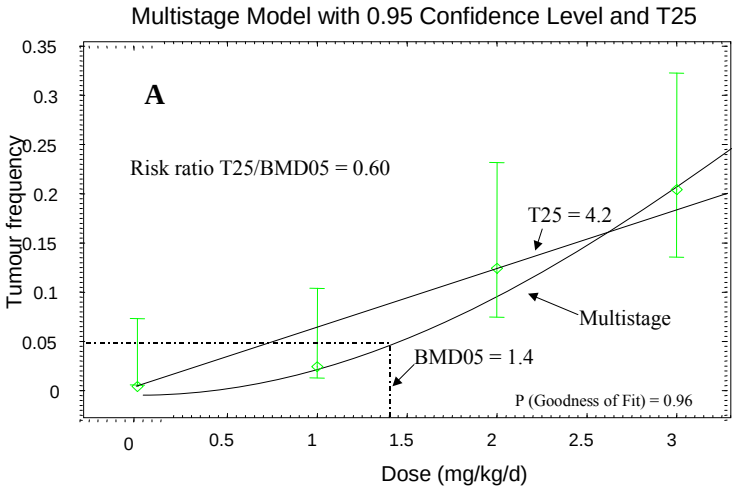
A comparison of the T25 and BMD05 by Crump et al. (2000) showed that in case of linear or close to linear dose response relationships the results of the two procedures are virtually identical. For sub- or supralinear dose response relationships estimates by the T25 method were at maximum two-fold higher or lower than the BMD05 method, respectively. In agreement with this, when the T25 method for quantitative risk characterisation was compared with the linearised multistage model or the LED10 method using the same data sets, the risks differed in most cases only slightly. Thus, when calculating the ratios of the risks obtained with the T25 and the other methods and plotting the frequency of the ratios, the means and medians were 1.2 and the 5% and 95% percentile equal to 0.50 and 2.0 (Sanner et al., 2001).

In the following examples, T25 and BMD05 have been calculated from the same data set. In addition, the p-values for Goodness of Fit and the risk ratios T25/BMD05 are given. The risk ratio T25/BMD05 is calculated from the formula: T25/(BMD05 x 5). The human lifetime daily dose representing a certain lifetime cancer risk as determined with the T25 method divided with the "risk ratio" will represent the corresponding human lifetime daily dose as determined with the BMD05 method. Thus, if the risk ratio is >1, it imply that the dose representing a certain risk determined with the BMD05 method would be less than that determined with the T25 method.

When considering the risk ratio it is important to remember that the variation in doses given a certain tumour response between different experiments even in the same laboratory with the same strain of animals may vary with a factor between 2 and 4.

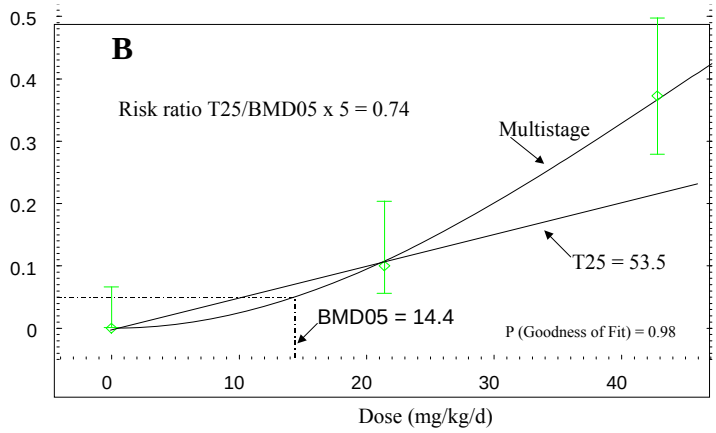
Example R. 8-3 Illustration A

0	MG/KG/D	0/50
1	MG/KG/D	1/50
2	MG/KG/D	6/50
3	MG/KG/D	10/50



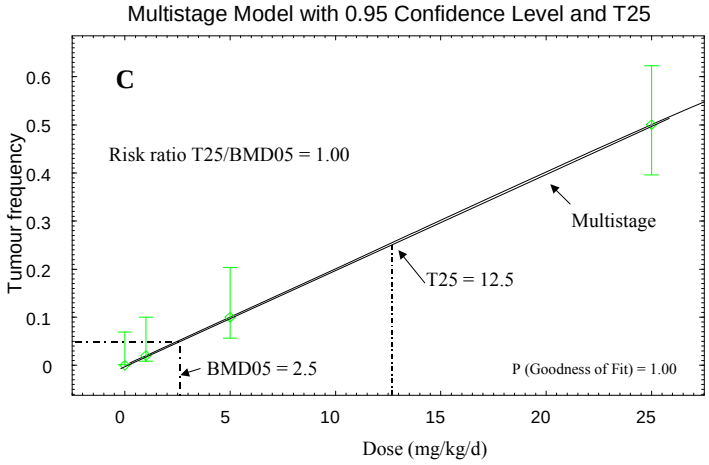
Example R. 8-4 Illustration B

0	MG/KG/D	0/52
21.4	MG/KG/D	5/50
42.8	MG/KG/D	19/51



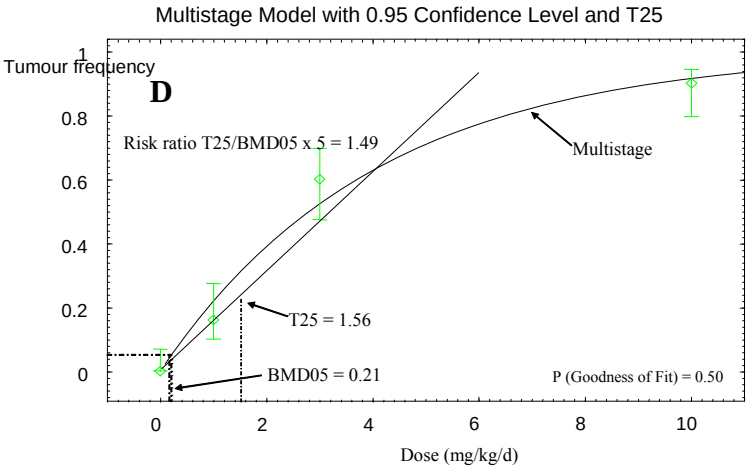
Example R. 8-5 Illustration C

0	MG/KG/D	0/50
1	MG/KG/D	1/50
5	MG/KG/D	5/50
25	MG/KG/D	25/50



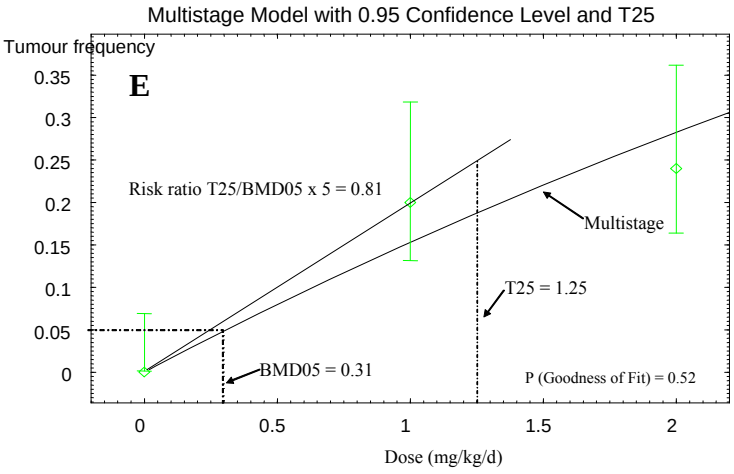
Example R. 8-6 Illustration D

0	MG/KG/D	0/50
1	MG/KG/D	8/50
3	MG/KG/D	30/50
10	MG/KG/D	45/50



Example R. 8-7 Illustration E

0	MG/KG/D	0/50
1	MG/KG/D	10/50
2	MG/KG/D	12/50



Illustrations A and B represent sublinear dose-response relationships. In illustration A the risk ratio is 0.60 indicating a 40% difference in the two methods. In illustration B only 2 dose points were available. The difference between the two methods is negligible.

Illustration C represents a close to linear dose relationship. The two methods give identical results.

Illustrations D and E represent supralinear dose-response relationships. In illustration D the difference between the two methods is 50%. It should be noted that the low dose point falls below the calculated curve, if the response had been larger (12/50) giving a frequency falling on the curve in illustration D, the risk ratio would decrease to 1.09 and the p-value for Goodness of Fit increase to 0.75. In illustration E only 2 dose points were available. The difference between the two methods is negligible.

The present exercise support the view presented in the main text that the BMD05 may offer an advantage in the case of sub- and supralinear dose relationship if a good set of quantitative data are available, i.e. a control population and at least three dose levels. In cases when the model shows a good fit, it is recommended that the point estimate of the BMD05 is calculated in addition to the T25. In cases where the evidence of sub- and supralinear dose-response is convincingly demonstrated also by other types of evidence in the dose-range studied e.g. formaldehyde, the BMD05 method should represent the method of chose.

APPENDIX R. 8-7 Derivation of a DMEL for Non-Threshold Carcinogens: Comparison of the "linearised" and the "large assessment factor" approach

Introduction

For the purpose of comparing the two methods of DMEL derivation (linearised approach and large assessment factor approach) it is not considered obligatory to calculate the relevant dose descriptors (BMD, T25) on a substance-specific basis. For the following examples a T25 of 10 mg/kg/d (oral rat carcinogenicity study) is assumed. The BMD10 equivalent to the T25 would be 4 mg/kg/d (T25 / 2.5 = 4). For the following exercise a two-fold less critical BMD10 of 8 mg/kg/d, and a BMDL10 of 2 mg/kg/d is assumed.

BMDL10 = 2 mg/kg/d

T25 / 2.5 = 4 mg/kg/d

BMD10 = 8 mg/kg/d

T25 = 10 mg/kg/d

Enclosed you will find the following examples:

	Type of Exposure Pattern	Relevant dose descriptor
(A)	General Population (oral exposure)	T25
(B)	General Population (oral exposure)	BMD10 resp. BMDL10
(C)	Workers (exposure by inhalation) Current sequence of extrapolation	T25
(D)	Workers (exposure by inhalation) Modified sequence of extrapolation ($10m^3$-approach)	T25

(A)

GENERAL POPULATION (ORAL EXPOSURE)

RELEVANT DOSE DESCRIPTOR: T25

Step 1: **Derivation / Identification of the relevant dose descriptor for carcinogenicity**		
Based on the results of an oral rat carcinogenicity study a T25 of 10 mg/kg/d is assumed.		
	"Linearised" approach	**"Large assessment factor" approach**
Relevant Dose descriptor	**T25 (rat, oral)** **10 mg/kg/d**	**T25 (rat, oral)** **10 mg/kg/d**

Step 2: **Modification of the relevant dose descriptor**		
	"Linearised" approach	**"Large assessment factor" approach**
For this scenario (general population, oral exposure) there is no need for a modification factor	1	1
Corrected Dose Descriptor	**Corrected T25** **10 mg/kg/d**	**Corrected T25** **10 mg/kg/d**

Step 3: Application of assessment factors to get the DMEL		
	"Linearised" approach	**"Large assessment factor" approach**
Interspecies extrapolation For the "linearity" approach only the allometric scaling factor of 4 is applied	4	10
Intraspecies extrapolation	not applied	10
Nature of the carcinogenic process	not applied	10
Point of comparison	not applied	10
2.5 in cases where the T25 is used instead of the BMDL10 (EFSA draft 07.04.2006)	not applied	2.5
High to low dose extrapolation	25,000 (linearity, 1:100,000) 250,000 (linearity, 1:1.000.000)	not applied
Calculation of DMEL (corrected T25 divided by overall assessment factor)	10 mg/kg/d / (4 * 25,000) = 0.0001 mg/kg/d 10 mg/kg/d / (4 * 250,000) = 0.00001 mg/kg/d	10 mg/kg/d / 25,000 = 0.0004 mg/kg/d
DMEL (based on T25) **associated with a lifetime cancer risk of very low concern**	**0.1 µg/kg/d (linearity, 1:100,000)** **0.01 µg/kg/d (linearity, 1:1.000.000)**	**0.4 µg/kg/d** **("Large assessment factor" approach)**
Overall quality of DMEL	**To be described narratively**	**To be described narratively**

(B)

GENERAL POPULATION (ORAL EXPOSURE)

RELEVANT DOSE DESCRIPTOR: BMD10 and BMDL10

Step 1: **Derivation / Identification of the relevant dose descriptor for carcinogenicity**		
Based on the results of an oral rat carcinogenicity study a BMD10 of 8 mg/kg/d and a BMDL10 of 2 mg/kg/d is assumed.		
	"Linearised" approach	**"Large assessment factor" approach**
Relevant Dose descriptor So far, for the "linearity" method it is proposed to use the BMD, while the EFSA method uses the BMDL.	**BMD10 (rat, oral)** **8 mg/kg/d**	**BMDL10 (rat, oral)** **2 mg/kg/d**

Step 2: **Modification of the relevant dose descriptor**		
	"Linearised" approach	**"Large assessment factor" approach**
For this scenario (general population, oral exposure) there is no need for a modification factor	1	1
Corrected Dose Descriptor	**Corrected BMD10** **8 mg/kg/d**	**Corrected BMDL10** **2 mg/kg/d**

Step 3: Application of assessment factors to get the DMEL		
	"Linearised" approach	**"Large assessment factor" approach**
Interspecies extrapolation For the "linearity" approach only the allometric scaling factor of 4 is applied	4	10
Intraspecies extrapolation	not applied	10
Nature of the carcinogenic process	not applied	10
Point of comparison	not applied	10
2.5 in cases where the T25 is used instead of the BMDL10 (EFSA draft 07.04.2006)	not applied	not applied
High to low dose extrapolation	10,000 (linearity, 1:100,000) 100,000 (linearity, 1:1.000.000)	not applied
Calculation of DMEL (corrected BMD/BMDL divided by overall assessment factor)	8 mg/kg/d / (4 * 10,000) = 0.0002 mg/kg/d 8 mg/kg/d / (4 * 100,000) = 0.00002 mg/kg/d	2 mg/kg/d / 10,000 = 0.0002 mg/kg/d
DMEL (based on BMD/BMDL) **associated with a lifetime cancer risk of very low concern**	**0.2 µg/kg/d (linearity, 1:100,000)** **0.02 µg/kg/d (linearity, 1:1.000.000)**	**0.2 µg/kg/d** **("Large assessment factor" approach)**
Overall quality of DMEL	**To be described narratively**	**To be described narratively**

(C)

WORKERS (EXPOSURE BY INHALATION)

RELEVANT DOSE DESCRIPTOR: T25

DEFAULT SEQUENCE OF EXTRAPOLATION

Step 1: **Derivation / Identification of the relevant dose descriptor for carcinogenicity**		
Based on the results of an oral rat carcinogenicity study a T25 of 10 mg/kg/d is assumed.		
	"Linearised" approach	**"Large assessment factor" approach**
Relevant Dose descriptor	**T25 (rat, oral)** **10 mg/kg/d**	**T25 (rat, oral)** **10 mg/kg/d**

Step 2: **Modification of the relevant dose descriptor**		
	"Linearised" approach	**"Large assessment factor" approach**
Route-specific bioavailability: 50% oral absorption 100% absorption by inhalation	50 / 100	50/100
Adjustment of route of exposure: from rat (oral) in mg/kg/d to rat inhalation (0.8l/min/kg, 8h): 0.384 m^3/kg/8h	1 / 0.384	1 / 0.384
Activity-driven differences: At rest / light activity: 6.7 /10 in line with the „10 m^3" approach	6.7 / 10	6.7 / 10
Differences between occupational and lifetime exposure conditions 7/5 * 52 /48 * 75 / 40 = 2.8	2.8	2.8

Calculation of modified dose descriptor	T 25 of 10 mg/kg/d multiplied by 50/100 * 1/0.384 * 6.7/10 * 2.8 = 24.4 mg/m³	T 25 of 10 mg/kg/d multiplied by 50/100 * 1/0.384 * 6.7/10 * 2.8 = 24.4 mg/m³
Corrected Dose Descriptor	**Corrected T25** **24.4 mg/m³**	**Corrected T25** **24.4 mg/m³**

Step 3: **Application of assessment factors to get the DMEL**		
	"Linearised" approach	**"Large assessment factor" approach**
Interspecies extrapolation Allometric scaling implicitly taken into account	1	2.5
Intraspecies extrapolation	not applied	5
Nature of the carcinogenic process	not applied	10
Point of comparison	not applied	10
2.5 in cases where the T25 is used instead of the BMDL10 (EFSA draft 07.04.2006)	not applied	2.5
High to low dose extrapolation	25,000 (linearity and 1:100.000) 5,000 (linearity and 5:100.000)	not applied
Calculation of DMEL (corrected T25 divided by overall assessment factor)	24.4 / 25,000 = 0.001 mg/m³ 24.4 / 5,000 = 0.005 mg/m³	24.4 / 3.125 = 0.008 mg/m³
DMEL (based on T25) **associated with a lifetime cancer risk of very low concern**	**1 µg/m³ (1:100.000, linear)** **5 µg/m³ (5:100.000, linear)**	**8 µg/m³** **("Large assessment factor" approach)**
Overall quality of DMEL	**To be described narratively**	**To be described narratively**

(D)

WORKERS (EXPOSURE BY INHALATION)

RELEVANT DOSE DESCRIPTOR: T25

MODIFIED SEQUENCE OF EXTRAPOLATION ("10 m³ APPROACH")

Step 1: **Derivation / Identification of the relevant dose descriptor for carcinogenicity**		
Based on the results of an oral rat carcinogenicity study a T25 of 10 mg/kg/d is assumed.		
	"Linearised" approach	**"Large assessment factor" approach**
Relevant Dose descriptor	**T25 (rat, oral)** **10 mg/kg/d**	**T25 (rat, oral)** **10 mg/kg/d**

Step 2: **Modification of the relevant dose descriptor**		
	"Linearised" approach	**"Large assessment factor" approach**
Route-specific bioavailability: 50% oral absorption 100% absorption by inhalation	50 / 100	50/100
Body weight of 70 kg	70	70
Breathing volume of 10 m³ for workers (light activity, 8 hours)	1 / 10	1 / 10
Differences between occupational and lifetime exposure conditions 7/5 * 52 /48 * 75 / 40 = 2.8	2.8	2.8
Calculation of corrected dose descriptor	T 25 of 10 mg/kg/d multiplied by 50/100 * 70 * 1/10 * 2.8	T 25 of 10 mg/kg/d multiplied by 50/100 * 70 * 1/10 * 2.8

	= 98 mg/m³	= 98 mg/m³
Corrected Dose Descriptor	**Corrected T25** **98 mg/m³**	**Corrected T25** **98 mg/m³**

Step 3: **Application of assessment factors to get the DMEL**		
	"Linearised" approach	**"Large assessment factor" approach**
Interspecies extrapolation Allometric scaling implicitly taken into account	4	10
Intraspecies extrapolation	not applied	5
Nature of the carcinogenic process	not applied	10
Point of comparison	not applied	10
2.5 in cases where the T25 is used instead of the BMDL10 (EFSA draft 07.04.2006)	not applied	2.5
High to low dose extrapolation	25,000 (linearity and 1:100.000) 5,000 (linearity and 5:100.000)	not applied
Calculation of DMEL (corrected T25 divided by overall assessment factor)	98 / 100,000 = 0.001 mg/m³ 98 / 20,000 = 0.005 mg/m³	98 / 12,500 = 0.008 mg/m³
DMEL (based on T25) **associated with a lifetime cancer risk of very low concern**	**1 µg/m³ (1:100.000, linear)** **5 µg/m³ (5:100.000, linear)**	**8 µg/m³** **("Large assessment factor" approach)**
Overall quality of DMEL	**To be described narratively**	**To be described narratively**

APPENDIX R. 8-8 Acute toxicity

In principle, the DNEL for acute toxicity is derived in the same way as the DNEL for repeated-dose toxicity, although special consideration needs to be given to the nature of the endpoint and the type and quality of the data available. Acute toxicity includes effects which occur after a single exposure, and those effects may be transient and reversible (e.g. narcosis) or may be irreversible (e.g. irreversible organ damage or possibly effects on the developing foetus). They may appear immediately at the exposure, or after some time. As well as acute toxicity studies, the findings from toxicity studies with repeated dosing (e.g. repeated-dose toxicity studies, reproductive or developmental toxicity studies) need to be considered to ensure that all the possible consequences of acute exposures are identified.

The establishment of an acute toxicity DNEL set for effects occurring after a single exposure of a few minutes up to 24 hours is not only cumbersome (there is no established accepted methodology) and resource-intensive but probably unnecessary, as the long-term DNEL is normally sufficient to ensure that these effects do not occur. It is therefore proposed that if an acute toxicity hazard (leading to C&L) has been identified, a DNEL for acute toxicity is only established for the effects of peak exposures as these peaks can be significantly higher than the average daily exposure and the long-term DNEL may be insufficient to limit them. Overall, therefore, a DNEL for acute toxicity should be derived if an acute toxicity hazard (leading to C&L) has been identified and there is a potential for high peak exposures. High peak exposures are usually assessed for the inhalation route only, so this guidance outlines how to set acute toxicity DNELs for the inhalation route. Although, peak exposures in theory may also occur for the dermal and oral routes, these are not normally assessed, so the establishment of acute toxicity DNELs for dermal and oral peak exposures appears superfluous. However, on a case-by-case basis an ‘acute’ dermal or oral DNEL can be set for comparison with single exposure events. No detailed guidance is given for setting oral or dermal acute toxicity DNELs, but the principles are the same as those described below for setting inhalation acute toxicity DNELs. The registrant needs to use expert judgement in setting acute toxicity DNELs, and provide justification for why they are needed and how they have been derived.

It should be stressed that the actual daily dose is *independent* of the exposure frequency. This means that if for a certain scenario, worker or consumer exposure is for instance only for a number of days per year, the exposure value is the actual dose on the exposure days, and not the daily dose averaged out (and thus divided!) over the whole year.

Identification of the typical dose descriptor

Traditionally, acute toxicity tests in animals have primarily used mortality as the main observational endpoint, usually in order to derive a $LC(D)_{50}$ value. In many cases there will be little information available on the cause of death or mechanism underlying the toxicity, and only limited information on pathological changes in specific tissues or clinical signs, such as changes in the activity or behaviour of the animals. Using mortality as a starting point for derivation of a DNEL ignores the possibility of sub-lethal, but still potentially serious, toxicity at lower doses. Indeed, mortality is often the most severe expression in a continuum of toxicity with a common underlying mechanism. For example, lethality induced by solvents is the severe manifestation of CNS depression, which would be observed through clinical signs at much lower doses. Ideally, the sub-lethal toxicity occurring at lower doses should be considered as a more rational starting point than mortality for derivation of a DNEL.

Many acute toxicity studies on chemicals of low toxicity are performed as limit tests, with only a single dose, and provide very limited information on toxicity. In studies with more toxic chemicals several dose levels may have been used and therefore will provide more information on the dose-response relationship at lower dose levels. The clinical and pathological observations at the lower end of the dose-response curve may be particularly useful in identifying NOAEC(L)s or LOAEC(L)s for toxicity, other than mortality. For systemic toxicity, there could be some evidence of target organ toxicity if pathological findings are documented, or signs of toxicity from clinical findings. For instance, a decrease in activity or a change in normal behaviour may be indicative of CNS depression.

Findings from repeated-dose studies with more detailed investigations and acute effects observed in other studies (including in vivo mutagenicity tests, neurotoxicity studies or other specialised studies) may provide additional relevant information to help refine the N(L)OAEC from acute toxicity studies. Repeated dose studies may even provide a N(L)OAEC for effects due to acute exposure. For example, mortalities and signs of distress occurring soon after onset of exposure of a repeated dose study should be interpreted as acute toxicity and could be used as the basis for identifying a N(L)OAEC for acute toxicity. However, if the N(L)OAEC from acute toxicity studies is considered to be unreliable or uncertain due to limited reporting of findings and if no N(L)OAEC for effects due to acute toxicity can be identified from a repeated dose study, then the use of a N(L)OAEC for effects due to repeated exposure may be more appropriate, in particular if short-term repeated-dose studies (e.g. subacute studies, range-finding studies, developmental toxicity studies, etc.) are available. One possible exception to the use of a N(L)OAEC for effects due to repeated exposure to assess the effects of single exposures is for chemicals that bioaccumulate or cause enzyme induction on repeated exposure as in these cases such a N(L)OAEC may be of no relevance to acute toxicity.

Human evidence, such as epidemiological studies, case reports of poisoning or episodes of acute toxicity at work, or information from medical surveillance, can be very important for the assessment of acute toxicity and can provide evidence of effects that are undetectable in animal studies, for instance induction of symptoms such as headaches and nausea. There may be case-reports of human poisoning incidents, which are usually single-exposure events, either deliberate ingestion or during incidents/accidents. The reliability of exposure assessments in such reports needs careful consideration as there is often substantial uncertainty, but these data may give valuable information on the acute toxicity in humans, allow the identification of human NOAEC(L) or LOAEC(L) values and give some indication of the relative sensitivity between humans and animals.

In addition to acute systemic effects, some substances may cause local effects on the respiratory tract following a single exposure via the inhalation route. Acute local effects on the respiratory tract could be due to either or both of two different toxicological phenomena: sensory irritation or cytotoxicity/tissue damage. Only the derivation of a DNEL for acute cytotoxicity on the respiratory tract will be dealt with under this endpoint. The derivation of a DNEL for sensory irritation will be dealt with the endpoint of respiratory tract irritation. For acute cytotoxicity on the respiratory tract, the severity of the local effects is usually proportional to the concentration/dose level; in such a situation, therefore it may be possible to identify a NOAEC or LOAEC for these effects from pathology or clinical observations from either animal studies or human data.

However, there are cases where the available data are too poor (e.g. only an LC_{50} is available with no data to inform on the magnitude of the severity factor or only an LD_{50} for a different route of exposure is available) to allow setting a sufficiently robust DNEL. For these cases, and for substances being acutely very toxic (T+), a qualitative risk characterisation is recommended (see below).

Use of alternative data

If chemical-specific testing data for assessing acute toxicity are not available, the use of alternative data (in vitro data, (Q)SAR, read-across) should be considered.

Currently there are no validated in vitro methods for acute toxicity, although several cytotoxicity assays are undergoing a validation process and may be useful for range-finding and prioritisation purposes. These in vitro assays are not suitable for quantitative purposes on their own, but they could provide supplementary information, useful in the interpretation of in vivo studies.

For some specific effects (e.g. CNS depression) it might be also possible to obtain a (Q)SAR prediction from a validated model. However a QSAR prediction within the applicability domain from a validated QSAR model for acute toxicity may only be suitable for hazard identification and normally not for quantitative purposes.

Read-across from one or more structural analogues could be an option. If the use of read-across is considered appropriate, it should then be possible to select an equivalent starting point, i.e. a NOAEC(L) or LOAEC(L) for sub-lethal effects or an $LC(D)_{50}$ value.

Modification of the dose descriptor

The potential for peak exposures is of most concern for workers, but it can also be of importance for consumers. Hence, the exposure assessment should always consider the possibility for such peak exposures, particularly for the inhalation route of exposure, as these peaks could potentially be significantly above the average daily exposure level. As the long-term DNEL is normally derived for a daily reference period (e.g., 8 hours/day) rather than for a cumulative period, and as the degree/magnitude of the repeated dose toxicity response normally depends on the degree/magnitude of the acute toxicity response, the derivation of an acute toxicity DNEL for the same reference period as that of the long-term DNEL is not pragmatic. Thus, if a DNEL for acute toxicity needs to be established, as determined based on the toxicological profile of the chemical concerned, this should be derived only for a specified fraction of the daily exposure duration (usually 15 minutes for workers).

It should be noted that, as the reference period of the acute toxicity DNEL (e.g. 15 minutes) is likely to differ from the exposure duration in the experimental (animal or human) study from which the N(L)OAEC or $LC(D)_{50}$ was identified, the derivation of such a DNEL, particularly for the inhalation route of exposure, might involve time scaling. Before correcting the starting point to account for time extrapolation, on a case-by-case basis it should be judged whether this is appropriate. For example, if the relevant effect is deemed to be more concentration rather than dose dependent (which is not always the default position to take for local cytotoxic effects), then the duration of exposure is likely to be of little consequence, and hence, time extrapolation would be inappropriate. If the starting point was identified from a repeated-dose toxicity study, involving exposure for several hours per day for several days, time scaling is relevant if the effect occurs soon after onset of the exposure. However, it might be considered that time scaling is inappropriate if the effect arise after days of exposure as this would involve extrapolation 2- to 3-fold orders of magnitude below the experimental exposure period, leading to a highly uncertain estimate.

If time extrapolation is considered valid, then the most appropriate approach is to make use of the modified Haber's law ($C^n \times t = k$, where 'C' is the concentration, 'n' is a regression coefficient, 't' is the exposure time and 'k' is a constant) according to which the relationship between exposure concentration and exposure duration for a specific effect is exponential.

In order to estimate the value of the exponent n, empirical exposure concentration-exposure duration relationships for the relevant effect, which require the availability of good quality studies with several exposure durations, need to be established. In the absence of suitable data for deriving n, a default value of n=1 for extrapolating from shorter to longer exposure durations and a default value of n=3 for extrapolating from longer to shorter exposure durations should be used as these values lead to the most conservative estimates. These defaults are consistent with those laid out in US guidance on setting emergency standards for major accident hazards (US NRC, 2001) and are based on the observation that n lies in a range of 1 to 3 from an analysis of approximately 20 structurally diverse chemicals with established concentration-time relationships for lethality (ten Berge et al., 1986).

In addition to time scaling, it might be also necessary to convert the dose descriptor into a correct starting point to take account of differences in routes of exposure between experimental animals and humans (in the exceptional case of route-to-route extrapolation), possible differences in absorption between routes and between experimental animals and humans and potential differences in respiratory volumes between experimental animals (usually at rest) and humans (light activity in case of workers) as described in the general guidance (Section R.8.4.2).

Application of assessment factors to the correct starting point to obtain the acute-toxicity-specific DNEL

In general, the investigation of toxicity is less extensive and detailed in acute toxicity studies compared to repeated-dose studies and reporting is in most cases limited to overt signs of toxicity (i.e. clinical signs) because histopathology, clinical chemistry, urinalysis, haematology and detailed motor activity are not normally performed. In view of this, one should consider the possibility that a lower NOAEC(L) or LOAEC(L) would have been determined if more detailed histopathology etc had been conducted. Overall, therefore, when a NOAEC(L) or LOAEC(L) from acute toxicity studies is used as the starting point for the derivation of the DNEL for acute toxicity, careful consideration should be given as to whether an additional (over and above those described in the general guidance) assessment factor should be applied to account for these deficiencies.

When only an $LC(D)_{50}$ value is available or when all the dose/concentration levels tested produced mortality, there is substantial uncertainty regarding the toxicity at lower doses and no reliable basis to judge a dose which would not cause any toxicity in humans. If this is the case, the whole toxicological database should be considered to try to make a reasonable prediction of what the toxicity will be like at lower doses and a rough estimate of the likely threshold. Using this information it might be possible to determine the size of an additional severity factor to be applied to the $LC(D)_{50}$ value in the derivation of the DNEL to cover the significant inherent uncertainties.

It should be emphasized that some of the current standard protocols, such as the fixed-dose procedures and the acute toxic class methods, use evident or severe toxicity (clinical signs of toxicity indicating the moribund condition or severe pain of the test animals) in place of mortality and identify the discriminating dose (dose causing evident toxicity) rather than a specified lethal dose. The discriminating dose may be seen as a severe LOAEL, thus requiring a somewhat (depending on the severity of the effects) lower AF than if a LD(C)50-value had been the starting point (see box 4 and 5 of the flow chart below).

If no information on thresholds for acute toxicity is available, two options are envisaged. It may be concluded, based on exposure-driven considerations (REACH Annex VI, see step 4), that more data are required to address the inherent uncertainties regarding the risks of acute toxicity. Alternatively, a practical approach, which has a long tradition in the occupational health arena, could be considered.

Depending on the steepness of the dose-response curve for the repeated dose effects, the DNEL for acute toxicity could be set for a reference period of 15 minutes at 1-5 times the value (default 3) of the long-term DNEL. The steeper the dose-response relationship, the smaller the multiplying factor. This approach is particularly valid if similar mechanisms of actions may be involved in the responses to single and repeated exposure, but can be used as a precautionary approach also in other cases. Although it is recommended that a qualitative RC should be performed for very toxic, (T+)-labelled, substances (see box 8 in Figure R. 8-5), aiming at avoiding/minimising exposure, an indicative acute toxicity DNEL may also be set for these substances based on this approach to assist in the performance of the qualitative RC. It should be noted that for very toxic substances, if this approach is adhered to, the acute toxicity DNEL should be set at only 1, max 2 times the value of the long-term DNEL, as it will result in a more conservative value.

In relation to acute local cytotoxic effects on the respiratory tract, the DNEL should be derived by applying the same assessment factors as described for local effects in the general guidance.

If the starting point has been derived by using read-across from one or more structural analogues, the additional uncertainty deriving from using these data may be addressed by selecting an additional assessment factor.

When it comes to the final (total) assessment factor, it may become large if the available data introduces many uncertainties into the assessment. In that case, the registrant may choose to perform a qualitative assessment and to introduce relevant risk management measures.

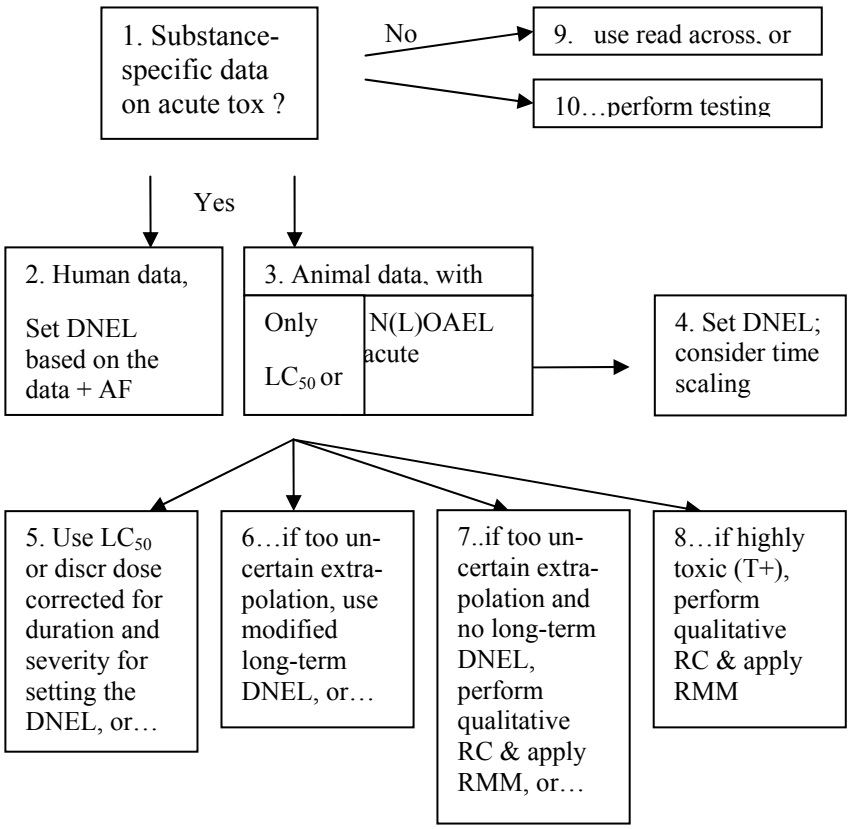
In summary, for some substances, notably substances for which an acute toxicity hazard (leading to C&L) has been identified and for which the exposure assessment (the tentative exposure scenario) has predicted high peaks (because of, e.g., high volatility or specific use patterns), the long-term DNELs may not ensure a sufficient level of protection after peak exposure. Particular account should be taken of health effects which are not of the same type as those which drive the long-term DNEL. Still, all of the available evidence and not only the acute toxicity studies should be used to determine the most appropriate toxicological effect on which to base the derivation of the DNEL for acute toxicity. As a rule of thumb, a $DNEL_{acute}$ should be set for acutely toxic substances if actual peak exposure levels significantly exceed the long-term DNEL. For such cases, a $DNEL_{acute}$ need to be set and assessed in relation to the peak exposure levels that humans may experience.

In general, the ideal starting point for the derivation of the acute toxicity DNEL should be the NOAEC(L) or LOAEC(L) for sub-lethal effects, such as local irritation (e.g., respiratory irritation caused by cytotoxicity) or CNS depression (if identified). The starting point might have to be modified to correct for time scaling, potential differences in routes of exposure between experimental animals and humans (in case of route-to-route extrapolation), possible differences in absorption between routes and between experimental animals and humans and potential differences in respiratory volumes between experimental animals (usually at rest) and humans (light activity in case of workers). The selection of the assessment factors to be applied to the starting point should be performed in the usual way, through a systematic consideration of all the uncertainties in the effect assessment.

A schematic representation of how to derive the acute toxicity DNEL for different dose descriptors and a flow-chart describing the decision tree that should be followed when deriving the acute toxicity DNEL are given below (Table R. 8-5 and Table R. 8-20). However, depending on the available data, it may not always be possible to set an acute toxicity DNEL. In those cases a qualitative approach can be chosen, as described in boxes 7-8 of the decision-tree.

In the explanatory text corresponding to the boxes in the decision-tree figure, it has been assumed that it is an inhalation DNEL that is to be set.

Figure R. 8-5 Decision tree for setting an acute inhalation toxicity DNEL.



Box 1 Is there substance-specific data on acute toxicity available for the substance? If human data are available, proceed to box 2. Acute oral or inhalation animal toxicity data should be present already from the 1 t/y level, usually allowing setting of an acute toxicity DNEL based on animal data (box 3). Still, all available data should be considered when setting the acute toxicity DNEL. If no substance-specific data are available, consider if there are other data that can be used for setting the acute toxicity DNEL (box 9), or consider performing a relevant acute toxicity study (box 10).

Box 2 In some cases there may exist human data (e.g., occupational experience of CNS depression, epidemiology, case studies, or reports from poison centres) on the toxicity of the substance that will allow setting a DNEL. These data are generally surrounded by high uncertainties, for instance because of unclear exposure-situations or co-exposure to other chemicals. However, after a case-by-case evaluation of the data, they can sometimes still be used. In a first step, the dose descriptor may need to be time-scaled (by the modified Haber's law). In a second step, an AF for intra-species variation is usually needed, when setting the $DNEL_{acute}$.

Box 3 Acute animal toxicity data are available. Assess whether a N(L)OAEC for sub-lethal effects can be set, which would allow proceeding to box 4 and setting a reliable $DNEL_{acute}$. If only data on lethality is available, proceed to box 5, 6, 7, or 8.

Box 4 If the acute animal inhalation study (normally 4 hours exposure time) includes dose-response relationships for relevant sub-lethal effects, a N(L)OAEC for these effects may be identified. In a first step, the N(L)OAEC is corrected to a dose descriptor representing 15 minutes exposure using the modified Haber's law. There is usually also a need to correct for differences in respiratory volumes between experimental animals (usually at rest) and (working) humans. In the next step, appropriate AFs are used when setting the DNEL, i.e., AFs for inter- and intra-species variation, as well as, if needed, an AF for LOAEC to NOAEC extrapolation. If only a limited number of end-points have been examined in the acute toxicity study (e.g., not the relevant effect, if known), then an additional assessment factor to account for the inherent deficiencies of the study should be applied when deriving the $DNEL_{acute}$.

Acute systemic toxicity data may sometimes only be available for other routes of exposure (e.g., oral) than the relevant route of human exposure. If the oral acute toxicity study includes dose-response relationships for sub-lethal effects, then an oral N(L)OAEL for these effects may be identified. The oral N(L)OAEL could be modified into an inhalation N(L)OAEC using route-to-route extrapolation (Section 8.4.2). However, it is noted that this procedure introduces significant uncertainties especially in relation to what inhalation time-frame this extrapolated N(L)OAEC would represent, and the procedure is therefore discouraged.

Box 5 An inhalatory $DNEL_{acute}$ is derived by applying a large assessment factor for severity of effect to the (time-scaled*, if needed) LC_{50}-value. There is no scientific basis for a default value of this assessment factor (for the extrapolation of a lethal concentration into a NOAEC). Still, a default AF of 100 is suggested as a starting point, but the factor can be modified in light of the whole toxicity database for that substance provided that a sound justification is given. In addition, AFs for inter-and intra-species variation may be needed when setting the DNEL. The approach and the resulting DNEL is uncertain, and it should be acknowledged in the risk characterisation. It should have high priority to revise the DNEL when better data become available.

If acute systemic toxicity data are only available for other routes of exposure (e.g., oral), and the dose-descriptor is a LD_{50}-value, a DNEL can in theory be calculated using R-t-R-extrapolation and application of AFs (as above). However, such a DNEL would generally be so uncertain, that a quantitative risk characterisation is questionable. Rather, a qualitative risk characterisation is recommended (see box 7). In the qualitative risk characterisation, the potency of the substance (as judged by its C&L) should be considered when deciding on the RMM/OCs needed to ensure control of risks (see Section E.3.4).

Box 6 If the setting of an acute DNEL based on the lethality data is considered to involve too large uncertainties, a $DNEL_{acute}$ can be set based on the long-term inhalation DNEL (if available). The long-term DNEL shall then be modified by multiplying with a factor of 1-5, where the size of this factor depends on the potency and the dose-response curve of the substance. In spite of scientific uncertainties with this approach (e.g., that different target organs/systems may exist after long-term versus peak exposure), it can be useful in the assessment of acute risks. This approach is not recommended if long-term data are only available for another route of exposure than the human exposure route for which the $DNEL_{acute}$ needs to be derived.

Box 7 If the derived inhalation $DNEL_{acute}$ is highly uncertain (e.g., because of R-t-R extrapolation + extrapolation from an LD_{50} value) and a long-term DNEL is not available, then a qualitative risk characterisation shall be considered. In the qualitative risk characterisation, the potency of the substance (as judged by its C&L) should be considered when deciding on the RMMs/OCs (see Section E.3.4) needed to ensure control of risk. Preferably, the RMM should ensure that peak concentrations exceeding the long-term DNEL will not occur. If the substance is used in consumer-available products, it must be made sure that the risks are controlled.

Box 8 For acutely very toxic substances (T+), a qualitative risk characterisation shall be considered, since the high toxicity is sufficient to warrant a strict control of any potential exposure (as any $DNEL_{acute}$ for these substances are likely to be very low). In the qualitative risk characterisation, the potency of the substance (as judged by its C&L) should be considered when deciding on strict RMM (e.g., specific PPE/OCs) needed to ensure control of risks (see Section E.3.4). Basically, the RMM should ensure that peak concentrations exceeding the long-term DNEL will not occur.

Box 9 It may also be possible to set a DNEL based on read-across from chemically related substances. Thus, if acute toxicity data are available for a substance(s) with a similar structure and physico-chemical properties (and toxicity profile, if such data are available) it can be used for setting the acute toxicity DNEL. The data may need time scaling[14], and there is a need for AFs for severity (if LC_{50} data), inter- and intra-species variation, and on a case by case basis for the uncertainties introduced by the read across as such.

Box 10 If there are no chemical-specific data and read-across is not possible, then testing is required.

The risks of acute toxicity for humans via the oral or dermal route should normally be assessed using the long-term DNEL. However, on a case-by-case basis, the risks of acute toxicity for humans via the oral route, and sometimes even the dermal route, may be assessed using specific acute DNELs also for these routes. This may apply if an acute toxicity hazard via the dermal or oral route (leading to C&L) has been identified and high (peak/single) exposures through these routes, exceeding the long-term DNEL, have been predicted to occur in the tentative ES. The approaches that can be used to derive these dermal or oral acute toxicity DNELs are in principal the same as those outlined above to derive inhalation acute toxicity DNELs.

Table R. 8-20 Derivation of the acute toxicity DNEL

Based on different types of substance-specific data. If instead read across is used to identify the dose descriptor, in addition to the steps described in the table below, an additional assessment factor to address the uncertainty deriving from using read-across (on top of those described below) should be considered.

Dose descriptor	Modification of the dose descriptor (over and above that described in the general guidance)	Selection of assessment factors (over and above those described in the general guidance)
N(L)OAEL from animal acute toxicity studies	Consider time scaling by applying the modified Haber's law	Consider additional factor to account for the inherent deficiencies of acute toxicity studies
N(L)OAEL from animal repeated dose toxicity studies	Time scaling might be inappropriate	
$LD(C)_{50}$ or discriminating	Consider time scaling by applying the	Apply additional severity factor;

[14] The N(L)OAEC is corrected to a dose descriptor representing 15 minutes exposure using the modified Haber's law. There is usually also a need to correct for differences in respiratory volumes between experimental animals (usually at rest) and (working) humans.

dose from animal studies	modified Haber's law	determine its magnitude by considering the whole tox database
N(L)OAEL from human data	Consider time scaling by applying the modified Haber's law	
N(L)OAEL for local cytotoxic effects on the respiratory tract	Consider time scaling by applying the modified Haber's law	

REFERENCES

US NRC. National Research Council 2001. Standing Operating Procedures for Developing Acute Exposure Guideline Levels for Hazardous Chemicals. National Academy Press, Washington, DC, USA.

ten Berge WF, Zwart A, Appelman LM (1986) Concentration-time mortality response relationship of irritant and systemically acting vapour and gases. J Hazardous Materials 13:301-309.

APPENDIX R. 8-9 Skin and eye irritation/corrosion and respiratory irritation

Several types of data can be used to assess the irritant or corrosive effects of a substance. However, DNELs for irritation/corrosion can only be derived if dose-response information is available. Acute, sub-acute or sub-chronic toxicity studies in animals by the dermal or inhalation route may be able to provide this information, provided that symptoms of irritation/corrosion are recorded and reported in relevant studies. Also human data, primarily from reliable evidence of symptoms caused by occupational exposures, may have dose-response information.

When no DNEL can be derived, a qualitative approach to assessing and controlling the risks is necessary. This can be the case when only the following types of data are available: pH, in vitro data on skin and eye irritation/corrosion, in vivo data with no information on a dose/response, or QSAR. From these types of data, only qualitative information, either showing or not showing an effect, can be obtained. Sometimes the potency categories on the irritation and corrosion can also be obtained (see Table R. 8-21 and Table R. 8-22).

The definition of irritation and corrosion are based on the scoring of signs as given in the relevant test guidelines (B.4 and B.5 of Annex V of Directive 67/548/EEC). In summary, according to the test guideline B.4, 0.5 ml of undiluted test substance is topically administered to the skin of three rabbits and kept semi-occluded for 4 hours. Thus, no dose/concentration-response relationship can be obtained from this test. After 24, 48 and 72 hours, the signs of irritation/corrosion are recorded. Grading of the signs of erythema echar formation and oedema (from 1 to 5) is given, as indicated in this test guideline. Similar guidance on the eye irritation/corrosion can be found in the test guideline B.5. The grading as described in test guidelines B.4 and B.5 is also the basis for classification according to Annex VI of Directive 67/548/EEC (need to be revisited when the EU Regulation implementing GHS enters into force).

A DNEL for irritation/corrosion can be obtained from an acute or repeated dose dermal study as described below, only when observed skin reactions in these tests are graded as described in guideline B4. Therby the criteria for DNEL derivation are consistent with the classification criteria. In practice, grading the skin responses at different dose levels would allow one to have a basis for setting a DNEL for irritation/corrosion.

Identification of the dose descriptor

Corrosive and irritant effects of substances on the skin and eye, as well as in the respiratory tract are usually concentration-dependent. Occasionally, it may be possible to derive a non-irritant dose/concentration (NOAEL/C) or the lowest irritant dose/concentration (LOAEL/C) of the substance concerned. This dose-descriptor can be used to derive an irritation-specific DNEL, which might be expressed as a dose or as a concentration. For consistency, the term dose descriptor is used in this chapter, although in many cases irritant/corrosive concentration has been found in the relevant studies. By this it follows that the exposure estimation should include information on the concentration of the substance relevant to each use pattern, to allow for comparison with a potential concentration-based DNEL. Further guidance on preparing the exposure assessment is given in the Part D).

The available data should be assessed to determine the applied concentration/dose that shows no irritant/corrosive effects on the skin, eye or respiratory tract, taking into account data from both humans and animals. Where the data are of sufficient quality to determine a no-effect concentration/dose (expressed e.g. concentration, percentage or mg/kg body weight), this

NOAEC/L shall be taken to be the concentration/dose descriptor for deriving the irritation-specific DNEL.

Where data are available from both well-documented human experience and from adequately reported animal studies, due account shall be taken of the human data and all other relevant information in deriving the DNEL. In particular, where it is known that experimental animals express a differing sensitivity to the corrosive/irritant effects of a substance than humans, the data from humans shall take precedence in deriving the DNEL.

When applicable (e.g. in case of corrosive substances), two dose descriptors should be identified: one for the irritant effect and another for the corrosive effect. Depending on what concentrations are used in different exposure situations, one may need either or both of these dose descriptors for setting DNELs. For example hydrogen peroxide above 5% is irritating to eyes, above 35% it is a skin irritant, whereas above 50%, it is corrosive to the skin.

Skin

A DNEL for skin irritation/corrosion may be derived if a NOAEL/C or a LOAEL/C can be identified from *dermal acute, sub-acute or sub-chronic studies in animals.* It should be noted that the observations from acute skin irritation tests are not always predictive for the occurrence or absence of skin irritation after repeated exposure (2).

In these studies information on the appearance of the skin and histopathological data may be available. The current test guidelines support extracting and reporting such data as they recommend e.g. that systematic and individual observation of effects should include changes in skin (3). According to the guidelines histopathological examination is made on normal and treated skin of the high dose group and controls and, when necessary, also in other groups (3).

It may be possible to observe irritation/corrosion in an *acute dermal toxicity study* as well. The relevant guideline indicates that cage-side observations should include changes in fur, eyes and mucous membranes, and also the respiratory system. Note that histopathology is not systematically made, as it is in sub-acute and sub-chronic studies.

If human data are available, a NOAEL/C or a LOAEL/C may be identified from these data.

The NOAEL/C for skin irritation/corrosion would be the highest dose/concentration that did not cause dermal irritation/corrosion in the relevant animal study or in the human cases/data. It is assumed that at the higher dose levels clear signs of skin irritation were observed. Note that since the relevant tests, i.e. acute, sub-acute and sub-chronic dermal toxicity studies are made to observe the systemic toxicity, the local effects on the skin may not always be reported in detail. If irritation or corrosion was observed in more than one study, the highest (relevant) NOAEL/C below the lowest (relevant) LOAEL/C should be selected.

Respiratory system

A distinction between the two mechanisms of respiratory irritation, cytotoxicity and sensory irritation should be considered, whenever possible. Cytotoxic effects may be observed in humans or in animal studies. Sensory irritation is experienced by humans and when reported in sufficient extent and detail, it can be the basis for the DNEL. Sensory irritation is generally regarded as a more sensitive endpoint than cytotoxicity, and consequently data on sensory irritation (if available) should be given priority. The inter-species AF of 2.5 proposed below aims to capture some of the uncertainty caused by this difference between animal and human data. It could be reminded that several of the current OELs are based on respiratory irritation and therefore this endpoint might be the leading human health effect in the risk characterisation.

In principle, sensory irritation can be observed also in animals as a decrease of the respiratory rate (Alarie assay). However, the quantitative use of these data is not generally accepted and therefore, data from this assay should be used on a case-by-case basis as an acute effect to derive a short-term (15 min) DNEL only.

For irritant cytotoxic effects on the respiratory tract it may be possible to derive from the available data (either in humans or in animals) a dose descriptor, i.e. a non-irritant concentration (NOAEC) or the lowest irritant concentration (LOAEC) (expressed either as ppm or mg/m^3) for setting the DNEL.

Where data are available from both well-documented human experience and from adequately reported animal studies, due account shall be taken of the human data and all other relevant information in deriving the DNEL. In particular, where it is known that experimental animals express a differing sensitivity to the corrosive/irritant effects of a substance than humans, the data from humans shall take precedence in deriving the DNEL.

If only animal data are available, a DNEL for irritation/corrosion of the respiratory tract may be derived from a NOAEC or a LOAEC which can occasionally be based on the acute, sub-acute or sub-chronic inhalation studies.

The NOAEC would be the highest concentration that did not cause respiratory irritation in the acute, sub-acute and sub-chronic inhalation toxicity study or in the human cases/data. It is assumed that at the higher dose levels clear signs of respiratory irritation were observed. If irritation was observed in more than one study, the highest (relevant) NOAEC below the lowest (relevant) LOAEC should be selected.

Eye

In acute, sub-acute and sub-chronic toxicity studies, the eyes of the animal are not intentionally exposed and the symptoms are not systematically reported. Therefore, normally there is no basis for quantitative assessment of the eye irritation/corrosion from these studies. In case signs of eye irritation/corrosion are observed in inhalation toxicity studies, and dose-response information is available, it may be possible to identify a NOAEC or a LOAEC and derive a DNEL. Also, if human data are available, a NOAEC or a LOAEC may be identified from these data.

Human data

Human data may sometimes by used for setting the irritant/corrosive *concentration* for skin or eye or irritant concentration in the respiratory tract. Concentration-response relations or threshold concentration in humans have been observed for some chemicals, for which there are sufficient evidence and case reports/clinical cases e.g. from occupational exposures. For some groups of chemicals, such as common solvents, peroxides and acids, the irritant concentration in a liquid or in the air has been characterised based on the human evidence. Also information provided by poison information centres can be useful.

Open literature and the company-based occupational health surveillance of the relevant industries should be used to find out if data is available, which enable characterisation of these effects in quantitative terms (e.g. Medical Surveillance reviews of U.S Department of Labour (4) and Chemical Hazards review UK Health Protection Agency (5)).

Relationship of the DNEL to substance specific limits for classification

In the harmonised classification and labelling, substance-specific provisions (i.e., specific concentration limits) for irritancy/corrosivity to the skin and/or eyes, based on human data, might exist for the substance of concern (6). In these cases, it is obvious that there are data available allowing setting a N(L)OAEL. The registrant can either use the old assessments, or, if having access to this data, perform a new analysis of it, in search for a N(L)OAEL.

Modification of the dose descriptor

Skin

The dose descriptor for dermal irritation/corrosion may come from dermal sub-acute or sub-chronic studies. In these studies, the dose is reported in the unit mg/kg of body weight/day. This needs to be modified to enable comparison with the human exposure, generally expressed in mg/cm^2/day. Taken that:

- the average weight of rats is 250g (200-300g),
- the dose is applied over an area which is approximately 10% of the total body surface, and
- the total body surface of rats is on the average 445 cm^2 (363 to 527 cm^2).

the generic modification from the $NOAEL_{test}$ (in mg/kg of body weight) to $NOAEL_{modified}$ (in mg/cm^2/day) will be

$NOAEL_{test} * 0.25/44{,}5 = NOAEL_{modified}$

For example if the highest dose not causing irritation/corrosion was 100 mg/kg bw in the sub-acute study, the modified dose descriptor would be

100 mg/kg*0.25kg/44,5 cm² = 0.56mg/cm^2 = 560µg/cm^2

The specific figures from the respective test report should be used when available. For the other test animal species (rabbits and guinea pigs) ranges of weight can be found in the OECD TG 410.

If the data on dose descriptor is the highest concentration (based on e.g. human data), which does not irritate human skin or eye, the exposure estimation should address the concentration of the substance in the relevant use, and no modification is needed.

Respiratory system

If the dose descriptor is the highest air concentration, which does not cause respiratory irritation (NOAEC) or the lowest air concentration which causes respiratory irritation (LOAEC) from either an animal inhalation toxicity study or from human data, the exposure estimation should address the air concentration in work-places and in consumer uses, and no modification is needed.

Some modification may be necessary in order to take into account the different respiratory rates of the experimental animal (at rest) and the human. The increased respiratory rate of a worker, whose work may be physically demanding, should be considered. However, these modifications would only apply when there is evidence that the inhaled dose or duration of exposure, and not the concentration, drive the appearance of the effect.

Eye

Usually, quantitative assessment of the eye irritation/corrosion will not be possible, because only qualitative data from the relevant in vitro or in vivo studies are available. Only occasionally, signs of eye irritation/corrosion may be observed in animal inhalation toxicity studies or in humans. If a NOAEC or a LOAEC can be identified, this does not need to be modified.

Application of assessment factors to obtain the DNEL

Dermal irritation/corrosion

Since the mechanism (direct chemical reactivity with cell membranes) of skin irritation/corrosion is considered to be the same in experimental animals and in human, no **inter-species** AF should be applied to the NOAEC or to the LOAEC. Mechanisms that are considered to apply to both human and experimental animal are cytotoxicity, sometimes observed as reduced cell viability and release of inflammatory mediators. Chemical reactivity of the substance is relevant in terms of interaction with the cell membranes (7, 8 ,9). Furthermore, since irritation and corrosion are local effects and metabolism in skin tissues is limited, there are no species differences in toxicokinetics. Concerning corrosivity, the similarity of mechanism in animal and human is more obvious, since corrosive substances induce destruction of the full thickness of the skin, manifested e.g. as ulcers, bleeding and later as alopecia, and scars.

To cover the **intra-species** variation, the default AF of 5 for workers and 10 for general population should be applied to the concentration/dose descriptor.

If the dose descriptor is a LOAEC, an AF should be considered to account for uncertainty related to the extrapolation of the LOAEC to the NAEC (see Section R.8.4.3).

When weight of the evidence approach or uncertainty analysis shows that the data are uncertain, an AF to account for uncertainties related to the quality of the database (according to Section R.8.4.3) can be applied.

Respiratory irritation

For effects on the respiratory tract, whether the mechanism indicates that the effect seen is a simple destruction of membranes due to the physico-chemical properties (e.g. pH) of the chemical concerned or whether a local metabolic process is involved, further kinetic and dynamic considerations still apply. Given that there could be significant quantitative differences in deposition, airflow patterns, clearance rates and protective mechanisms between humans and animals and when there is no data to inform on this uncertainty, it is prudent to assume that humans would be more sensitive than animals to effects on the respiratory tract. In such a situation, a

chemical-specific interspecies AF or the default AF of 2.5 should be applied, as would be the case for systemic effects.

To cover the **intra-species** variation, the default AF of 5 for workers and 10 for general population are applied to the concentration descriptor. If the dose descriptor is a LOAEC, an AF to account for uncertainty related to the extrapolation of the LOAEC to the NAEC (see Section R.8.4.3) should be considered.

Use of data on sensory irritation from the Alarie test may on a case-by-case basis be used for a short-term DNEL-value; and should include weighing all available human and animal data on sensory irritation. The concentration inducing a reduction in respiratory rate of 10 % (i.e., the RD10) is proposed as the threshold concentration for inducing a reduced respiratory rate in mice, which can be used as a starting point to derive a threshold for a biologically significant sensory irritation in humans (10). Depending on the available substance-specific data, the default AF(s) may be adjusted.

Eye irritation

Quantitative human data on eye irritation may be available for some substances, thus, no **inter-species** AF is applied. If the dose descriptor has been identified from an animal inhalation toxicity study, it is proposed that no interspecies AF is applied, as the mechanism (direct chemical reactivity with cell membranes) of eye irritation/corrosion is considered to be the same in animals and humans.

To cover the **intra-species** variation, the normal AF of 5 for workers and 10 for general population are applied to the dose descriptor.

If the dose descriptor is a LOAEC, an AF to account for uncertainty related to the extrapolation of the LOAEC to the NAEC (see Section R.8.4.3) should be considered.

Risk characterisation

Dermal irritation/corrosion

Depending on the available data, risk characterisation is based on concentrations or dose (see Section "Identification of the dose descriptor" above).

In risk characterisation, the derived DNEL(s) for irritation and/or corrosion should be compared with the exposure levels, expressed in the same unit as the DNEL.

In some cases quantitative data may be available from acute, sub-acute or sub-chronic studies. After modification of the dose descriptor into a relevant dose unit (usually mg/cm^2 for skin effects or concentration), and applications of relevant AFs, the resulting DNEL is compared with the exposure levels, either estimated using, e.g., the RISKOFDERM model or measured (given in mg/cm^2).

If the exposure is lower than the DNEL(s), it can be assumed that at that specific exposure no irritation or corrosion of the skin of exposed persons would occur. Exposure might occur more than one time per day and may lead to accumulation of the substance in the skin or accumulation of an effect e.g. due to injury/scaling of stratum corneum. In that case, it is recommended that data from sub-acute/sub-chronic studies or on human data from repeated exposures is used for risk characterisation. On the other hand, when it is clear from the exposure data only rare and short term

exposures occur, then an appropriate short term/acute DNEL for the dermal effects should be used, when available.

Respiratory irritation

In risk characterisation, the derived DNEL for respiratory irritation should be compared with the exposure, which should also be expressed as (work-place) air concentration of the substance in the specific use. The air concentration of the substances (in mg/m^3) can be obtained by measuring it or using a reliable deterministic modelling tool. More guidance on the preparing the relevant exposure assessment can be found in Part D.

If the exposure is lower than the DNEL, it can be assumed that at that specific exposure no respiratory irritation of an exposed person would occur.

Eye irritation

In some cases human data on the irritant/corrosive concentration in the eye may be available. That would be used when setting the DNEL. If the concentration to which humans are exposed is lower than the DNEL(s), it can be assumed that at that specific exposure no irritation or corrosion of the eye of exposed persons would occur.

Qualitative assessment and potency information

If there is no dose-response data and hence a DNEL cannot be derived, a qualitative assessment is necessary. The qualitative assessment would usually include an assessment of the potency of the substance or severity of the effect. A distinction should be made, at the minimum, between irritants, corrosives and strong corrosives. These distinctions are made to enable selection of the adequate RMM/OCs (see Section E.3.4).

Strong acidity/alkalinity (pH is <2 or >11.5) can be used to identify skin and eye corrosives. However, the DNEL or other quantification of the effect can not be based on physico–chemical data on irritation /corrosion.

According to the literature, QSAR tools might enable quantification of the irritation potential using e.g. primary irritation index. Thus, Multicase and TOPKAT provide categorisation of the effect, i.e. they give a distinction between weak/mild/moderate and severe irritation. However, so far none of the QSAR tools can be recommended for the regulatory use and a DNEL can not be based on such data.

In the skin and eye irritation tests, a nominal dose of undiluted test substance is tested. Therefore, the tests would not provide a dose-response data, which would support determination of the DNEL. Potency data will be provided by these tests; see Table R. 8-21 for potency categories of specific in vivo and in vitro tests.

Table R. 8-21 Skin and eye irritation/corrosion: potency categories from in vitro and in vivo tests and from some other types of data

Type of test or other data	Possibility to observe and use potency/severity categories
In vitro , membrane barrier test method, skin corrosion (OECD 435)	Subcategories 1A, 1B and 1C, and non corrosive can be obtained. Categories are based on time of exposure and time of observation (11,12). [1)] For categorisation in Table E.3-1 in Section E.3.4, subcategory 1A is for strong corrosives (equivalent to R35), while subcategories 1B and 1C are for corrosives (equivalent to R34).
In vitro test guideline B.40 (Human Skin Model test, Rat skin TER assay)for corrosivity (B40 covers 430 and 431 OECD test guidelines)	Results from these tests do not allow sub-categorisation (see Section R.7.2)
In vivo test guidelines B.4 and B.5 for skin and eye irritation/corrosion [2)]	Scoring is reported for one concentration, which is normally 100% or the test substance (12). Using the scoring, the distinction can be made between strong corrosives (R35), corrosives (R34) and skin irritants (R38); for the eye, distinction can be made between serious damage to eyes (R41) and irritation to eyes (R36) (see also Section E.3.4).
Read-across	Depends on the quality of the data on the "source" chemical

[1)] Three potency categories as requested by the UN GHS (1);

[2)] Validation of a refined Draize test called rabbit low volume eye test (LVET) is ongoing. In LVET, same grading scale as in B.5 is used.

After considering the available qualitative data on irritation and corrosion, the key study is selected, and this should give an indication of the potency. If the relevant data comes only from human evidence, the potency information is not always available.

Respiratory system

Qualitative animal or human data on respiratory irritation, if available, would usually lead to selection of adequate RMMs as instructed in Section E.3.4.

Normally, assessment of substances for the corrosivity in the respiratory tract would not be necessary, because corrosion in the respiratory tract would cause acute toxicity and the hazard assessment would be made accordingly. There may be cases then a corrosive substance is diluted and would therefore need to be assessed for respiratory irritation.

Table R. 8-22 Types of quantitative and qualitative data.

Note that read-across, which has to be justified, could apply to all endpoints. See Table R. 8-21 and Section R.7.2 for more details of relevant tests.

Endpoint	Quantitative data on the endpoint from	Qualitative or potency data from
Skin irritation	Sub-acute and sub-chronic dermal toxicity studies and human data	Test B.4, some pre-validated in vitro tests guidelines
Skin corrosion	Human data [1)]	Test B.4, B.40
Eye irritation	Human data	Test B.5
Eye corrosion	Human data [1)]	Test B.5
Respiratory irritation	Acute, sub-acute and sub-chronic inhalation toxicity studies and human data	Human data
Respiratory corrosion	Not relevant, see the chapter "respiratory system" above.	Not relevant, see the chapter "respiratory system" above.

[1)] Animal systemic toxicity studies are not mentioned here, because signs of corrosion are normally not observed in them. If there are indications of corrosivity, a substance would not be tested for systemic toxicity by any route of exposure.

REFERENCES

(1) Globally Harmonised System of Classification and Labelling of Chemicals (GHS) available at: http://www.unece.org/trans/danger/publi/ghs/ghs_rev00/English/GHS-PART-3e.pdf

(2) Monique A. J. Rennen, Kirsti Nordheim, Geert F. Houben and Cees de Heer. (2002) Prediction of Local Irritant Effects after Repeated Dermal and Respiratory Exposure to Chemicals Regulatory Toxicology and Pharmacology , 36, 3, 253-261.

(3) Test guidelines B.8, B.9, B.28 and B.29 of Annex V to Directive 67/548/EEC at http://ecb.jrc.it/documents/Testing-Methods/ANNEXV

(4) U.S. Department of Labor. Medical surveillance – Formaldehyde. Accessed 7.7.2007. http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10078

(5) UK Health Protection Agency. Chemical Hazards review. Accessed 7.7.2007 http://www.hpa.nhs.uk/chemicals/compendium/

(6) Examples concentration dependent classifications can be found in Commission Directive 2004/73/EC Annex I, and in the Risk Assessment Reports prepared according to the Regulation 793/93.

(7) Osborne R, Perkins MA. (1994). An approach for development of alternative test methods based on mechanisms of skin irritation. Food Chem Toxicol. ;32(2):133-42.

(8) MA Perkins, R Osborne, FR Rana, A Ghassemi and MK Robinson. (1999). Comparison of in vitro and in vivo human skin responses to consumer products and ingredients with a range of irritancy potential. Toxicological Sciences, Vol 48, 218-229,

(9) Kiran Kodithala, A. J. Hopfinger, Edward D. Thompson and Michael K. Robinson. (2002). Prediction of Skin Irritation from Organic Chemicals Using Membrane-Interaction QSAR Analysis. Toxicological Sciences 66, 336-346

(10) Peter M.J. Bos (2007). The mouse sensory irritation bioassay (Alarie-test): Its relevance for AEGL-derivation. SIR-advisory report 11449A00. RIVM, De Bilt, The Netherlands.

(11) In vitro Membrane Barrier Test Method for Skin Corrosion, OECD Guideline for Testing of Chemicals: 435, July 19, 2006 http://www.oecdbookshop.org/oecd/results.asp?SF1=SeriesIdentifier&lang=EN&ST1=SER-00611E1

(12) Test guidelines B.4 and B.5 of the Annex V to Directive 67/548/EEC

http://ecb.jrc.it/testing-methods/annex5/

APPENDIX R. 8-10 Skin sensitisation

Skin sensitisation is generally regarded as a threshold effect, although in practise it may be very difficult to derive a threshold and to set a DNEL.

Data that can be used to inform on potency and possibly thresholds for sensitisation may come from human experience/studies, the mouse local lymph node assay (LLNA), or other animal tests (e.g., GPMT and the Buehler test). While in case of data derived from the GPMT and the Buehler test normally only a qualitative assessment can be performed, the human data and data derived from the LLNA may be in some cases used in a more quantitative manner. In terms of quantification however, there often are considerable uncertainties related to the underlying data to be used, as well as in determining the appropriate assessment factors.

Thus, the general approach to sensitizers could be viewed as a two-step procedure involving;

1. a qualitative approach (by using potency categorisation and following the approach described in Section E.3.4 to define the risk management measures (RMMs) and operational conditions (OCs),
2. and setting a DNEL (if possible) to judge the remaining/residual likelihood of risks after these RMMs and OCs are implemented

RMMs and OCs may be chosen in relation to the potency of the sensitizer. The more potent is the sensitizer (2), the more stringent measures to control exposure are required. The qualitative approach is discussed in Section E.3.4.

However, in cases where the data base gives reliable information on the dose-response, a DNEL can be set to judge the remaining/residual likelihood of risks after the RMMs/OCs are implemented. In analogy with the DMEL approach for carcinogens, the sensitisation DNEL may be useful when preparing the chemical safety assessment to judge the remaining/residual likelihood of risks. Based on such judgement the registrant may need to further refine the way he uses or recommends to use the substance by revising the relevant tentative exposure scenario(s) for use of the substance.

Examples on relevant RMMs and OCs in relation to the sensitisation potency (if known), are given in Section E.3.4. For cases where the data are good enough to allow setting a DNEL, as judged by expert judgement, guidance is given below on how to proceed.

Note that the quantitative assessment of sensitizers and derivation of DNELs does not have any implications on the classification and labelling of sensitising substances or preparations containing sensitizers.

Guidance for potency categorisation (for qualitative approaches) and for setting DNELs that can assist in judging the remaining/residual likelihood of risks after the RMMs and OCs are implemented

Sensitisation usually refers to the induction of an immunological (hypersensitive) state following exposure to a chemical, such that subsequent exposure to that chemical (or cross-reactive chemical) will cause an allergic reaction (elicitation). Both induction and elicitation display a dose-response relationship and have a threshold. The threshold for induction can be defined as the highest level of exposure that fails to induce sensitisation (1). The threshold for elicitation can be defined as the highest level of exposure that fails to elicit a reaction in a previously sensitised subject (2).

The threshold for induction, a measure of intrinsic sensitising potency, can vary widely between sensitizers, with more potent sensitizers having a lower threshold. The relationship between sensitising potency and elicitation is not well characterised. The dose response relationship for the two differs as the dose required to induce sensitisation in a non-sensitised subject is usually greater than that required to elicit a reaction in a previously sensitised subject. These matters have been the subject of recent expert reviews (2). A threshold for skin sensitisation can be defined as concentration (in %) or a dose per skin area (in µg/cm2), which can be calculated from the concentration tested, the patch size and application volume. In fact, for most normal conditions of exposure it is the dose of sensitizer per unit (skin) area the critical determinant of whether sensitisation occurs (3, 4) which allows also the comparison of results derived from different test methods (e.g. animal and human) as well as a quantitative risk assessment. It has been shown however, that when the exposed area becomes very small (less than 1 cm^2), the size of the exposed area becomes the critical determinant.

In cases where a threshold for induction or elicitation can be derived from the available data, risk characterisation could be conducted using the approach taken for threshold effects. However, in practice it may be difficult to derive a threshold for either induction or elicitation and in these cases a threshold approach will not be possible. Elicitation thresholds seem to correlate poorly with induction potency (2, 5). There is large variation in elicitation thresholds between people, which depends on the sensitising potency of the substance, duration, site and extent of exposure, condition of the skin and a very important determinant is also the extent to which the sensitisation has been acquired (2, 5, 6). The EU Expert Group on skin sensitisation nominated by the Technical Committee of Classification and Labelling (2) concluded that it would not be appropriate to define elicitation thresholds as a function of skin sensitising potency. An attempt should however be made as far as possible to describe the magnitude of the risk for each exposure scenario based on qualitative considerations, expressed as high or low concern.

In view of the difficulties in deriving thresholds for elicitation, the description of the quantitative and semi-quantitative approaches to define the potency (to be used in a qualitative approach as described in Section E.3.4) and derive DNELs for skin sensitizers in the following chapters will focus only on the induction phase of skin sensitisation.

Potency categorisation and identification of the typical dose descriptor

Animal data

Some animal tests, such as the local lymph node assay (LLNA) and guinea pig tests (e.g. open epicutaneous test) can provide information on the potency for induction and/or elicitation in animals and possibly also information on the dose-response relationships (4). In particular the dose-response data generated by the LLNA makes this test more informative than the guinea pigs assays for the assessment of potency. The LLNA has been shown to correlate relatively well with the limited human data for induction thresholds and may therefore be used in a predictive manner (7, 8). Data derived by the reduced LLNA (rLLNA) can not be used for assessment of potency and for derivation of a threshold.

Potency information if available can be used in qualitative risk characterisation and for recommendation of appropriate RMMs (see Section E.3.4). Proposals for potency categorisation of sensitizers based on the LLNA, Guinea pig maximisation test (GPMT) and the Buehler test have been put forward by Kimber et al. (6), ECETOC (26), by the EU Expert Group on skin sensitisation (1, 2) and a proposal is currently under development by the Expert group on sensitisation under the Task Force on harmonisation of Classification and Labelling at the OECD. In case of the LLNA, potency is categorised based on the (Effect Concentration 3) EC3 value (see details on derivation of

EC3 below), while in the case of the GPMT and the Buehler test, potency is based on the percentage of positive animals in relation to the induction concentration tested. Potency categorisation as proposed by the EU Expert Group on skin sensitisation (1, 2) is now being used in the EU on a case by case basis to set specific concentration limits for sensitising substances in preparations. Data derived from the LLNA can be used for assignment of skin sensitizers in specific potency classes, while there are more limitations in using data from the GPMT and the Buehler test for this purpose. The higher uncertainty associated with the latter methods are particularly due to the single induction dose regime employed. However, data from dose-response studies would reduce the degree of uncertainty. More details and guidance on potency categorisation is given in the report from the EU expert group (1) as published in Basketter et al., 2005 (2).[15]

Potency categorisation as proposed by the EU expert group on skin sensitisation (1, 2) is shown in the following tables.

Table R. 8-23 Potency categorisation based on LLNA

Category	EC3 (%)
Extreme	≤ 0.2
Strong	> 0.2 - ≤ 2
Moderate	> 2

Table R. 8-24 Potency categorisation based on GPMT:

Intradermal concentration employed during induction phase (%)	Incidence of sensitisation (30 – 60%)	Incidence of sensitisation (≥ 60%)
≤ 0.1	Strong*	Extreme
> 0.1 - ≤ 1	Moderate*	Strong*
> 1	Moderate	Moderate*

*acknowledged by the EU expert group that this categorisation is associated with high degree of uncertainty (1, 2)

Table R. 8-25 Potency categorisation based on the Buehler test

Concentration employed during induction phase (%)	Incidence of sensitisation (15 – 60%)	Incidence of sensitisation (≥ 60%)
≤ 0.2	Strong*	Extreme
> 0.2 - ≤ 20	Moderate*	Strong*
> 20	Moderate	Moderate*

*acknowledged by the EU expert group that this categorisation is associated with high degree of uncertainty (1, 2)

[15] This guidance will be included in the REACH implementation project 3.6 which will develop guidance on classification and labelling under the future EU Regulation on classification, labelling and packaging of substances and mixtures, implementing the Globally Harmonised System (GHS) in the EU..

In the LLNA, the endpoint of interest is the proliferative response in the draining lymph nodes of mice exposed topically to the test material. Substances that elicit a 3-fold or greater proliferative activity compared to the controls (induce a stimulation index (SI) of 3 or more) are considered to be skin sensitizers. The amount of chemical required to induce a SI of 3 is the Effect Concentration 3 (EC3), which can be estimated from the LLNA dose-response data. Often linear interpolation of the points in the dose response curve which lie immediately above and below the SI of 3 is proposed, but more advanced statistical approaches basing conclusions on the characteristics of the dose response curve and variability of results are also used (9, 10, 11). Since the EC3 is usually expressed as concentration (%) it needs to be converted to dose per skin area (µg/cm^2). The EC3 in (µg/cm^2) can be calculated as shown below, by considering the dose volume of 25 µl (according to the standard LLNA protocol), and an estimated application area of 1 cm^2 for the mouse ear. Assuming the density of the liquid is 1, a conversion factor to be applied to the EC3 (%) of 250 is calculated by converting 25 µl/cm^2 into µg/cm^2. The EC3 (%) is then converted to EC3 (µg/cm^2) as shown in the formula below:

EC3 [%]*250 [µg/cm^2/%] = EC3 [µg/cm^2]

EC3 value can be considered as the LOAEL for induction.[16] The EC3 value (expressed in µg/cm^2) can be used for quantitative assessment to derive no-effect levels (including DNEL), providing that relevant assessment factors (AFs) are applied. Examples of quantitative assessments for specific skin sensitizers in defined exposure situations are available (12, 13, 14, 15, 16) and this approach has been used, to identify safe exposure levels for a range of skin sensitising chemicals, such as for example fragrances and preservatives in consumer products (17, 18, 19, 20, 27). It has to be noted however, that in these cases the weight of evidence approach (WoE) using human and animal data (LLNA) has usually been applied (note that for industrial chemicals historical human data are usually not available). Furthermore the approach was specifically developed for specific types of chemicals and for well defined exposure situations (e.g. exposure to a specific type of consumer products) resulting in inclusion of certain considerations related to exposure (that under REACH should be considered as part of the exposure assessment) in the calculation of the no-effect levels. Based on this, the approach can not be directly applied to industrial chemicals in general; therefore guidance in particular in relation to the use of AFs, is given in the following sections.

Human data:

Human data will normally take preference over animal data, although the reliability of such data should be carefully assessed, particularly if derived from old studies, and may be rejected in favour of animal data. Lack of positive findings in humans should not normally overrule positive and good quality animal data.

[16] In some papers it is suggested that the LLNA EC3 value is close enough to the human NOAEL and that therefore it can be used as a surrogate for the NOAEL (19, 26, 28).

Human data on induction thresholds is normally not available and testing for induction of sensitisation in humans is no longer conducted on ethical grounds. However, there may be data available from historical predictive testing to inform on potency, a threshold/NOAEL or a LOAEL which should be considered on a case-by-case basis. Due to standard exposure conditions, data from historical predictive tests (e.g. human repeat insult patch test (HRIPT) and human maximisation test (HMT)) may provide information on potency for induction. Thresholds from reliable historical human predictive tests can be used in combination with the LLNA data in a weight of evidence (WoE) approach to set a NOAEL/LOAEL for induction of sensitisation (12, 13, 14, 15, 19, 20). The NOAEL/LOAEL from human predictive tests should be calculated from the concentration of the substance tested, the patch size and the application volume and should be expressed as function of dose per unit skin area. The NOAEL from such tests should be the dose at which no sensitisation in the exposed people has occurred, while the LOAEL sometimes has been proposed to be the dose at which 5% (or < 8%) of the exposed people were sensitised.[17]

As already mentioned above, new experimental testing for hazard identification in humans, including HRIPT and HMT, is not acceptable for ethical reasons, therefore historical information from this type of studies will be available for a limited number of chemicals. Furthermore, the quality/reliability of the results from these studies should be carefully checked in particular in relation to the number of people tested (21).[18]

Potency information if available can be used in qualitative assessment and for recommendation of appropriate RMMs (See also Section E.3.4).

Testing humans with pre-existing contact allergy to determine sensitisation to a particular substance is done extensively as part of clinical examinations. Evidence of skin sensitising activity derived from such tests demonstrates the (previous) induction of skin sensitisation to that substance or cross-reaction with a chemically very similar substance. Moreover, clinical examinations are usually not designed to determine elicitation thresholds, as the dose used is the one giving response in the majority of sensitised subjects. However, clinical data should be used for qualitative assessment.

The potency of a chemical could be evaluated by comparison of the incidence of skin sensitisation in the human population with the exposure situation, if known. For example, if for a certain substance a high incidence of contact allergy is observed in an exposed population in relation to relatively low degree of exposure, this could be considered as an indication that the substance is a strong sensitizer while in cases where a low incidence of contact allergy among exposed individuals in relation to high degree of exposure would be observed, this would be an indication that a substance is a weaker sensitizer (21). However, normally exposure is not sufficiently well defined and positive findings in epidemiological studies (population based studies, data from contact dermatitis patients, studies/data from occupational groups and outbreak population studies) can only provide evidence sufficient for hazard assessment.

Potency of induction cannot be directly derived from human elicitation threshold data from diagnostic clinical studies (e.g. patch test dose-response data, Repeated Open Application Test (ROAT)), however, a low elicitation threshold could indicate high potency and vice versa (21).

---

[17] These "human LOAEL threshold values" are taken from two different publications (7, 8).

[18] For the HRIPT a large number of people are required in each test, to reduce the 95% confidence interval for the test result (22). In different publications, different acceptable number of people tested in HRIPT can be found (e.g. 100 (19) or 150 – 200 (22)).

In vitro data

No officially adopted EU-OECD in vitro tests for skin sensitisation are currently available. Several in vitro assays to detect sensitising properties of a chemical are currently under development for the area of epidermal bioavailability, chemical reactivity and cell-based assays. At the moment, *in vitro* tests may be used only as supportive evidence in combination with other types of data for the identification of skin sensitizers.

Non-testing data

Non-testing methods for skin sensitisation include grouping of chemicals (read-across/chemical categories), chemistry considerations and (Q)SARs. Detailed information on the application of these methods is given in Section R.7.3. A number of (Q)SARs for skin sensitisation are reported in literature, which include local and global (Q)SARs and expert systems (see Section R.7.3). The available (Q)SARs may be suitable mostly for hazard identification, in particular in a WoE approach. With some QSARs the potency as well as the EC3 value can be estimated. It should be noted however that in order to be used for regulatory purposes the validity and adequacy of the (Q)SARs need to be established (see Section R.6.1).

One approach that can be used when lacking information to assess skin sensitisation is to perform read-across if experimental information is available for substances structurally closely related to the investigated substance. In relation to potency estimation and derivation of a DNEL for induction, the use of local QSARs (their scope is characterised by a mechanistic reactivity domain) could be useful in particular if applied in the so called mechanistic read-across (22, 23) in which by assigning the substance to the appropriate reaction applicability domain and by quantification of it's reactivity/hydrophobicity relative to the known sensitizers in the same domain for which experimental information is available it would be possible to predict, within a range, the likely sensitisation potential, including potency. The assumption behind this mechanistic read-across is that sensitisation potential is always related to a combination of reactivity and hydrophobicity.

Application of assessment factors (AFs) to the correct starting point to obtain the induction-specific DNEL

As for other toxicological endpoints, depending on the method used, default AFs for inter- and intra-species variation, for dose response uncertainties (considering the uncertainties in the dose descriptor as the surrogate for the true no-adverse-effect –level (NAEL) and uncertainties in extrapolation of the LOAEL to NAEL, the EC3 is considered as the LOAEL for induction[19]) and uncertainties related to the quality of the whole database need to be considered when calculating the DNEL based on a NOAEL/LOAEL/EC3 (see Section R.8.4.3).

Skin sensitisation specific AFs

AF for vehicle or matrix effect

The available information on potency of various sensitizers comes usually from studies using a simple matrix (usually as recommended in the testing guideline), but the actual human exposure might involve exposure to sensitizers in a different or more complex matrix which might increase the potential for induction of sensitisation (e.g. matrix with irritant or/and penetration enhancing properties). In these cases the application of an additional AF of 1-10-fold should be considered, depending on the information available on the vehicle or matrix relevant for human exposure (14). If human exposure is expected in a matrix with no penetration enhancers or irritants, an AF of 3 might be sufficient, or if the matrix is very similar to the matrix used to determine the NOAEL/LOAEL/EC3 and is not expected to increase the potential for induction of sensitisation, the AF may be reduced to 1 (14).

AF for different exposure conditions

In addition, on a case by case basis an additional AF (1 – 10 fold) should be considered to account for specific exposure condition considerations (that are not considered in the exposure assessment). This concerns situations when the experimental set up (animal or human) differs from actual human exposure conditions, by e.g. different parts of the body being exposed, differences in skin integrity caused by specific human activities, occlusion of the exposed skin and differences in exposure frequency between the animal/human study and actual human exposure situation (14). It is important to consider that repeated exposure may lead to induction of skin sensitisation at exposures lower than experimentally derived induction threshold. This was observed in some animal and human studies (24, 25). It is not clear whether this occurs with most or only some sensitising chemicals, and which mechanism is involved (e.g. slow release of initially bound chemical to the upper skin layers, accumulation of the substance). Thus, the application of an AF to account for the uncertainty related with repeated exposure needs to be considered in the derivation of the DNEL.

The application of skin sensitisation specific AFs should be decided by expert judgement and justified on a case by case basis. Therefore, in some cases it might be decided that AFs higher than the ranges proposed above are appropriate.

Derivation of the induction specific DNEL for skin sensitisation

a) *Derivation of the induction specific DNEL based on LLNA data only*

[19] In some papers it is suggested that the LLNA EC3 value is close enough to the human NOAEL and that therefore it can be used as a surrogate for the NOAEL (19, 26, 28)

The EC3 value expressed in dose/unit area of exposed skin (e.g. µg/cm^2) can be considered as the LOAEL for induction[20]. By the application of the relevant AFs, a DNEL can be derived expressed in µg/cm^2/day. Application of the relevant default AFs (see Section R.8.4.3) and skin sensitisation specific AFs should be considered as described in the section on the application of assessment factors above. EC3 data generally correlate well with human skin sensitisation thresholds derived from historical predictive testing; however there are cases where this correlation is poor and the two values may differ by 10-fold or more. In view of this variation, the default AF of 10 for interspecies variation (see Section R.8.4.3) should be used, unless there is evidence (e.g. from a close analogue of the substance in question) of good correlation between the EC3 and human NOAEL/LOAEL. Therefore, on a case by case basis the interspecies AF could be lowered.

b) *Derivation of the induction specific DNEL based on* WoE using *LLNA and historical human predictive test data (HRIPT or HMT)*

For substances for which both the LLNA and historical human predictive test data of good quality are available, the DNEL can be derived by the WoE approach. A reliable NOAEL from a well conducted HRIPT would have precedence over the LLNA EC3 value or a NOAEL from HMT. The application of the relevant default AFs (see Section R.8.4.3) and skin sensitisation specific AFs should be considered as described in the section on the application of assessment factors above.

In cases where there is good agreement between the LLNA EC3 value and the NOAEL/LOAEL derived from good quality historical human predictive tests, the lowest threshold value should be used to derive the DNEL and in this case there would be no need to apply an interspecies AF.

In cases where there is significant discrepancy between good quality historical human data and animal data (an order of magnitude or more) and the HRIPT NOAEL would be higher than the LLNA EC3 it should be carefully considered whether the higher HRIPT NOAEL is sufficiently robust to override the lower LLNA EC3 value. If this is not the case, the lower value derived from the LLNA should be used to derive the DNEL. In any case, an interspecies AF would not be needed.

In case there are both LLNA data and data from historical human predictive studies available, but human data are not considered to be reliable, the LLNA EC3 value should be used as the basis for DNEL and the same AFs as in case when no human data is available should be applied.

c) *Derivation of the induction specific DNEL based on read-across from structurally related substances*

It might be possible to derive a DNEL based on the read-across from structurally related substances for which experimental data are available. The same AFs as described in the previous two sections would in this case need to be applied; however, on a case by case basis an additional AF to account for the uncertainty related to the use of the read-across should be considered.

Risk characterisation

In risk characterisation, the derived induction specific DNEL (expressed in µg/cm^2/day) is compared with the estimated exposure, which has to be expressed in µg/cm^2/day. This information needs to be obtained in all exposure assessments concerning sensitizers, using appropriate exposure

[20] In some papers (26) it is suggested that the LLNA EC3 value is close enough to the human NOAEL and that therefore it can be used as a surrogate for the NOAEL (19, 26, 28)

assessment models (e.g. RISKOFDERM, CONSEXPO) or measurements. When performing exposure assessment it should be considered that the exposure might occur more than once per day or repeatedly during a longer period of time and may lead to accumulation of the substance on the same skin site.

If the DNEL exceeds the exposure, it can be assumed that at that specific exposure no induction in a non-sensitised person would occur. However it should be noted, that at this exposure level, a reaction in a previously sensitised person could still occur.

REFERENCES:

(1) ECB (2003). Report from the Expert Working Group on Sensitisation. Ispra, 18-19 April 2002. Ispra, 26 May 2003 (ECB/81/02 Rev 3).

(2) Basketter DA, Andersen KE, Lidén C, van Loveren H, Boman A, Kimber I, Alanko K and Berggren E. (2005a) Evaluation of the skin sensitising potency of chemicals using existing methods and considerations of relevance for elicitation. *Contact Dermatitis*, 52: 39 - 43.

(3) Friedmann PS. (2006) Contact sensitisation and allergic contact dermatitis: immunobiological mechanisms. *Toxicology Letters* 162, 49-54, 2006

(4) Kimber I, Basketter DA, Berthold K, Butler M, Garrigue J-L, Lea L, Newsome C, Roggeband R, Steiling W, Stropp g, Waterman S and Wiemann C. (2001) Skin sensitization testing in potency and risk assessment. *Toxicological Sciences* 59, 198-208.

(5) Scott AE, Kashon ML, Youcesoy B, Luster MI, Tinkle SS. (2002). Insights into the quantitative relationship between sensitization and challenge for allergic contact dermatitis reactions. *Toxicology and Applied Pharmacology* 183, 66-70.

(6) Kimber I, Basketter D, Butler M, Gamer A, Garrigue J-L, Gerberick GF, Newsome c, Steiling W, Vohr H-W. (2003). Classification of contact allergens according to potency: proposals. *Food Chem. Toxicol.* 41, 1799-1809.

(7) Schneider K, Akkan Z. (2004).Quantitative relationship between the local lymph node assay and human skin sensitization assays. Regulatory Toxicology and Pharmacology 39, 245-255.

(8) Basketter DA, Clapp C, Jefferies D, Safford RJ, Ryan CA, Gerberick GF, Dearman RJ and Kimber I. (2005b) Predictive identification of human skin sensitisation thresholds. Contact Dermatitis, 53, 260 - 267.

(9) Basketter DA, Lea LJ, Dickens A, Briggs D, Pate I, Dearman RJ, Kimber I. (1999a) A comparison of statistical approaches to the derivation of EC3 values from local lymph node assay dose responses. *J Appl Tox* 19: 261-266.

(10) Basketter DA, Lea LJ, Cooper K, Stocks J, Dickens A, Pate I, Dearman RJ, Kimber I. (1999b). Threshold for classification as a skin sensitizer in the local lymph node assay: a statistical evaluation. *Food and Chemical Toxicology* 37: 1167-1174.

(11) Van Och FMM, Slob W, de Jong WH, Vandebriel RJ, van Loveren H. (2000). A quantitative method for assessing the sensitizing potency of low molecular weight chemicals using a local lymph node assay: employment of a regression method that includes determination of the uncertainty margins. *Toxicology* 146: 49-59.

(12) Gerberick GF, Robinson MK, Ryan CA, Dearman R, Kimber I, Basketter DA, Wright Z, Marks JG. (2001a). Contact allergenic potency: correlation of human and local lymph node assay data. *American journal of contact dermatitis* 12; 156-161.

(13) Gerberick GF, Robinson MK, Felter SP, White IR, Basketter DA.(2001b): Understanding fragrance allergy using an exposure-based risk assessment approach. *Contact Dermatitis* 45: 333-340.

(14) Felter SP, Robinson MK, Basketter DA, Gerberick GF. (2002): A review of the scientific basis for uncertainty factors for use in quantitative risk assessment for the induction of allergic contact dermatitis. *Contact Dermatitis* 47: 257-266.

(15) Felter SP, Ryan CA, Basketter DA Gilmour NJ, Gerberick GF. (2003). Application of the risk assessment paradigm to the induction of allergic contact dermatitis. Regulatory *Toxicology and Pharmacology* 37; 1-10.

(16) Basketter DA, Jefferies D, Safford BJ, Gilmour NJ, Jowsey IR, McFadden J, Chansinghakul W, Duangdeeden I, Kullavanijaya P. (2006). The impact of exposure variables on the induction of skin sensitization. *Contact Dermatitis* 55: 178-185.

(17) Zachariae C, Rastogi R, Devantier C, Menne T and Johansen JD. (2003). Methyldibromo glutaronitrile: clinical experience and exposure-based risk assessment. *Contact dermatitis* 48: 150-154.

(18) Basketter DA, Angelini G, Ingber A, Kern P, Menne T. (2003). Nickel, chromium and cobalt in consumer products: revisiting safe levels in the new millennium. *Contact dermatitis* 49: 1-7.

(19) Api AM, Basketter DA, Cadby PA, Cano M-F, Graham E, Gerberick F, Griem P, McNamee p, Ryan CA, Safford B. (2006). Deramal Sensitization Quantitative Risk Assessment (QRA) for fragrance ingredients. Technical dossier. March 15, 2006 (revised May, 2006).

(20) Basketter DA, Clapp CJ, Safford BJ, Jowsey IR, McNamee PM, Ryan CA, Gerberick GF. (2007). Preservatives and skin sensitisation quantitative risk assessment: risk benefit considerations. *Dermatitis.* Submitted

(21) European Commission (2000). Opinion concerning the predictive testing of potentially cutaneous sensitising cosmetic ingredients or mixtures of ingredients adopted by the SCCNFP during the 11th plenary session of 17 February 2000. http://ec.europa.eu/health/ph_risk/committees/sccp/docshtml/sccp_out102_en.print.htm

(22) Roberts DW, Patlewicz G, Kern PS, Gerberick F, Kimber I, Dearman RJ, Ryan CA, Basketter DA, Aptula AO. (2007a). Mechanistic applicability domain classification of a Local Lymph Node Assay dataset for skin sensitisation. *Chem. Res. Toxicol.* Submitted

(23) Roberts DW, Aptula AO, Patlewicz G, Pease C.(2007b). Chemical reactivity indices and mechanism-based read across from non-animal based assessment of skin sensitisation potential. *Journal of Applied Toxicology.* Submitted

(24) De Jong WH, ter Beek M, Veenman C, de Klerk A, van Loveren H. (2005). Effect of repeated and prolonged exposure to low concentrations of low molecolar weight chemicals on local lymph node responses. RIVM report 340300001/2005: 1-20.

(25) Griem P, Goebel C, Scheffler H. (2003). Proposal for a risk assessment methodology for skin sensitization based on sensitization potency data. *Regulatory Toxicology and Pharmacology* 38: 269-290.

(26) ECETOC. (2003). Contact Sensitization: Classification according to potency. Technical Report No. 87.

(27) Api AM, Basketter DA, Cadby PA, Cano M-F, Graham E, Gerberick F, Griem P, McNamee p, Ryan CA, Safford B. (2007). Dermal Sensitization Quantitative Risk Assessment (QRA) for fragrance ingredients. *Regulatory Toxicology and Pharmacology.* Accepted for publication

APPENDIX R. 8-11 Respiratory sensitisation

Respiratory hypersensitivity includes asthma and other respiratory conditions, irrespective of the mechanism (immunological or non-immunological) by which they are caused (see also Section R.7.3). The induction phase of respiratory sensitisation can be described as the process of rendering the airways unusually sensitive (hypersensitive) such that following subsequent inhalation exposure an asthmatic reaction might be elicited associated with classical symptoms of airway narrowing, chest-tightening and bronchial restriction. The mechanisms of immunologically- mediated respiratory hypersensitivity are described in Section R.7.3.

As for skins sensitisation, there is evidence that for respiratory sensitisation the dose-response relationships exist, although these are frequently less well defined. Nothing or little is known about the dose-response relationships in the development of respiratory hypersensitivity by non-immunological mechanisms.

At present there are no validated or widely accepted animal or *in vitro* test protocols to detect respiratory sensitisation or to determine the induction or elicitation thresholds. Data from the local lymph node assay (LLNA) or other tests used to identify skin sensitizers (e.g. the guinea pig maximisation test (GPMT) or Buehler test) may be of some use in characterising likely respiratory sensitising activity. Currently the view is, that most of the chemicals that are known to cause respiratory allergy are able to elicit positive responses in these tests. Therefore for example, for chemicals that test positive in these tests it can not be wholly excluded that would not also cause sensitisation of the respiratory tract after inhalation or dermal exposure. However to confirm this, further testing would be needed as not all skin sensitizers identified in these tests are deemed to be respiratory sensitizers. On the other hand, a chemical which is negative in these tests (at an appropriate test concentration) most probably also lacks the potential to cause respiratory allergy. More information on the animal data that could be used for identification of respiratory sensitizers is given in Section R.7.3.

Due to the lack of an appropriate predictive animal study, at present hazard identification is based on human data. This data could be derived from consumer experience and consumer tests, records of worker's experience, published case reports and epidemiological studies. Types of human data and studies that can be used for identification of respiratory sensitizers are further described in Section R.7.3.

Currently available methods do not allow determination of threshold and establishment of a DNEL. Therefore for substances classified as respiratory sensitizers only qualitative assessment as described in Section E.3.4 can be performed.

APPENDIX R. 8-12 Reproductive toxicity

The hazard assessment and derivation of DNEL-values for a substance should be based on the data requirements described in the integrated testing strategy. This section addresses guidance for derivation of DNEL-values at the different levels of knowledge. This should enable the registrant to consider appropriate risk management measures also in the period waiting for further data to be generated (cf. REACH Annex I, 0.5). It should be stressed however, that the DNEL-values that are hampered by uncertainty due to lack of data should not be used as a substitute for the required testing.

Identification of the typical dose descriptor for DNEL calculation

From the data on reproductive toxicity a DNEL value for effects on fertility ($DNEL_{fertility}$) as well as for developmental toxicity ($DNEL_{development}$) should be derived.

The methodology for the DNEL calculation for reproductive toxicity is similar to the methodology as described for repeated dose toxicity. However, reproductive toxicity also includes effects which may occur after one single exposure in a susceptible window during foetal development (e.g. malformations and functional deficits). Most often it is not known from the data whether the reproductive effect has occurred after single or repeated exposures. However, given the conservative nature of the proposed methodology, a DNEL value for reproductive toxicity should be sufficient to ensure that adverse effects do not occur following high short-term exposures, which should not result in an exceeding of the daily DNEL value.

Usually, the various aspects of reproductive toxicity are considered to be effects with underlying dose threshold mechanisms and a NOAEL or LOAEL value should normally be provided from the available data, though the threshold dose for specific aspects of reproductive toxicity is not always easy to identify. If the data allow calculation of a benchmark-dose value (BMD_x) (see Section R.8.2) this can possibly be used as a starting value instead of a N(L)OAEL value. In the rare case that a NOAEL has been derived from well-reported and reliable human data, this should be considered for the DNEL calculation, but generally a value from a study conducted in animals will be used.

Non-threshold effects

When it is known that genotoxicity is the underlying mechanism for the reproductive toxicity of a substance, it is prudent to assume that a threshold dose/concentration cannot be identified. In such cases, a DNEL value can not be derived and instead the assessment should be based on the approaches described for genotoxic substances (see Sections R.8.5 and R.8.6).

Derivation of DNEL, application of assessment factors

A number of studies may provide relevant information in relation to reproductive toxicity. However, the different studies may provide different levels of certainty with respect to the evaluation of reproductive toxicity. Since reproductive toxicity is a complex endpoint expert judgement using an overall weight of evidence approach considering all available data is crucial when performing safety assessment for this endpoint.

Therefore the choice of a specific assessment factor in relation to qualitative and quantitative uncertainties should be decided on a case-by-case basis. The recommendations for specific assessment factors given below are therefore given as intervals with the intention that the risk assessor should evaluate the available information and justify the choice of the assessment factor.

OECD TGs for identification of reproductive toxicity

OECD TG 414 (prenatal developmental toxicity study)

OECD TG 415 (1-generation study) + future extended version

OECD TG 416 (2-generation study)

OECD TG 426 (developmental neurotoxicity study)

Information from one or several of these current guideline studies can be used with confidence to identify substances as being, or as not being, toxic to reproduction in relation to the endpoints addressed in the test and thus can be used for risk assessment and DNEL-calculation. From the studies the relevant N(L)OAELs should be identified for effects on fertility and development, and $DNEL_{fertility}$ and $DNEL_{development}$ calculations should be performed according to the general rules concerning conversion of the dose descriptor and the use of assessment factors (see Sections R.8.4.2 and R.8.4.3). Consideration should be given as to whether the derived DNEL is relevant for males or females, the pregnant or the lactating mother, the newborn or the child.

Aspects concerning maternal toxicity and developmental toxicity

Differences in the sensitivity to general toxicity between the pregnant or lactating female and the non-pregnant animal may be apparent from comparison of reproductive toxicity test results with findings in (sub-)chronic tests. Maternal toxicity may be observed in fertility and/or developmental toxicity assays at dose levels lower than those obtained from general repeated dose toxicity studies. In such cases, consideration should be given to the identification of pregnant women as a more vulnerable population. If such differences are apparent, the adequacy of the DNEL-value for repeated dose toxicity must be assessed with respect to the pregnant or lactating female.

Particular attention should be given to the relationships between dose level and adverse effects on reproduction compared to other systemic toxicity. In cases where developmental toxicity is seen only in the presence of maternal toxicity, DNEL-calculations must be conducted with respect to both the developing offspring and the mother. This should be the case even when a causal association between maternal and developmental toxicity has been demonstrated and it has been concluded that developmental toxicity would not occur in the absence of maternal toxicity. When deciding on assessment factors for the $DNEL_{development}$ calculation, the developing offspring should be the focus of attention. This is important for the protection of the unborn child. The effects in the mother may be mild and reversible, attracting a low level of concern leading to a low overall assessment factor for the DNEL calculation, whereas the effects in the offspring at similar exposure levels may have serious long-term consequences and a higher overall assessment factor may be warranted leading to a lower DNEL value in order to protect against possible developmental toxicity.

Severity of effects

The extent and severity of the effects seen at the LOAEL in reproductive toxicity studies may in some cases be very marked, e.g. extensive foetal or offspring death, major malformations, severe functional defects in the offspring, infertility or severe effects on the reproductive system. This should be reflected in the use of an appropriate assessment factor to account for the uncertainty related with the 'dose-response relationship' (see Section R.8.4.3.1).

OECD TGs 421 and 422 (reproductive toxicity screening studies)

The purpose of the Reproduction/Developmental Toxicity Screening Test (OECD TGs 421 and 422) is to provide information of the effects on male and female reproductive performance such as gonadal function, mating behaviour, conception, development of conceptus and parturition. It should be regarded as an in vivo screening assay and is not designed as an alternative or a replacement of the reproductive toxicity studies (OECD TGs 414 and 416). These screening tests are not meant to provide complete information on all aspects of reproduction and development such as that obtained from a two-generation reproduction study (OECD TG 416). In particular, the post-natal effects associated with prenatal exposure (such as undetected malformations affecting viability or functional effects) or effects resulting from post-natal exposure or exposure during lactation are not covered in these studies. Furthermore, the exposure duration in these studies does not cover a full spermatogenic cycle and the number of animals per dose group is limited.

A positive result in OECD TG 421/422 may be considered sufficient for the calculation of a $DNEL_{fertility}$ and/or a $DNEL_{development}$; however, an additional assessment factor of 2 to 5, decided on a case-by-case basis, should generally be used to take account of the lower sensitivity of the study, unless there is evidence to support that the lower sensitivity is not relevant for the effect mechanism of the substance (e.g. specific teratogenic effects that are the result of a known mechanism of action).

A negative result in OECD 421/422 test may lower the concerns for reproductive toxicity, but can not provide reassurance of the absence of this hazardous property. However, it can provide the basis for deriving a $DNEL_{fertility}$ and/or $DNEL_{development}$ from the highest dose level tested and by application of an additional assessment factor of 2 to 5, decided on a case-by-case basis that should account for the limitations of this study (see also Section R.7.6.4). Such a DNEL would be relevant only at the Annex VIII level (10 – 100 t/y) and below.

Aspects concerning maternal toxicity and developmental toxicity and severity of the reproductive effects as described above should be considered as well.

Other specific data concerning reproductive toxicity

Uterotrophic assay

Hershberger assay

Pubertal male/female assay

Intact male assay

The Uterotrophic and Hershberger assays, presently being internationally evaluated under the OECD Test Guideline Program, appear reliable in identifying substances with (anti)oestrogenic or (anti)androgenic mode of action (see Section R.7.6.4.1). The mechanisms and the effects on most of the target tissues are highly relevant for humans. The assays provide in vivo N(L)OAELs for the endpoints examined. For several endocrine disrupters (EDs), the dose levels causing effects in these assays seem to be of a similar magnitude or higher than those causing effects in reproductive and developmental toxicity studies such as the OECD two generation study. Negative data from these assays cannot be used to conclude on the absence of reproductive toxicity effects, because these may arise through many other mechanisms than the specific endocrine mechanisms included in the assays.

A number of assays in experimental animals may provide information on effects on production of steroids (see also Section R.7.6.4.1). The pubertal assays and the intact male assay may also provide information about the potency of the compound in vivo. Effects on the various endpoints included in these assays can be considered adverse and/or as representing an effect on a mechanism relevant for humans.

The mechanistic assays mentioned above are not part of the standard testing requirements. Although they are considered predictive for identification of adverse effects that may be seen in more comprehensive studies, they are not definitive tests. Positive and negative data from these studies may be used in a weight of evidence in combination with other data to fulfil the requirements for hazard and safety assessment, but can not be used as such as the basis for the calculation of the DNEL in relation to reproductive toxicity. Positive data from these tests are also valuable as triggers in a testing strategy.

More information on the above mentioned studies is given in Section R.7.6.4.

Repeated dose toxicity (RDT) studies

OECD TGs 407/410/412 (repeated 28D studies)

OECD TGs 408/409/411/413 (repeated 90D studies)

OECD TGs 452/453 (chronic studies)

OECD TG 451 (carcinogenicity studies)

OECD TG 424 (neurotoxicity study in rodents)

Information from repeated dose toxicity studies may be of some value for safety assessment of reproductive toxicity as these studies provide information on some of the many end-points relevant for reproductive and/or developmental toxicity, i.e.:

- changes in weight of reproductive organs following repeat dosing e.g. uterus, testis, epididymis – dependent on magnitude of effect, maybe in the presence or absence of histopathological changes
- histopathological changes in reproductive organs following repeat dosing e.g. ovaries, testis, prostate gland, uterus
- sperm parameters or oestrus cyclicity, if assessed
- indications of adult neurotoxicity: clinical observations, weight changes, histopathological changes in the brain

From RDT studies only the effects in relation to reproductive organs (i.e. fertility end-point) can be useful for safety assessment for reproductive toxicity. It should, however, be noted that the sensitivity of repeated dose toxicity studies for detecting effects on reproductive organs may be less than reproductive toxicity studies because of the lower number of animals per group (i.e. 5 per sex compared to 20-25 per sex). In addition, a number of cases have demonstrated that effects on the reproductive system may occur at lower doses during the development of foetuses and young animals than in adults (1, 2).

Consequently, in cases where there are indications of adverse effects on the reproductive organs of adult animals a two-generation study (OECD TG 416) may be triggered. In the period waiting for further tests an increased/ additional assessment factor as described below should be used for the $DNEL_{fertility}$ calculation.

To account for the lower sensitivity of the RDT studies for detecting effects on reproductive organs due to few animals in the exposure groups and a possible increased sensitivity in the developing foetuses and young animals an additional assessment factor of 2 to 5 should be considered on a case by case basis (e.g. where there are substantiated indications for adverse effects on the reproductive organs of adult animals). In cases where effects on reproductive organs are seen in the RDT study the size of the assessment factor should be chosen in consideration of the specific data. For example, in relation to histopathological changes in testis found in a RDT study there is empirical evidence that these effects often occur at lower levels than impaired male fertility observed in e.g. a 2-generation study (3). This is due to the big sperm reserve capacity of the experimental animals. Thus, when histopathological changes in testis are found in a 28D RDT study then it may be sufficient to use an assessment factor to account for the difference in exposure duration (subacute to sub-chronic exposure; see Section R.8.4.3.1). The latter assessment factor would not be needed when the data are from a sub-chronic or chronic study. It should be noted however that this DNEL would only be relevant to male fertility, and does not address the lack of data with respect to potential effects on developing foetus or females.

For deriving a $DNEL_{fertility}$ from a N(L)OAEL based on adverse effects on female reproductive organs from RDT studies the same overall aspects should be considered as described above (i.e. the decreased sensitivity of the study and the limited scope of the RDT studies in relation to addressing the reproductive toxicity end-points).

A RDT study showing no adverse effects on reproductive organs is ***not*** considered to provide sufficient information for a DNEL calculation for fertility or other reproductive effects.

REFERENCES

(1) ECB (2004). European Union risk assessment report: dibutyl phthalate. European Commission, Joint Research Centre. Volume 29 (http://ecb.jrc.it/esis/)

(2) ECB (2003). Risk assessment: bis(2-ethylhexyl) phthalate. Consolidated Final report: September 2003 (http://ecb.jrc.it/esis/)

(3) Mangelsdorf I., Buschman J., Orthen B. (2003). Some aspects relating to the evaluation of the effects of chemicals on male fertility. Regulatory Toxicology and Pharmacology 37: 356-369.

APPENDIX R. 8-13 Deriving DNELs, when a community/national occupational exposure limit (OEL) is available

The following guidance applies in situations where

- ***an EU indicative occupational exposure limit (IOEL) has been adopted***
- ***an EU binding occupational exposure limit (BOEL) has been adopted***
- ***a national occupational exposure limit has been adopted***

EU INDICATIVE OCCUPATIONAL EXPOSURE LIMIT (IOEL)

Background

Indicative occupational exposure limit values are health-based, non-binding values, derived from the most recent scientific data available and taking into account the availability of measurement techniques. They set threshold levels of exposure (with corresponding reference time period) below which, in general, no detrimental effects are expected for any given substance after short term or daily exposure over a working life time. They are European objectives to assist employers in determining and assessing risks, e.g. in accordance with Article 4 of Directive 98/24/EC. IOELs are adopted at EU level by Commission Directives.

IOELs have been set for around 100 substances. Information on the setting of the IOELs and the recommendations of scientific committee on occupational exposure limits (SCOEL) are available on the website of Directorate General for Employment, Social Affairs and Equal Opportunities (DG EMPL): http://ec.europa.eu/employment_social/health_safety/index_en.htm)

Registrant's obligations

When an EU IOEL exists the registrant may, under conditions as described below, use the IOEL in place of developing a DNEL. Alternatively the registrant should, in accordance with the requirements of REACH, derive a DNEL following the steps outlined in the hazard assessment section of REACH Annex I.

A registrant is allowed to use an IOEL as a DNEL for the same exposure route and duration, unless new scientific information that he has obtained in fulfilling his obligations under REACH does not support the use of the IOEL for this purpose. This could be because the information obtained is more recent than the information that was used to support setting the IOEL at EU level and because it leads to another value being derived which requires different risk management measures (RMMs) and operational conditions (OCs).

When the registrant has obtained new scientific information which indicates that the IOEL does not provide the appropriate level of protection required by REACH, then the registrant should develop a DNEL based on this new information whilst also taking account of the scientific information that was used as a basis for the adoption of the IOEL. In this case, the registrant may wish to provide details of the new scientific information to DG EMPL who will take this into consideration as part of the normal procedures for reviewing IOELs. In any case, this data will be submitted to the European Chemicals Agency in the framework of registrations of substances ≥ 10 tonnes/year/registrant and can be retrieved there for such regulatory purposes as well.

When the registrant is using a substance in a way that leads to other exposure routes or exposure durations than the exposure route and duration on which the IOEL is based (typically derived for

inhalation exposure over 8 hours per working day (TWA) and/or short term exposures, typically of 15 minutes duration (STEL)) or if other human populations are exposed, the relevant DNELs should be derived. For example, in the case when the use may lead to dermal or oral exposure of the population at large or vulnerable sub-populations, DNELs to cover these situations will be required.

The use of the IOEL in place of developing a DNEL, or the derivation of a DNEL when there is already an IOEL, has to be documented in the registrant's Chemical Safety Report.

EU BINDING OCCUPATIONAL EXPOSURE LIMIT (BOEL)

Background

BOELs reflect socio-economic and technical feasibility factors in addition to toxicological information taken into account when establishing IOELs. BOELs have been set for 4 substances. Information on the setting of the BOELs is available on the website of DG Employment, Social Affairs and Equal Opportunities (DG EMPL):

http://ec.europa.eu/employment_social/health_safety/index_en.htm)

Registrant's obligations

When a BOEL exists the registrant cannot use it in place of a DNEL without an evaluation of the scientific background for setting the BOEL. Consequently, information and toxicological evaluations of health effects used for setting the BOEL may, as for IOELs, be used and taken into account in deriving the DNEL.

NATIONAL OCCUPATIONAL EXPOSURE LIMITS

Background

For any substance for which an IOEL has been established at Community level, Member States must establish a national OEL which takes account of the IOEL and is in accordance with national legislation and practices. The IOEL transposed into national legislation can be higher or lower than the EU IOEL.

For any substance for which a BOEL has been established at Community level, Member States must establish a corresponding national BOEL. The national BOEL can be lower than the Community BOEL, but not higher.

Finally, Member States may set national OELs for other substances than those included in Community legislation. Various approaches may be used; in some cases the OELs are purely health based values and in other cases they may take into account feasibility factors. In total, national OELs have been set in various Member States for around 600 substances (in addition to the substances with EU IOELs and BOELs).

Registrant's obligations

A registrant cannot use a national OEL in place of a DNEL without an evaluation of the scientific background for setting the national OEL. However, in cases where toxicological information and evaluations of health effects used for setting the national OEL are documented and available, this may, as for IOELs, be used and taken into account in deriving the DNEL. In this evaluation, the approach used for setting the national OEL should be compared to the approach for deriving DNELs as described in the in the main body of this chapter, and any differences in approach should be taken into account.

IMPLICATIONS FOR DOWNSTREAM USERS

All employers (regardless if they are registrants or downstream users under REACH), are also responsible for compliance with any existing national OEL for the substance at their own facilities.

A downstream user receiving an Exposure Scenario from his supplier should implement the OCs and RMMs recommended for ensuring control of risks, unless he chooses to perform his own chemical safety assessment (CSA).

If the downstream user decides not to implement the OC and RMMs (e.g. if he has not informed his supplier about his use) described in an Exposure Scenario supplied to him, he will have to develop a CSA for his own use and notify the Agency. In developing the CSA, the downstream user may use the DNEL provided by the supplier or he may choose to follow the above guidance for setting a DNEL based on EU and national OELs or set the DNEL in accordance with Annex I of REACH.

APPENDIX R. 8-14 Evaluating carcinogenicity risk levels; a review of decision points

Evaluating carcinogenicity risk levels; a review of decision points that are used or have been discussed in some different countries, organisation, and committees.

Decision points discussed for the general population

The World Health Organisation (WHO) considers in their drinking water quality guidelines for genotoxic carcinogens a lifetime cancer risk for consumers of less than $\mathbf{10^{-5}}$ to represent a so-called tolerable risk (1).

In connection with the EU Air Quality Directive and the EU Drinking Water Directive a $\mathbf{10^{-6}}$ lifetime risk is used as a starting point for the derivation of limit values for the general population (2).

In the US, risks lower than $\mathbf{10^{-6}}$ is in general considered to constitute an acceptable risk for the general population (3).

In the EU risk assessments of industrial chemicals carried out under Regulation 793/93 the limited experience points towards a decision point in the order of $\mathbf{10^{-5}}$ to $\mathbf{10^{-6}}$ for the general population (Table R.8-26).

In a CSTEE (EU Scientific Committee) opinion on methylene chloride (4) the CSTEE has expressed that a lifetime cancer risk of $<\mathbf{10^{-5}}$ is the generally accepted decision point for an acceptable risk to the general population.

In the context of remediation of contaminated sites (soil), some countries have expressed specific decision points. Canada uses a lifetime cancer risk of $<\mathbf{10^{-5}}$ as the decision point for acceptable risk (i.e., "essentially negligible cancer risk") (5). In France, $\mathbf{10^{-5}}$ is the target value, and a risk of $\mathbf{10^{-4}}$ (or higher) is unacceptable for contaminated sites (6). In the US, risks lower than $\mathbf{10^{-6}}$ is the goal, but this target may be adjusted up to $\mathbf{10^{-4}}$ depending on site-specific circumstances (7).

In summary, the cancer risk decision points used for *lifetime* exposure of the general population are generally in the range of $\mathbf{10^{-5}}$ to $\mathbf{10^{-6}}$.

Decision points discussed for workers

To achieve adequate control of risks in occupational settings, the Carcinogens and Mutagens Directive (2004/37/EC) requires effective risk management to prevent workers exposure to carcinogenic and mutagenic chemicals. Where it is not technically possible to prevent exposure then, by implementing the concept of the principles of good occupational hygiene practice, it should be reduced to as low a level as technically possible. Namely, through substitution, reduction of exposure, use in closed systems etc.

In addition, some EU Member States have applied lifetime cancer risk estimates in judging tolerable risk levels for workers. For instance, a lifetime cancer risk of $4\cdot\mathbf{10^{-5}}$ (which corresponds to 10^{-6} per working year, assuming 40 years employment) is the starting point in setting occupational limit values in the NL (by the Health Council), although this level may be proposed to be (temporarily) adjusted upwards (with $4\cdot\mathbf{10^{-3}}$ as un upper limit) depending on economical or technical reasons (by the Social Economic Council) (8).

UK HSE has expressed that they believe that an individual risk of death of one in a million per annum for both workers and the public corresponds to a very low level of risk and should be used as a guideline for the boundary between the broadly acceptable and tolerable regions (9).

Switzerland has used lifetime cancer risks in regulating asbestos, with a risk less than $\mathbf{10^{-3}}$ being tolerable, and less than $4 \cdot \mathbf{10^{-5}}$ being acceptable (10). The US Occupational Safety and Health Administration considers that a lifetime cancer risk for workers higher than $\mathbf{10^{-3}}$ represents an unacceptably high risk and their goal is to reduce this risk to less than $\mathbf{10^{-5}}$. However, many US industrial workplace standards (such as those of the American Conference of Governmental Industrial Hygienists) use values $\leq \mathbf{10^{-3}}$ as the threshold for tolerable/acceptable risk (11).

In the EU risk assessments of industrial chemicals carried out under Regulation 793/93 some genotoxic carcinogens have been assessed. The Technical Meeting of MS representatives under Regulation 793/93 agreed that a conclusion of concern should be drawn for all genotoxic carcinogens and the magnitude of the risk for each exposure scenario described as far as possible. In some cases quantitative risk estimates were included (either in the main body or as an annex to the risk assessment report) to assist in describing the risk[21]. It can be deduced from some of these reports that the cut-off between concern and low concern or residual risk is in the region of $\mathbf{10^{-5}}$ and $\mathbf{10^{-6}}$. The decision point for a few selected risk assessments can be found in Table R.8-26. The EU Scientific Committees CSTEE/SCHER have reviewed the first four reports below, and agreed with the conclusions of these reports.

In summary, the decision point for 'acceptable' *lifetime* (i.e., a working life of 40 years) cancer risk levels used for workers are generally around $\mathbf{10^{-5}}$ but higher or lower levels have been considered to be tolerable under certain circumstances.

[21] During the revision of the Technical Guidance Document for new and existing substances in 2005 guidance on quantitative risk assessment for non-threshold carcinogens was introduced proposing a linearised risk estimate approach know as the T25 method, as well as a margin of exposure (MoE) approach using a large extrapolation factor.

Table R. 8-26: Experiences of decision points used in cancer risk assessments of industrial chemicals in the EU

Chemical	Year of RAR	Conclusion
o-anisidine	1999/ 2000	For consumers, a lifetime cancer risk $\leq 10^{-7}$ represents low concern, whereas $\geq 10^{-6}$ represents concern
aniline	2001-2004	Considering a likely overestimated occupational risk (it is uncertain whether it is a genotoxic mechanism), 10^{-4} is used as decision point for immediate action (i.e., risk reduction).
2,4-toluenediamine, 4-methyl-m-phenylenediamine	Draft 2006	The decision point for workers is a lifetime cancer risk of 10^{-5} for very low concern or concern
2,3-epoxypropyltrimethyl-ammonium chloride	Draft 2005	A life-time cancer risk for workers of 10^{-6} and higher leads to concern.
benzene	Draft 2006	The decision point for workers is a lifetime cancer risk of 10^{-5} for low risk situations ('low concern'). No decision point is set for the general population.
2-nitrotoluene 4-nitrotoluene	Draft 2007	A life-time cancer risk for workers/humans of less than 10^{-5} is considered a low concern
2,4-dinitrotoluene	Draft 2007	A life-time cancer risk for workers/humans of less than 10^{-5} is considered a low concern. For the general population $>10^{-5}$ is concern, whereas $<10^{-7}$ is of low concern.

REFERENCES

1. Guidelines for drinking-water quality, WHO, 2006.

2. The Drinking Water Directive (DWD), Council Directive 98/83/EC

3. Federal Register, USA, Vol. 63, No. 157, 1998.

4. Opinion on TNO report "Methylene chloride: Advantages and drawbacks of possible market restrictions in the EU". Opinion adopted at 17th CSTEE plenary meeting, 5 September, 2000.

5. Federal Contaminated Site Risk Assessment In Canada. Part 1:Guidance on Human Health Preliminary Quantitative Risk Assessment (PQRA), Appendix B, Health Canada.

6. Emmanuel Normant, French Environment Ministry, http://www.clarinet.at/policy/

7. Role of the Baseline Risk Assessment in Superfund Remedy Selection Decisions, US EPA, 1991.

8. Jolanda Rijnkels, Health Council of the Netherlands, Presentation at DG Employment and ACSH Workshop, October 2006.

9. Reducing risks, protecting people. HSE's decision-making process (http://www.hse.gov.uk/risk/theory/r2p2.pdf)

10. Schweizerishe Unfallversicherungsanstalt: Grenzwerte am Arbetsplatz 2005. Schweiz 2004.

11. American Conference of Governmental Industrial Hygienists (ACGIH). TLVs and BEIs. 2002.

APPENDIX R. 8-15 Use of human data in the derivation of DNEL and DMEL

Introduction

Human data have been used in the risk assessment of many chemicals; e.g. under the Existing Substance Regulation (EEC 793/93), for setting Occupational Exposure Limits (OELs), and for cancer risk assessment by the International Agency for Research on Cancer (IARC). The merits of the use of human data as a source of hazard information are that the mode of action is usually relevant, i.e. the same mechanism applies as in the larger population that should be protected and no inter-species safety or assessment factors are needed. In addition to that, human data in most cases originate from exposure levels comparable to those in the target population and relate to a pertinent route of exposure.

Under REACH human data, when available and relevant, are used in the human hazard assessment and as part of the Chemical Safety Assessment as described in Annex I of the REACH Regulation. Furthermore, according to the provisions of Annex XI of the REACH Regulation "Historical Human Data"[22] can be used to adapt the standard testing requirements of Annexes VII to X provided that the quality of the data is properly assessed and judge to be appropriate.

The human health hazard assessment as defined in Annex I of REACH comprises four steps, the last one of which is the derivation of DNELs:

Step 1: Evaluation of non-human information.

Step 2: Evaluation of human information.

Step 3: Classification and labelling.

Step 4: Derivation of DNELs.

The above-mentioned steps 1 to 3 shall be undertaken for every effect for which information is available. Step 4, derivation of DNELs, shall be undertaken by integrating the results of steps 1-3. If justified by the exposure scenario(s), a single DNEL may be sufficient. However, taking into account the available information and the exposure scenario(s) of the CSR it may be necessary to identify different DNELs for each relevant human population (e.g. workers, consumers) and possibly for certain vulnerable sub-populations (e.g. children, pregnant women) and for different routes (oral, inhalation, dermal) and durations (acute, long-term) of exposure.

It may not always be possible to derive a DNEL for an end-point. This may be the case when a substance exerts its effect by a non-threshold mode of action (e.g. carcinogenicity through genotoxic mechanism) or the threshold cannot be reliably identified (e.g. sensitisation and irritation). For such carcinogens, and assuming that there are data allowing it, the registrant may develop a DMEL as stated in Section R.8.1.1.

The process of deriving a DNEL/DMEL from human data is divided in nine phases as shown in Figure R.8-6 below. Each of these phases is described in detail in the following sections. Phases 1 to 5 are identical for DNEL and DMEL derivation and therefore both derived effect levels are dealt with together. However, Phases 6 to 8 differ significantly in case a DNEL or a DMEL is derived.

[22] Historical Human Data is a term used by REACH. It refers to already available human data.

For these phases a different section has been developed for DNEL and DMEL with the specificities of each case. In Phase 9 all effect levels (derived either from animal data (AD) or from human data (HD)) are brought together for selection of the critical DNELs/DMELs to be taken to the risk characterisation.

Phases 1 to 5 are common for all types of human data, independent of the type of the effect. However, some specific guidance for DNEL derivation from human data on acute toxicity, irritation/corrosion and sensitization are given in APPENDICES R.8-8, R.8-9, and R.8-10.

For using purely qualitative human data, which is not within the scope of this Appendix, the reader is referred to Sections R.8.6 and Guidance E.

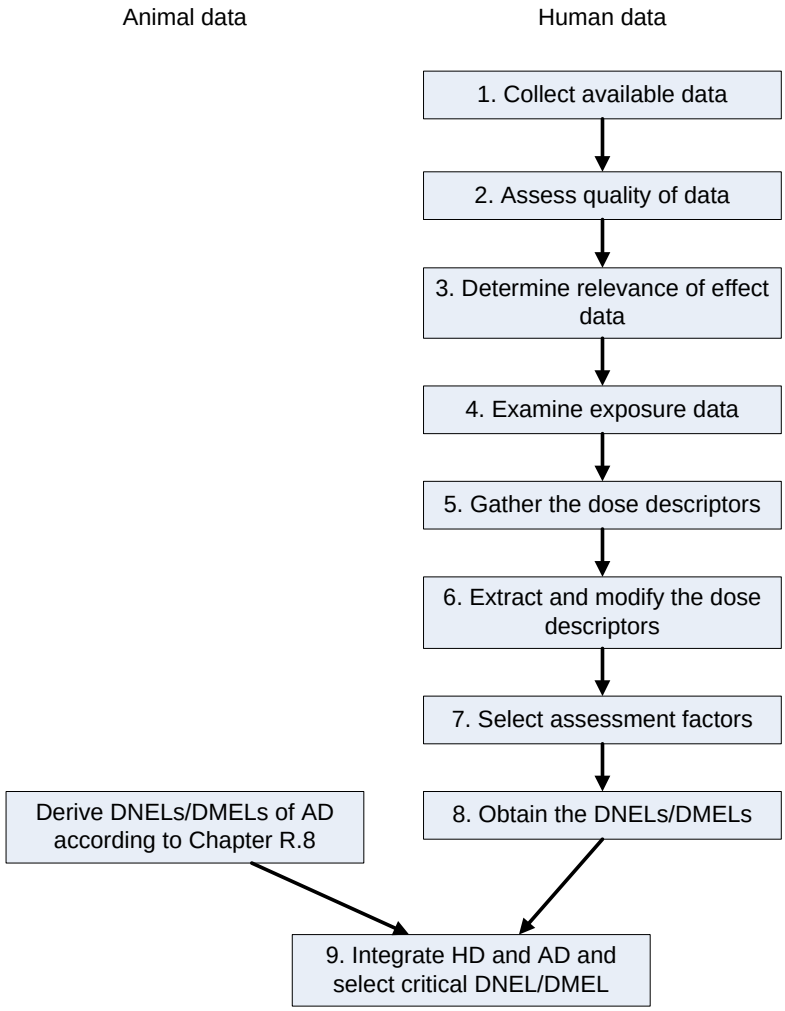


Figure R. 8-6 Illustration of the process for DNEL/DMEL derivation from human data.

The use of human data for risk characterisation requires a high level of expertise in relevant scientific fields (e.g. epidemiology, industrial hygiene, risk assessment). It is not possible to provide simple and explicit guidance on every detail; this guideline rather aims to provide the general framework as a basis for expert judgment. This guidance has been made as practical as possible bearing in mind the needs of an "average" user. In order to make the structure and logic of the guidance as clear as possible, each section includes a short reminder of what should be available when the phase is completed. With these reminders the user of the guidance can make sure that the phase is completed and adequate recording has been done.

Interpretation and assessment of epidemiological and other human data necessitate a good understanding of the inherent methodological issues. It is recommended that an epidemiologist by training or another person with relevant expertise on use of human data is involved in the assessment at least in the following situations:

1) the derivation of DNEL/DMEL includes an extensive amount of epidemiological studies or other human data

2) the derivation of DNEL/DMEL is mainly based on use of human data

3) human data not previously published in peer-reviewed scientific journals are being applied

4) the derivation of DNEL/DMEL includes adjustments or conversions of the original exposure information or statistical re-processing of original data in order to derive the necessary quantitative exposure estimates.

(Phase 1) Collection of available data

In this section the types of human data that can be used for derivation of the DNEL/DMEL and their availability are briefly described. Assessment of the quality of these data and evaluation of exposure data are covered in the relevant sections below.

The focus of many parts of this guidance is on the epidemiological data. It may be the most reliable type of human data but it would not be available for a very large number of chemical substances. When there are epidemiological data e.g. on carcinogenicity, reproductive toxicity or organ toxicity, before they are used, they should be carefully evaluated for their relevance in the overall risk assessment, including setting the DNEL/DMEL. The basic approach to the use and evaluation of human data is described in Chapter R.4 for all main types of human data, i.e. analytical epidemiology studies, descriptive or correlation epidemiology studies, medical surveillance data and case reports (in very rare justified cases also existing controlled studies in human volunteers). As underlined in Chapter R.4, a weight of evidence (WoE) approach is essential for risk assessment based on epidemiological and other human data. The availability, sources, use and value of human data vary according to the effect. Some endpoint-specific features of human information and its sources as well as evaluation of the human data and information requirements to be fulfilled are outlined in Chapter R.7, in the respective effect-specific Sections R.7.2. to R.7.7.

Published epidemiological data can be identified by searching Medline or other Life Science and Biomedical databases. Human data other than epidemiological studies can come from e.g. case reports, clinical studies, occupational disease registries or other occupational surveillance schemes and from poison centre information. In principle all types of toxic effects can be reported in such studies; however, in many cases they address acute and/or local effects. These data can be obtained e.g. from open literature by searching the relevant publication databases (see above), from occupational health units or occupational medicine clinics. Industry sources of unpublished monitoring and surveillance may be important although further efforts are needed to improve uniformity and harmonisation or these data (Ecetoc 2004). IPCS has made an attempt to demonstrate the feasibility of collecting harmonised human data from poisons information centres on a multi-national basis (IPCS 2005a, Ecetoc 2004), which increases the possibility of having these data available in the same form from different countries and with a sufficiently large number of cases. The use of harmonised reporting formats and terminology, developed originally by the IPCS (INTOX Data Management System) has been suggested for aggregating poison centre information. Finally, data collected under other EU or international regulatory processes, e.g. medicines or biocides, could be useful for some chemical substances.

Major chemical companies often have a routine medical surveillance system in place to monitor and manage employee health. Periodic routine medical examinations are offered to employees and these data are maintained in medical files in order to perform clinical practice and provide good quality clinical consultation. Apart from general medical surveillance programs, targeted programs exist to monitor employee health of those with potential exposure to certain chemicals. For example, employees with potential for exposure to benzene may be invited to participate in frequent medical examinations focused on the potential health outcomes such as changes in haematological parameters. This kind of data can also be useful either for setting the DNEL or as an element of the weight of evidence assessment.

The collection of data should be done in a non-discriminatory way as regards the nature of observations made in the individual study. This means that neither negative nor positive results should be preferred in the data collection phase. It is known that the studies with positive results tend to be more easily published than those without an observed effect. Even though statistical methods like the funnel plot have been developed to identify publication bias in meta-analysis

(Egger et al 1997) it is difficult to completely overcome this problem when searching the published literature. On the other hand published studies have undergone a review process that aims to ensure a high quality. When using unpublished data (e.g. company surveillance data) an objective approach to collecting data is equally important.

Result of Phase 1: The existing human data on the substance have been collected from the relevant sources. A list of studies has been prepared. It may be useful to indicate for each item in the list: the reference indicating the type of source (published article, published review or report, unpublished data) and the type of health effect observed.

(Phase 2) Assessment of the quality of the human data

Historical human data, such as epidemiological studies on exposed populations, accidental or occupational exposure data and clinical studies, shall be considered in this phase. The value of the data for a specific human health effect depends, among other things, on the type of analysis and on the parameters covered and on the magnitude and specificity of the health effect and consequently the predictability of the effect. Annex XI of the REACH Regulation sets out the following general criteria for assessing the adequacy of the data in view of using human data for the adaptation of the standard testing requirements:

1) the proper selection and characterisation of the exposed and control groups;

2) adequate characterisation of exposure;

3) sufficient length of follow-up for disease occurrence;

4) valid method for observing an effect;

5) proper consideration of bias and confounding factors; and

6) a reasonable statistical reliability to justify the conclusion.

It is recommended to use the same assessment criteria for the use of human data in the DNEL/DMEL derivation.

Epidemiology is the study of how often and why diseases occur in different groups (Coggon et al. 1997 bmj-online). Comprehensive guidance on both evaluation and use of epidemiological evidence for risk assessment is provided by Kryzanowski et al. (WHO 2000). Specific guidance on criteria to assess the quality of epidemiological studies can be found in text books of epidemiology (Checkoway et al. 2004, Hernberg 1992, Rothman and Greenland 1998).Chapter R.4 and Chapter R.7 of the Guidance on Information Requirements and Chemical Safety Assessment also describe some general and effect-specific quality aspects. The most important points to be addressed when assessing and documenting the above-mentioned six adequacy criteria are described below. Please note that the focus in this section is on the assessment of the quality of epidemiological studies. Although many of the quality assessment aspects (e.g. reliability of exposure and outcome information) are similar to other types of human data, there are also important differences, for example when case reports are assessed. This is further discussed later on in this Phase. APPENDICES R.8-8 to R.8-10 also contain useful information on the use of human data concerning acute toxicity, irritation and sensitisation.

1) The proper selection and characterisation of the exposed and control groups

Most commonly epidemiological studies are categorised into cohort (longitudinal), case-control (case-referent) and cross-sectional studies. Due to their non-experimental nature they are almost invariably subject to bias of one sort or another. A crucial aim of the study design is to keep bias to the minimum. Various types of bias are of concern (Sackett 1979) and some modern epidemiological domains like molecular epidemiology have their special considerations of bias (Vineis and McMichael 1998). Nevertheless the main types of bias to be avoided are selection bias (e.g. the so called healthy worker effect[23]) and information bias (e.g. recall bias). Both the selection and recruitment of the study subjects and the collection of information concerning their exposures and diseases should be done in such a way that they do not introduce bias in the difference in the disease rates between the exposed and unexposed groups (cohort design) or the difference in the occurrence of exposure between the cases and the controls (case-control design). In a cross-sectional design, an important aspect is the temporal relationship between exposure and health effect. Many types of effects are known not to occur immediately after exposure and original members of the study population may leave the study during the latent period of the health effect. Cross-sectional studies are considered to be suited to study acute effects or effects that do not lead to serious overt disease that would result in affected subjects leaving the exposure environment. A good epidemiological study describes and justifies the selection and recruitment procedures of the study subjects (exposed/unexposed cohorts or cases and controls), sampling, number of study subjects, non-response/non-participation, completeness of follow-up and the measures undertaken to ensure a comparable ascertainment of exposure and disease status between the different study groups.

2) Adequate characterisation of exposure

The relevant exposure parameter (mean level, peak level, duration, cumulative dose) depends on the health outcome and exposure setting and should be justified. In the analysis exposure can be considered as a continuous or categorised variable and the choices made should be described and justified. Multiple measurements will increase the accuracy of the exposure information especially in long-term exposures or when the exposure variation is high (see Phase 4). Errors in measuring exposure can be an important source of bias in epidemiology, especially concerning exposures further in the past and mixed exposures. In a good epidemiological study the methods to assess exposure are clearly described and their validity is assessed in case previously non-validated methods are used. Furthermore, exposure data should be quantitative and exposure categories well defined, in order to allow a DNEL/DMEL to be obtained from that study.

Comparing results from several studies may necessitate conversion of exposure parameters. Sometimes more qualitative exposure categories need to be converted to quantitative estimates by using exposure data from other sources. Combining results of several studies may involve additional statistical modelling and new analyses. In all the above situations the choices and compromises made should be clearly described and their potential impact preferably assessed by sensitivity analyses using varying parameter values. The uncertainty resulting from such

[23] The healthy worker effect results from three main phenomena: the healthy hire effect refers to the fact that the healthiest individuals are the most likely to get a job and the healthy worker survivor effect refers to the healthiest individuals being the most likely to continue in the job. The third phenomenon refers to changes in life that are related to the employment.

approximations needs to be assessed and may need to be taken into account when applying the assessment factors (see Phase 4 and Phase 7).

3) Sufficient length of follow-up for disease occurrence

The latency time between exposure and the occurrence of the health outcome varies from less than one day (e.g. acute irritation) to decades (e.g. most malignant diseases). A good epidemiological study describes and justifies the length of follow-up in longitudinal studies and the exposure time windows used in case-control studies.

4) Valid method for observing an effect

Measuring disease occurrence in populations requires diagnostic criteria. For practical reasons the criteria used in epidemiological studies usually differ at least somewhat from the criteria used in clinical practice. Many health outcomes represent more of a continuum rather than a dichotomous phenomenon and therefore standard predefined criteria need to be established for classifying the disease status of the study subjects. The types of health effects for which human data exist vary from acute effects to long term effects. The occurrence of health effects can be determined in various ways. Self filled-in questionnaires, clinical examinations or queries from already existing databases (e.g. causes of death databases or cancer registries) can be used. The quality of the health effect data depends on the data collection methods used. Standardised and validated data collection or diagnostic techniques with satisfactory sensitivity and specificity should have been used. It is crucial that the reliability of health effect data collection techniques is the same for the exposed and non-exposed populations. Errors in measuring disease status can be an important source of bias in epidemiology. In a good epidemiological study the diagnostic criteria used are clearly described and their validity is assessed in cases where previously non-validated criteria are used.

For a given chemical substance, several health effects may be relevant. The studies addressing these different health effects may have used quite different methods to ascertain the occurrence of that health effect. Each of these should be assessed for the validity in order to be sure that adequate information is available in the later phases of the DNEL/DMEL derivation process.

5) Proper consideration of bias and confounding factors

Bias should be minimised in the study design and the biases that cannot be avoided should be identified, assessed for their potential impact and taken into account when interpreting the results. Confounding occurs when the exposed and non-exposed populations have different background disease risks. Confounding could occur, for example, if the difference in the occurrence of the health effect results from differences in the age or gender distributions or life style rather than from differences in the chemical exposure between the exposed and unexposed populations. If taken into account in the study design, the effect of confounding factors can be controlled in the analysis. A good epidemiological study describes the confounders that could be controlled in the analysis, and how they could be controlled , and estimates the potential impact of confounders that could not be controlled.

6) A reasonable statistical reliability to justify the conclusion

Usually the results of epidemiological studies are expressed in terms of risk estimates characterising the difference in risk of disease between the exposed and unexposed populations. Parameters like relative risk (RR), odds ratio (OR), standardised incidence or mortality rate (SIR or SMR) are used. More generally, the statistical analysis aims to describe the study results of individual observations as meaningful numerical values (mean, median etc.). Control of confounding is also an important element of the statistical analysis.

Even after biases have been minimised and confounding controlled, these mean values or risk estimates may be unrepresentative just by chance (random error). In general terms the statistical precision of the risk estimate, i.e. the narrowness of its confidence intervals, is inversely proportional to the number of cases observed, i.e. the larger the study size the more precise the risk estimate in statistical terms. In epidemiology, it is preferable to base the statistical inference on the point estimate of the risk and its confidence intervals. P-values from statistical testing are also used. Nevertheless, p values have inherent problems due to being dependent not only on how much the result deviates from the so called null hypothesis, but also on the sample size in which this deviation is observed. In addition to the point estimate and its statistical confidence intervals, the actual methods of statistical analysis and the statistical models used should be known and justified. In addition to the statistical precision and significance, the internal consistency of the results is an issue (e.g. did all the analyses support an association, was it observed in all subgroups, was there a dose-response relationship?).

Although the above six quality criteria are formulated especially from the point of view of the assessment of the adequacy of a single epidemiological study, they can be applied to assess the adequacy of the relevant aspects of other types of historical human data, e.g. accidental exposure data and occupational exposure data (medical surveillance data). For example major chemical companies have implemented routine health surveillance programs to monitor and manage employee health. Although the information is not specifically collected for hazard assessment purposes, it can provide a useful database provided that the relevant points of the above-mentioned six quality criteria are met, i.e. unpublished surveillance data should meet the same quality criteria as published reports. It is especially important to address and document in a report the attendance rates and other selection mechanisms, quality of the health effect and exposure data, confounding and biases and the proper analysis of data.(see von Elm et al. 2007 for guidance on proper reporting of observational human studies). A drawback often reducing the usefulness of company surveillance data is the lack of an unexposed comparison group that would have been followed with an identical protocol.

Case reports can also provide crucial qualitative information or even quantitative information on the dose descriptor, but they generally pertain to less complex situations and instances where the causality of exposure and effect are immediately obvious. This is the case for example, when the health effect is acute, specific or preferably both. The quality criteria above can still be applied for the relevant parts. The quality criteria for study design do not apply as such but bias and confounding should, nevertheless, be considered and even very simple but reliable exposure information and health outcome information can suffice.

Bias and confounding are challenges in human data and therefore an important issue when assessing the quality of the data. Their potential influence, nevertheless, varies depending both on the specificity of the exposure-effect relation and on the latency period between exposure and effect. Therefore, in practice the quality requirements for specific effects (e.g. asbestosis) are different from those for multifactorial effects (e.g. lung cancer). The same applies for the difference between acute effects and long-term effects. These differences are also reflected in the requirements concerning the study design. Swaen (2006) and ECETOC (2009) have produced guidance on the quality aspects and use of information according to the type of effect and type of study. APPENDICES R.8-8 to R.8-10 also contain useful information on the use of human data concerning acute toxicity, irritation and sensitisation.

In all cases adequate and comprehensive enough documentation shall be provided to justify the use of human data in the derivation of DNEL/DMEL.

Result of Phase 2: Each human study has been characterised for its quality. Some studies with borderline quality, for example from the point of view of the quantitative nature of the exposure data, may be taken to the weight of evidence analysis, depending on the expert judgment. The studies with inadequate/low quality are unlikely to be of significant value in the DNEL/DMEL derivation.

(Phase 3) Evaluation of the relevance

An assessment of the likelihood of a **causal association** should be made for all endpoints or health effects identified in Phases 1 and 2. The causality assessment in epidemiology and its relationship with causality considerations in other fields of science has been extensively discussed (see e.g. Checkoway et al. 2004, Hernberg 1992, Rothman and Greenland 1998). The most often cited guidance on causal inference are the criteria described by Hill (1965). These criteria should not be used in a stringent manner in the sense that they all must be met. A too stringent causal inference approach will lead to false negative conclusions. A too loose application of the Hill criteria will lead to false positives. A practical approach to application of Hill's criteria with an analysis based on actual data has been introduced by Swaen and Amelsvoort (2009. As underlined in Chapter R.4, a weight of evidence approach is essential for risk assessment based on epidemiological data. The specific features of evaluation of human information are outlined in Chapter R.7 in the respective effect-specific Sections R.7.2. to R.7.7.

Due to its non-experimental nature, human data, unlike animal experiments, very seldom relate to exposure to a pure, clearly defined chemical substance. **Confounding from concomitant other exposures** as well as non-specific characterisation of the chemical substances in question are common challenges that need to be assessed when judging the relevance of the human data.

Epidemiological studies may differ in the extent to which they are focussed on testing a specific hypothesis. Studies targeting testing of a specific hypothesis with a specific exposure, a specific effect and an *a priori* specified statistical analysis protocol should be given more weight than studies with a more exploratory character. In general studies designed to test a specific hypothesis tend to have more extensive and reliable quantitative exposure information than studies with a more general hypothesis or studies with an exploratory aim.

The evaluation should include an assessment of whether **the available human information addressing the endpoints of interests is sufficient and consistent** with the tonnage driven data requirements necessary to fulfil the REACH obligations, or whether the knowledge provided by the human information still presents data 'gaps'. For example, in the case where sperm quality has been analysed in a group of male workers, it would be incorrect to regard that study as covering all the reproductive toxicity parameters. In many cases there are gaps in the human studies. Nonetheless, animal data often complement human data. The whole database available (animal and human data) should be sufficient to address the endpoints compliant with the data requirements specific for the tonnage level. Where the human data set is incomplete, in terms of covering all the relevant effects, a DNEL/DMEL should not be established based upon that data alone, when additional relevant animal data is available. In case the study is incomplete but of sufficient quality it should be taken to Phase 9 below.

Result of Phase 3: Each relevant finding in the human studies has been characterised for its relevance, i.e. the degree of certainty concerning the causal relationship between the exposure and effect, completeness in terms of coverage of relevant effects, and association of the effect with a specified chemical in the case of multi-exposure.

(Phase 4) Examination of the exposure data

Exposure data have already been addressed as one element of the quality assessment of Phase 2. In Phase 4, quantitative measures of exposure are identified, which are later used in extraction of the dose descriptors In principle, only human studies with quantitative information on exposure are useful in the process of setting a DNEL/DMEL. If the exposure data are instead of a descriptive and qualitative nature, the study results can be relevant for hazard identification, but usually not for quantitative risk assessment; (however, see Section on "Quantitative exposure data by modeling" below.). The exposure information varies according to the type of the study. In case reports or studies with a limited number of individuals, as for acute effects, the exact doses and other exposure characteristics may be known. In epidemiological studies on long-term effects, data are often less accurate and can contain relatively old measurements or other exposure data, which are difficult to validate. Mixed exposure to several agents is often a problem, which needs to be taken into account in the study design, and in the analysis and interpretation (see Phases 2 and 3).

Exposure conditions can vary substantially. The number of exposure measurements needed depends on the variability of the exposure conditions. If exposure is stable, with no significant variation over the workday, the season or between time periods, a few sample points can be adequate to characterise the exposure situation. However in reality, exposures usually vary from place to place and, from task to task. They may change over time (short-term and long-term) due to differences in production process, exposure reduction measures, and use of personal protective equipment.

The type of exposure information also varies. Sometimes the only known exposure parameter is that a person has been employed in a particular industry. More specific information would be the type of job the person has been doing in that industry and over which time period. Quantitative exposure characterisation can be made if industrial hygiene measurements are available. In general, industrial hygiene measurements can be done for various purposes. They can be done e.g. to identify sources of release or tasks with high exposure. In the latter case the results constitute an overestimate of general exposure at the workplace. Industrial hygiene measurements can also be conducted to provide a reliable picture of the exposure conditions at a specific work place. If the exposure measurements are collected by means of a systematic approach they are more valuable. It should be clear under which circumstances samples have been taken.

The precision of exposure measurements in estimating true exposure is not only determined by the number of measurements but also by the variability of exposure. Two aspects of exposure data are important for final interpretation of the findings. First, the internal validity should be satisfactory, meaning that the exposure data adequately describe the actual exposure situation. Internal validity depends on the sampling strategy and sampling frequency. Second, external validity should also be satisfactory. It relates to the comparability between the exposure conditions under investigation and the exposure conditions in other situations.

Money and Margary (2002) described a number of core principles to derive reliable and robust exposure assessments. They essentially describe three types of exposure data: actual data, analogous data and predicted exposure data derived from suitable validated models collected in a systematic manner. All three types of data can vary in quality and reliability.

Exposure data in a human study can be e.g.

- Measured data, which refers to ranges/categories of exposure (e.g. 0-10 ppm; 11-50 ppm, above 50 ppm). If these ranges are very wide, they may not be adequate for obtaining a DNEL/DMEL. When the sampling strategy and validity of exposure data is documented, it

can be used for obtaining the DNEL/DMEL. How dose descriptors are derived from exposure categories is explained in Phase 5 below.

- Qualitative exposure categories such as "no exposure", "low exposure", "medium and high exposure". As such this kind of data is not useful for setting a DNEL/DMEL, because no quantitative measure can be extracted from it.
- In case where biomonitoring values are available, where specific biomarkers can be clearly associated with the effects observed, they can be taken as dose descriptor. More guidance on the use of biomonitoring data in DNEL/DMEL derivation is given in Section R.8.1.2.7 and APPENDIX R.8-5.
- Measured analytical values associated both with effective dose/concentration and non-effective dose/concentration; e.g. for irritation, corrosion or, in some rare cases, for sensitisation. When representative and valid, this data can be used for obtaining a DNEL/DMEL.

Quantitative exposure data by modelling

In case qualitative exposure categories have been used in the original study, it may be possible on a case-by-case basis to obtain a quantitative estimate of exposure. More notably, there may be sufficient information in the human study on those exposure parameters that are needed for modelling the exposure. For example, the modelling tools referred to in Chapters R.14 and R.15 can sometimes be used. More sophisticated exposure modelling tools can also be used (IPCS 2005b). Exposure modelling in epidemiological or other human data requires specific expertise.

The information on the exposure parameters for modelling may be available from the same human study, or from a different study describing the operational conditions on that particular sector of industry or in those work tasks. Expertise on occupational hygiene is necessary for evaluation of relevance and reliability of this type of secondary data source.

It is emphasised that compensation of missing quantitative data by modelling results should only be done when that human study (with qualitative exposure data) is of good/sufficient quality. In case there are also other concerns (in addition to the exposure data) in relation to the quality, the study should not be used for setting the DNEL/DMEL. The aim of the modelling exercise is to "upgrade" or complement those human studies, which provide crucial evidence on a health effect that cannot be identified or adequately assessed based on other available animal or human studies. Thus, only studies with high relevance should be subject to this type of *ad hoc* exercise in the risk assessment.

Use of a job-exposure matrix (JEM) can in certain cases provide a reasonably robust tool for assessing the quantitative exposure levels linked to job titles or complete job title histories (Hoar 1983). In longitudinal epidemiological studies the construction of a job-exposure matrix has been shown to be a valuable means of using exposure information. It is only useful to construct a job-exposure matrix if (semi-)quantitative exposure information is available. The job-exposure matrix is based on homogeneous exposure groups, consisting of those jobs that are thought to be characterised by comparable exposure conditions. For each homogeneous exposure group the exposure intensity is estimated. Historical changes in the production process or work practices, resulting in changes in exposure, are taken into account and form a dimension of the matrix. The job-exposure matrix allows calculating cumulative exposure, but can also serve to stratify the study groups into subgroups with certain exposure characteristics, such as those exposed at least once over a certain concentration.

It is strongly recommended that in case modelling or application of a JEM is used *a posteriori* in order to generate the missing quantitative exposure data, that is done by an occupational hygienist or a similar expert, who has comprehensive knowledge of the relevance and use of various exposure parameters and is familiar with the modelling tool or JEM, which is used. The parameter choices made should be justified and documented in a clear and transparent manner. Sensitivity analyses should be performed in order to assess the effect of the parameter choices made and to justify the validity of the choices.

In human studies other than analytical epidemiological studies (case reports, medical surveillance data, etc) valuable quantitative exposure information can also be generated from more qualitative data (e.g. job titles, occupational histories). In such study designs a JEM is not needed, while access to the company information or more general industrial hygiene information concerning exposure levels is required (if not already used in the report available).

Result of Phase 4: The respective exposure levels (concentration or dose) have been assigned to each relevant health effect observed in the human studies. Additional information e.g. on the physicochemical properties of the substance, the pattern of use and the work tasks performed which affect exposure should be recorded here, as appropriate. It is recommended that the type of each exposure data is indicated: e.g., "measured worst case scenarios/high exposures", "representative measure data" or "modelling data". The studies which only have qualitative exposure data, and for which modelling cannot be applied are identified as such. They are put aside for the next phases and when relevant and crucial, considered again in Phase 9 in the weight of evidence analysis.

(Phase 5) Gathering the dose descriptors

Provided that according to Phases 1 to 4 there are sufficient human data on a health effect associated with exposure to a certain chemical, the most reliable dose descriptors for each health effect will have to be identified.

If for a given health effect, only one human study with quantitative dose descriptor information has been identified in Phases 1 to 4, the selection of the dose descriptor is straightforward, i.e. the dose descriptor is the exposure concentration or dose that has been assigned to that health effect observed in the human study. It is to be underlined that, if overall only one human study is available, the quality of the data needs to be of high standard in order to justify its use, unless animal or other test data give supporting evidence (see Phase 9). In case several studies are available for a single health effect the most reliable should be selected using a weight of evidence approach. In case comparable and good quality data are available from more than one study, a summary measure could be created by a meta-analysis or pooled analysis. A weight of evidence approach is essential for risk assessment based on epidemiological data to (a) assess (sources of) heterogeneity across the studies and (b) increase statistical stability of the risk estimates. A meta-analysis of published studies or a pooled analysis of original raw data provides the best basis for deriving an overall dose-descriptor. Meta- and pooled analyses can also take into account small studies, which - on their own - are not suitable for deriving dose-descriptors due to statistical instability. If a good summary of all evidence is not available, using an individual relatively large study may be an acceptable, but statistically less accurate alternative (in comparison to a meta-analysis or pooled analysis using all the studies). For some substances, a dose descriptor on the dose-response curve may be derived from a single good quality epidemiology study, if this is the only adequate study. Once the dose descriptors have been gathered, they should be collected in a table (see Table R.8-14 of APPENDIX R.8-1). Please note that table R.8-14 will often contain dose descriptors expressed in ranges/categories of exposure.

This is because epidemiological studies often relate the health effects to exposure categories. These categories are further processed in Phase 6.

Before actually deriving DNELs/DMELs on the basis of the derived dose descriptors, it is important to determine whether the substance exerts its effect by a threshold or by a non-threshold mode of action.

If the substance exerts its effect by a threshold mode of action, one or more DNELs will have to be derived for the different threshold endpoints, based on the most relevant dose descriptors for these endpoints. For non-threshold effects, for which in principle any level of exposure carries a risk, one or more DMELs could be derived instead (if data allow) on the basis of their most relevant dose descriptors.

The mechanism to derive DNELs differs substantially from the DMEL approach. For this reason, Phases 6, 7 and 8 of the DNEL/DMEL derivation process are described separately for each effect type (.i.e. threshold or non-threshold).

Result of Phase 5. The dose descriptors derived from relevant human data have been gathered (see Table R.8-14 in APPENDIX R.8-1). A decision has been made on the substance mode of action (threshold or non-threshold).

A.- DNEL DERIVATION FOR THRESHOLD EFFECTS

(Phase 6-A) Selection and modification of the relevant dose descriptors

6.A.1 Selection of the relevant dose descriptors

For threshold effects, i.e. health effects induced only above a certain exposure level, the aim is to find a NOAEL or LOAEL, more or less analogous to the procedure using animal data.. Some differences exist in the nature of such data from experimental vs. observational studies. In experimental data (e.g. animal tests) the doses are predefined and concern more or less exact values apart from each other (e.g. 100 mg/kg, 300 mg/kg, 1000 mg/kg) and the experiment usually does not provide observations on the occurrence of effects between these values. The highest dose without observed effect is identified as the NOAEL and the lowest dose with an observed effect as the LOAEL. The true threshold apparently lies somewhere between the two values. In the case of observational (human) data with exposure categories forming a continuum, i.e. the next category normally starting from where the previous category ended (e.g. 0- 5 ppm, 5 – 25 ppm, >25 ppm), the situation is a bit different from animal data. As NOAEL and LOAEL are based on absence or occurrence (e.g. a statistically significant increase in RR) of an "observed effect", the NOAEL category would be the highest category where an effect is not yet observed and LOAEL category as the lowest category where an effect is observed. Nevertheless, the true threshold value does not lie between the categories (as they form a continuum) but in one of the categories.

As a consequence, many study reports with quantitative exposure and effect data do not directly allow to establish the exact NOAEL, but only to approximate the exposure range the NOAEL lies in. If the exposure categories form a continuum, the upper exposure limit of the range of exposures in the no-effect category is the same as the lower limit of the range of exposure in the lowest category showing an effect. In the absence of more details on the distribution of exposures this value (i.e. the boundary between the two categories) should be used as a point estimate of NOAEL. In cases where the number of individuals in the NOAEL category is small or if there is indication that the exposure distribution is skewed towards the lower end of the category, a more conservative NOAEL may be justified. In that case, the average of the lower and the upper limit value of the NOAEL category could be used, based on expert judgement that should be explained in the dossier. If sufficient data are available, the average exposure of the individuals or the median exposure value of the NOAEL group may be a better choice for describing the exposure of the group.

This procedure applies to acute as well as long-term health effects (i.e. effects with a longer latency period). In case only a LOAEL can be identified, that value should be carried to the next Phases with an indication of the fact that it is a LOAEL and an adequate assessment factor should be applied in Phase 7. Using LOAEL with an assessment factor should also be considered when the identification of the "NOAEL category" is uncertain, for example if the number of observations in the "NOAEL category" is low.

The above considerations are written from the point of view of epidemiological studies. In acute specific effects (e.g. irritation), the data are sometimes reported also as simple frequency data over exposure ranges. The identification of NOAEL and LOAEL could then follow a similar reasoning. Case reports may also contribute to defining LOAEL values.

In case there was no effect at any of the exposure ranges, the study should not be used for derivation of the NOAEL, because there is no need to set a DNEL. However, see section on negative studies under Phase 9.

6.A.2 Modification , when necessary, of the dose descriptors to the correct starting point

In a few situations, the exposure situation from which the dose descriptor is obtained is not directly comparable to the exposure situation for which the DNEL is being derived in terms of exposure route, units and/or dimensions. In these situations, it is necessary to convert the dose descriptor into a correct starting point. This applies to the following situations:

1. If epidemiological data derive from another exposure route than the route to which the risk assessment has to be applied, a route-to-route extrapolation is necessary.
2. Differences in exposure conditions between the source population and the target population, e.g. differences in respiratory volumes, or intermittent versus continuous exposures etc.

In principle situation 1 above is rare in the case of human data as these data more or less by definition deal with a route of exposure relevant to humans. Nevertheless exposure routes of consumers and those exposed in the occupational setting (often the origin of studies available) may differ. If needed, the principles described in Section R.8.4.2 apply to both situations 1 and 2 above. It should be noted that modification is usually not needed in cases where human exposure is evaluated based on biological monitoring data (internal dose metric). If valid human data that relate the effect directly or indirectly to a biomonitoring metric are available, the calculation of $DNEL_{biomarker}$ values can be straightforward. See APPENDIX R.8-5 for more guidance on the derivation of DNELs using biomonitoring data.

After modification where necessary of the relevant dose descriptors, the corrected starting points should be collected in a table (see Table R.8-15 of APPENDIX R.8-1)

Result of Phase 6–A: For each threshold effect, one or more dose descriptors have been selected. The dose descriptors, after modification (if required), are collected in a table (see Table R.8-15 APPENDIX R.8-1)

(Phase 7-A) Selection and justification of the Assessment Factors

In the use of human data for DNEL derivation, assessment factors (AFs) associated with e.g. intraspecies variation, dose-response relationship and differences in exposure conditions are considered. Contrary to the case with experimental animal findings, there is no need to consider interspecies variation when using human data for DNEL development. Where human data are considered a suitable starting point for the derivation for a DNEL, then a partly similar set of considerations can be identified as those applied to experimental data (Section R 8.4.3.1 and Table 8-6). These aspects will be discussed under the following headings:

1. intraspecies differences;
2. differences in duration of exposure;
3. issues related to dose-response;
4. quality of human data

7.A.1 Intraspecies Differences

When human studies are used for derivation of the DNEL, intraspecies assessment factors are needed to account for the variability in the target population, which can be anticipated to be usually larger than that in the study sample. For example, there can be differences in toxicokinetics due to

slower excretion of the substance or due to a higher rate of transformation of the parent substance to a more toxic substance in some sub-populations, but also due to differing absorption rates or due to differences in toxicodynamics. The source population may comprise only/mainly healthy workers and the target population may include also e.g. sensitive target populations such as very young children, elderly people and persons having diseases (e.g. diabetics, people with kidney diseases). This would mean in practice that usually in the target population there are more sensitive people than in the source population and therefore the effect level in the target population could be significantly lower. Obviously, workers may develop diabetes, cardiovascular and other diseases just like the general population, but it is generally acknowledged that selection of workers will lead to a worker population, which is either healthier and/or more resistant to the physical and chemical stress factors of the work. Sensitivity of the human sub-population should be taken into account when establishing the DNEL. For example the possibility of higher sensitivity of children and pregnant women should be considered.

7.A.1.1. Selection of assessment factors

Human studies normally cover at least some of the human inter-individual variability. Use of AFs strongly depends on the human data that is available for obtaining the DNEL. Therefore, a pragmatic approach is described below for using appropriate intraspecies assessment factors, which are based on the specific human studies available. Adequate justification of selection of any AF should always be given.

In case specific intraspecies assessment factors cannot be justified with the human data available, the values of Table R.8-6 in Section R.8.4.3.3. should be used.

7.A.1.2. Use of standard assessment factors

According to Table R.8-6, the standard intraspecies assessment factors would be 5 for workers and 10 for the general population. These AFs are the same as those applied when using animal data as a starting point.

A standard assessment factor would be appropriate when the human study is small and the sample in the study is homogenous and therefore no significant part of human variability could be regarded as covered.

Examples of cases, where the use of standard AF is necessary are when:

(i) there are **one or two case studies/reports** with low number of individuals observed, or

(ii) there is **a small occupational surveillance study** with a sample of 10-20 workers who might have been selected so that healthy worker effect applies.

7.A.1.3. Deviation from standard assessment factors

It should be always examined whether there are substance specific data to justify deviation from the standard assessment factors. Some cases, where this deviation could be justified are specified below.

(i) In some cases, e.g. when a high inter-individual variation of susceptibility has been identified, assessment factors higher than the standard assessment factor may be needed. This could be the case for example when genetic polymorphism leads to a high variation in the level of the metabolizing enzymes. Those rare cases where an unusually large assessment factor is necessary for protection of children are described in Section R.8.4.3.1.

In case the human study is small and/or the sample is only representative of a particular sub-population, it should be considered whether that study should be used for the derivation of the DNEL. Obviously, a well conducted and relevant human study should not be rejected only, because the sample size is small Use of expert judgment by the registrant is necessary when that type of study is used for setting the DNEL. See a relevant example on hydrogen peroxide in APPENDIX R.8-16.

(ii) Use of AFs lower than the standard assessment factors is appropriate when it can be shown that some of the **factors that cause the intraspecies variation** in the target population, such as gender, age, nutritional status, health, susceptibility and genetic polymorphism have been covered in the study population. When this is the case, a value lower than the standard assessment factor should be selected and justified based on expert judgment.

(iii) In some cases, **substance specific information** might be available that can be used to justify special assessment factors. This information could be from **toxicokinetic and/or toxicodynamic studies** where variation in the human population has been measured. For example, when measurements in sufficient number of humans have shown that toxicokinetic and toxicodynamic factors, taken together, can be accounted by an AF between 2 and 5/10, that value can be used instead of "standard" or "lower" AFs. It should be acknowledged that the number of substances for which this information is currently available seems limited. It is also noteworthy that when substance specific information is obtained from studies where the sample size (number of people) is small (10-30), it is not justified to set a low AF, since the effects of human variability cannot be fully observed in a study with a relatively small sample size. In principle, the intraspecies variability for workers can be addressed in a smaller study sample, in comparison with a study that aims to cover the human variability in the general population. Guidance for the use of substance specific data and some examples are provided in the IPCS document 'Chemical-specific adjustment factors for interspecies differences and human variability' (IPCS, 2005c). Furthermore, as specified in APPENDIX R.8-4, **PBPK modeling data** can aid in the quantification of intraspecies variability, which may be caused by variation in anatomical, physiological and biochemical parameters with age, gender, genetic predisposition and health status. PBPK models can be used to quantify these, which would result in possible modification of AFs. IPCS is about to finalize guidance on these issues (for the progress of the project see http://www.who.int/ipcs/methods/harmonization/areas/pbpk/en/index.html).

(iv) There can be cases where the sample (i.e. the population) in a good quality human study is so heterogeneous and well characterised for different "aspects" of intraspecies variation that the use of a lower than the standard AF, i.e. 1-2, is justified. In the current experience, the number of substances with that kind of human data is not high. When e.g. the **sample size adequately takes into account the frequency of the effect and the study group is heterogeneous and the surveillance/study has an adequate duration**, it may be concluded that most of the intraspecies variation has been covered in that study. In an optimal case, justification for a low assessment factor could be based for example on a description of the demographic data of the study sample, such as age distribution, gender and diseases. In addition , if the NOAEL is obtained from a study where a susceptible group of people has been specifically addressed (e.g. the registration of respiratory effects in a group of persons with asthma or hyper-reactive airways) a reduced intraspecies AF may be more appropriate. In cases where e.g. children, elderly or sick people or people having a special diet were not represented or were excluded from the study sample, the use of a low AF would not be justified.

Use of low AFs should be considered on a case-by-case basis and it is acceptable only when this is supported by appropriate data, which is given in a transparent way. In APPENDIX R.8-16 an

example is given of a substance, for which the human database enables the use of low intraspecies assessment factors.

(v) When the effect seen in humans is associated with **biomonitoring** data such as urinary or blood level of a compound or its metabolite, again the toxicokinetic factor of intraspecies variation is accounted for and the AF could be the remaining 3.16 for the toxicodynamic variability. Biomonitoring data reflect the internal exposure and thus, toxicokinetic parameters influencing the internal/systemic bioavailability do not play a role. The factor of 3.16 would then be applied to the dose descriptor obtained from the biomonitoring study. (See also APPENDIX R.8-5)

See APPENDIX R.8-16 for examples of modification or deviation from the default intraspecies assessment factors.

7.A.1.4 Study Size

It is not possible to define minimum and/or recommended size of the study population, since (i) it will depend on the study type/aim and (ii) because study size as such does not provide assurance that sensitive subpulations and factors causing variation in the human population have been covered in the source study. Thus it would not be scientifically justified to give accurate sample sizes (to cover variability in the target population). It is notable that for e.g. substances covered under Existing Substance Regulation (ESR) (see APPENDIX R.8-16), for which the hazard evaluation primarily relied on human data, the epidemiological studies (for carcinogenicity or organ toxicity) generally had large sample sizes, i.e. thousands or tens of thousands of individuals. Smaller studies have been used in some of the ESR cases, for the purposes of hazard identification and classification and labeling.

Therefore, sample size of the source study should be considered together with an evaluation of whether the different "factors" that cause intraspecies variation have been addressed in the study. This means in practice that e.g. even large worker surveillance, when done in a homogenous group of workers, does not cover intra-species variation among general population. The homogeneity of the worker population could, for example, result from the healthy worker effect which can play a role both at hire and during the career (see Phase 3). Similarly a sample of general population that is limited to a region next to a source of the release of a substance, might be rather homogenous e.g. in terms of dietary habits, and ethnic background.

The issue of study size can also be addressed within the overall study quality (see Phase 2). In case the sample size is so small that it compromises the statistical power of the study or the reliability of the study, it would usually be more appropriate to use other data (for setting the DNEL) instead of that study, rather than try to correct the weakness of the study by additional AFs.

7.A.2 Duration of exposure

Provided that the information in a human study covers a sufficient time span, there is generally no need to introduce an assessment factor to account for differences in the duration of exposure for the study population and scenario under consideration (target population). Instead, the duration of the human study should be compared with the exposure situation in the target population. When considering the time span aspects of a study, one should check that from the point of view of the end-point studied, it covers both a sufficient duration exposure time and a sufficient follow-up time to observe the effect. The time of follow-up should be sufficient to cover the latency period between the exposure and the manifest disease or organ damage. This is especially important in case of cancer epidemiology and chronic organ toxicity studies

Whether the duration of a human study can be considered sufficient will also depend on the type of effect under consideration. Acute effects, e.g. effects on the central nervous system caused by solvents or skin and eye irritation/corrosion can usually be observed within a few days. The reversibility of the effects can also be observed during some more days. Development of most signs and symptoms of target organ toxicity can usually be observed if the exposure duration has been several years. A longer follow-up and/or a longer duration of exposure is necessary when carcinogenicity is studied. This is due to the latency period, which depending on the type of cancer, can even be of many decades.

If a reliable NOAEL for a chronic endpoint is available, this is the preferred starting point for a $DNEL_{long\text{-}term}$ and no assessment factor for duration extrapolation is needed regardless of whether the information is applied to workers or consumers. In some cases, the duration of exposure falls between traditional acute and chronic studies (such as depression of blood counts following days/weeks of exposure, i.e. they are observable effects of possible pre-clinical significance and serve as a surrogate measure for serious effects) and where a $DNEL_{long\text{-}term}$ must be derived. In such cases, an AF of 2 is suggested to be applied to the NOAEL (and which is consistent with past practice in this area[24]). A NOAEL for an acute endpoint (NOAEL following short term exposure only) should not be used as the basis for the derivation of a $DNEL_{long\text{-}term}$. If the study design does not allow to adequately address any latency of the observed effect, then these data should not be used for deriving a $DNEL_{long\text{-}term}$.

7.A.3 Dose-response relationship

The dose-response relationship and the shape of the dose-response curve for the endpoint of interest are important elements to be considered in the derivation of the DNEL.

As with animal data, consideration should be given to the uncertainties in the dose descriptor (NOAEL, benchmark dose…) as the surrogate for the true effect threshold, as well as to the extrapolation of the LOAEL to the effect threshold (in cases where only a LOAEL is available) and the extent and severity of the effects at the LOAEL.

Unlike in experimental animal data, in many human studies the response/effect will not be displayed at discrete exposure concentrations but within exposure categories/ranges. The cut-off points of these categories should be set in a transparent and scientifically sound manner instead of trying to create exposure categories so that a favourable result (i.e. a high NOAEL/LOAEL) can be obtained. Exposure categories and ranges of the original study should usually be kept.

In case the exposure data comes from e.g. worker surveillance studies or from case studies, the data would normally be more accurate than just wide exposure categories. It is recommended that the reliable exposure measures are compiled and that those, which best represent the conditions where a health effect was identified are used when setting the dose descriptor.

It is proposed that in the absence of more detailed information the lower boundary/limit of the lowest exposure category in which the most sensitive effect is still observed should be considered the LOAEL and the upper boundary/limit of the exposure range in which no statistically or biologically significant effect is observed should be considered the NOAEL. For further consideration of NOAEL and LOAEL values see also Phase 5 and Phase 6.A.1.

---

[24] Steven Fairhurst (1995) "The Uncertainty Factor In The Setting Of Occupational Exposure Standards" Ann. Occupational Hyg., 39: 375 - 385.

Some of the uncertainties associated with the reliability/accuracy of the dose-response relationship of a substance, such as dose/exposure spacing, group sizes and statistical methods, cannot be dealt with using formalised assessment factors. These uncertainties have to be addressed qualitatively. In cases where the uncertainties are major, the study should not be used for derivation of the DNEL. See more guidance on these issues e.g. in Phase 2 (data quality). The only major uncertainty in the dose-response relationship that is traditionally addressed with the application of assessment factors is the extrapolation of the LOAEL to the NOAEL when only a LOAEL is available.

It is proposed that when the starting point for the DNEL calculation is a LOAEL, an assessment factor ranging from 3 (as minimum/majority of cases) to 10 (as maximum/exceptional cases) is applied. An AF of 3 may be more appropriate for instance in situations, where the effects at the LOAEL are mild, or the LOAEL represents the lower boundary of the exposure range in which the effect is observed. Higher numerical values should be considered in situations where the effects at the LOAEL are severe and irreversible, or the shape of the dose-response curve is shallow or the quality of the study (e.g. group sizes, statistical methods, study design, exposure data) gives rise to uncertainties about the reliability of the identified LOAEL. It is especially important to apply a high assessment factor to a shallow dose-response curve, when dealing with incidence data[25]. This is because a large decrease in dose is needed to make sure that no individual is affected by a serious effect , e.g. cancer.

Case reports e.g. on acute effects must be evaluated within their context, i.e. against the background information concerning the exposure levels in this setting at the time when no cases were reported. A report of an unusual case of a disease or health effect would indicate that earlier exposure was below the NOAEL and that the exposure of the reported case is above the NOAEL. From such a report it can be concluded that the NOAEL must lie between the normal exposure condition and the unusual exposure condition leading to the induction of the effect. Appropriate assessment factors need to be applied in these instances and the $DNEL_{acute}$ will generally lie between the normal exposure condition and the exposure level responsible for the effect. If a serious health effect is reported (other than irritation or rash for example) the application of an assessment factor could lead to a DNEL below the normal exposure condition.

7.A.4 Quality of human data (including exposure data)

In principle, significant flaws concerning the quality criteria set in Phase 2 will lead to rejection of that individual study in the process of setting a DNEL Application of an additional AF may be necessary when a relevant and valuable set of human data with limited quality is used. For example data from a human study with qualitative (e.g. job title and factory information) or semi-quantitative exposure parameters was converted into quantitative exposure data with the help of some additional external exposure information. This conversion of exposures was, nevertheless considered to contain so much uncertainty that the study would not have been used as “a stand alone piece of information”. Nevertheless further qualitative data from other sources gave additional support for this study and it was concluded being the best available basis for derivation of the DNEL. In such a case it should be considered to add an additional assessment factor for quality of human data when using the dose descriptor from this study in the derivation of DNEL.

[25] In some studies you have yes/no data and a dose-response characterised by incidence (e.g. for carcinogenicity), whereas in other studies there is continuous data and a dose-effect relationship (e.g., RDT-data such as the effect of cadmium on urinary protein excretion).

As described earlier, also combining results of several studies may involve additional statistical modelling and new analyses. In such situations the choices and compromises made should be clearly described and their potential impact preferably assessed by sensitivity analyses using varying parameter values. Typically the uncertainty resulting from such approximations needs to be taken into account, case-by-case, when setting assessment factors.

7.A.5 Overall assessment factor and its application to the correct starting point

The overall AF is obtained by simple multiplication of individual assessment factors discussed in the previous paragraphs. Care should be taken to avoid double counting several aspects when multiplying the individual factors.

Result of Phase 7-A: Assessment factors for intraspecies variation, duration of exposure, dose-response and quality of the human data are assigned to each dose descriptor. The justification of the assessment factors is documented. The overall assessment factor for each dose descriptor is calculated by simple multiplication of the individual assessment factors of that dose descriptor.

(Phase 8-A) Obtaining the DNEL

This phase describes how the DNELs are obtained from dose descriptors and assessment factors.

Once the relevant dose descriptors have been selected for each endpoint and modified to the correct starting point (see Phase 6-A) and the overall assessment factor calculated for each of them (see Phase 7-A) an endpoint-specific DNEL will be derived by dividing each dose descriptor by its overall assessment factor.

All derived DNELs are collected in a table (see Table R.8-16 in APPENDIX R.8-1). In case there are more than one DNEL per endpoint, all of them are taken to Phase 9.

Result of Phase 8-A: For each relevant dose descriptor selected in Phase 6, a DNEL is calculated. This is done by dividing the dose descriptor by its overall assessment factor. All DNELs are summarised in a table (see Table R.8-16 in APPENDIX R.8-1) and taken to Phase 9.

B- DMEL DERIVATION FOR A NON THRESHOLD CARCINOGEN

"When no DNEL can be derived, the registrant has to conduct "a qualitative assessment of the likelihood that effects are avoided when implementing the exposure scenario" (REACH Annex I, Section 6.5). No DNEL can be derived for non-threshold mutagens/carcinogens as it is assumed that a no-effect-level cannot be established for these substances (either because there is no threshold or the threshold level cannot be determined). In such cases, and assuming that there are data allowing it, the registrant should develop a DMEL (derived minimal effect level), a reference risk level which is considered to be of very low concern. A DMEL derived in accordance with the guidance should be seen as a tolerable level of effects and it should be noted that it is not a level where no potential effects can be foreseen.

Contrary to the case for the risk assessment for threshold effects, by definition for non-threshold mutagens and carcinogens a dose without a theoretical cancer risk cannot be derived. Therefore the establishment of a reference risk level for the DMEL clearly is of societal concern and needs policy guidance. Although there is no EU legislation setting the 'tolerable' risk level for carcinogens in society, cancer risk levels have been set and used in different contexts (see APPENDIX R. 8-14 for various values previously applied within and outside the EU)[26].

Two quantitative risk assessment formats can be followed to derive a DMEL for a non-threshold carcinogen: the 'Linearised' approach, or the 'Large Assessment Factor' approach. Both formats are based on the same principal elements of risk extrapolation or risk evaluation using a dose-descriptor related to a risk estimate (an RR or a comparable measure such as an OR or an SMR). Because of different perceptions of the uncertainties involved in quantitative risk assessment and risk evaluation and of different approaches to risk communication, there may be a preference for one of these formats .

In the following sections of this document, Phases 6, 7 and 8 of the DMEL derivation process are explained in detail for the 'Linearised' approach. Additionally the general principles of the 'Large Assessment Factor' approach will be outlined. However, due to the lack of experience in the use of this method for the derivation of DMELs from human data, no further explanation will be provided on how to proceed on each phase.

The 'Linearised' approach

Some regulatory agencies including the US EPA, the Danish EPA and the Dutch Health Council basically follow this approach (US EPA, Danish EPA 2004, Dutch Health Council, 1989; *see also* Goldbohm et al., 2006). The aim, when using this approach, is to identify an exposure level that gives rise to a risk which is considered to be of very low concern. A review of carcinogenicity risk levels used or discussed by different organisations, countries and committees is given in APPENDIX R.8-14.

---

26 Should such EU legislation for setting a 'tolerable' risk level for carcinogens be developed in future, then DMELs need to be derived on that basis.

(Phase 6 B) Extraction and modification of the relevant dose descriptors

6.B.1 Selection of the relevant dose descriptors

For non-threshold effects, i.e. notably carcinogenicity through a genotoxic mechanism, the dose descriptor is usually derived from cohort or case-control studies reporting Relative Risks (RR) or comparable measures to describe a dose-response association. The RR is the ratio between the risk of the health effect in the exposed population divided by the risk in the unexposed population. Comparable measures are the standardised ratio, such as standardised mortality ratio (SMR) or standardised incidence ratio (SIR), which are conventionally used in cohort studies if the unexposed reference group is the general population. The odds ratio (OR), which is derived from case-control studies, is also a measure of relative risk. The dose descriptor of interest for derivation of a DMEL is the exposure level related to a RR (or comparable measure). In its most simple form, the dose descriptor represents the exposure level related to a relative risk observed in an exposed compared to an unexposed population. Ideally, it is based on the slope of the exposure-response function derived for the whole range of exposure levels observed in the study or based on pooled data from all available adequate studies by modelling. As default a linear relative risk model should be used. In this way, only a single RR per unit of exposure (i.e. slope factor) is obtained for a substance. Occasionally, a non-linear exposure-response model may be fitted to the data and used to derive the dose descriptor. When this is the case, the selection of the dose-response model should be clearly justified. A background and further explanation can be found in Goldbohm et al (2006).

6.B.2 Modification , when necessary, of the dose descriptors to the correct starting point

In a few situations, the exposure situation from which the dose descriptor is obtained is not directly comparable to the exposure situation for which the DMEL is being derived in terms of exposure route, units and/or dimensions. In these situations, it is necessary to convert the dose descriptor into a correct starting point. This applies to the following situations:

1. If epidemiological data derive from another exposure route than the route to which the risk assessment has to be applied, a route-to-route extrapolation is necessary.

2. Differences in exposure conditions between the source population and the target population, e.g. differences in respiratory volumes, or intermittent versus continuous exposures etc.

In principle situation 1 above is rare in the case of human data as these data more or less by definition deal with a route of exposure relevant to humans or a combination of routes. Nevertheless exposure routes of consumers and those exposed in the occupational setting (often the origin of studies available) may differ. If needed, the principles described in Section R.8.4.2 apply to both situations 1 and 2 above.

The exposure metric most often used in the analysis of the epidemiologic data is a cumulative exposure value including years of exposure, e.g. 'ppm-years'. For genotoxic carcinogens cumulative dose is thought to be the more relevant exposure metric than exposure concentration. Hence, a correction for *duration* of exposure is not needed if an adequate cumulative dose exposure metric is used.

It must be noted that in many instances epidemiological data on long term cancer risks from chemicals are derived from epidemiological studies on occupationally exposed cohorts. These risks need to be converted to continuous (24 hours per day, 365 days per year and 75 years long) exposure for the general population.

After modification, where necessary of the relevant dose descriptors, the corrected starting points should be collected in a table (see Table R.8-15 of APPENDIX R.8-1)

Result of Phase 6-B (Linearised approach): For each non threshold effect, one or more dose descriptors have been selected. The dose descriptors, after modification (if required), are collected in a table (see Table R.8-15 APPENDIX R.8-1)

(Phase 7-B) Selection and justification of the Assessment Factors

The next step in the derivation of a DMEL is to address uncertainties in the extrapolation of the study data to the real human exposure situation, taking into account variability and uncertainty. Clearly, the use of epidemiological data has advantages over the use of animal data since there is no need for interspecies extrapolation. Furthermore the extrapolation from high exposure (study data) to low exposure (level of exposure/risk of low concern) is usually done over a narrower range of exposure levels. Nevertheless, some assessment factors still need to be considered.

For DMEL derivation based on epidemiological studies, the following assessment factors will still need to be considered:

1. Intraspecies differences
2. Quality of the database (amount and quality of available information)

7.B.1. Intraspecies differences

Part of the population is suspected to be more susceptible to cancer due to differences in toxicokinetics (ADME) and to genetic properties (such as having specific polymorphisms). If it can be documented that these properties are equally distributed across the relevant population subgroups (e.g. workers vs. general population or age groups, men/women, healthy/ill), there is no need to use AFs for extrapolation of a DMEL derived from one subgroup (e.g. a worker population) to a DMEL for the general population. For non-threshold effects such as carcinogenicity through genotoxic mechanism, it is often assumed that different large population groups have similar susceptibility. Usually the human dose-response data is based on reasonably large epidemiological studies. However, exceptions may arise, e.g. if the human data are derived from populations with a different genetic background. The principles described in Phase 7 of DNEL Derivation (Section 7.1) should be applied to justify the selection of the intraspecies assessment factor. If there are data on some risk-related parameters that allow comparison of dose-response (relative risk estimates) between the general population and susceptible individuals, the additional analyses may be performed to adjust the general population estimate for susceptible individuals.

The 'Linearised' approach intrinsically takes into account that individuals may be exposed during different time periods and at different exposure levels during life. In many epidemiological studies, in particular occupational studies, cumulative exposure (cumulative exposure = exposure level * exposure duration, e.g. ppm years) is used as exposure metric. As the dose-response is, in general, considered to be constant over all age groups, extrapolation does not need an AF. However, exceptions to this general rule may be encountered. For example, it is known that breast tissue is more susceptible to unrepaired genotoxic damage in the period between menarche and first pregnancy, as during pregnancy breast cells differentiate. In this case, RRs may be higher depending on age. This should be solved by applying the life table method (see Phase 8), incorporating the higher RRs during adolescence and young adulthood and the lower RRs during the remaining life periods.

Duration of follow-up is addressed in Phase 2. It is assumed that only studies with a sufficient follow-up time will be used for DMEL derivation. Therefore an assessment factor for duration of follow-up is not needed.

7.B.2 Quality of the database

The quality of the individual studies available is assessed according to the criteria of Phase 2. When assessing the quality of the overall human data consisting of these individual studies, the same issues should be summarised. Especially the following should be carefully considered:

- The *amount* of available data, i.e. the size of the study (or studies) that are used when extracting the dose descriptor, determines the amount of random error in the risk estimates related to the dose descriptor. This uncertainty is usually represented by the confidence intervals that are routinely derived for such estimates. A pooled analysis or meta-analysis, when based on a substantially large database, has relatively small confidence intervals. An assessment factor larger than 1 may be applied if the selected risk estimate has wide confidence intervals (i.e. the uncertainty concerning the risk estimate related to the dose descriptor is large).

- Another source of uncertainty is derived from *uncontrolled biases* (e.g. confounding bias or healthy worker effect) in the data (see Phase 2). Evidently, data likely to be subject to serious bias should not be used for quantitative risk assessment at all. For example if selection bias or information bias (see Phase 2 point 1 and 5) could plausibly explain the main findings of the study. However, in less serious cases, the impact of a possible bias on the dose descriptor may be estimated[27] and compensated by an assessment factor.

- If there is reason to assume that the quantitative exposure-response relationship based on the epidemiological data is probably an underestimate or overestimate of the true association an appropriate assessment factor should be applied. An example of such a situation is when quantitative exposure estimates are lacking from a study and exposure level(s) were estimated from other sources to obtain a dose descriptor.

7.B.3 Overall assessment factor and its application to the correct starting point

The overall AF is obtained by simple multiplication of individual assessment factors discussed in the previous paragraphs. Care should be taken to avoid double counting several aspects when multiplying the individual factors.

Result of Phase 7-B (Linearised approach): Assessment factors for intraspecies variation and quality of the database are assigned to each dose descriptor. The justification of the assessment factors is documented The overall assessment factor for the dose descriptor is calculated by simple multiplication of the individual assessment factors of that dose descriptor.

27 A practical approach to assess the effect of possible uncontrolled biases on the risk estimate can be to apply sensitivity analyses postulating different levels of bias. A more sophisticated and reliable approach is to use probabilistic simulations to estimate bias, e.g. [Steenland and Greenland, 2004].

(Phase 8-B) Obtaining the DMEL

8.B.1 High to low dose extrapolation

The RR (whether or not corrected in Phase 6-B) must be projected onto the target population (workers or general population) to derive an Excess Lifetime Risk (ELR) at a given level of exposure. I.e. how many excess life time cases in absolute terms will result from a given relative estimate of risk (RR, OR, SMR or SIR) This necessitates the application of the relative risk estimates on actual population data (with person-year data and case occurrence data). There are two options to do this:

i) a simple direct method as described by van Wijngaarden and Hertz-Picciotto (2004) or the Dutch Health Council (1989), and

ii) a more sophisticated method including the use of a life table approach as described by e.g. Steenland *et al.,* 1998.

The direct method calculates the ELR as: ELR = Lifetime Risk * (RR-1). Lifetime Risk is the (background) risk of the relevant health effect in the target population to which the DMEL applies. The simple direct method results in some overestimation of the lifetime risk, in particular if the background risk in the target population is high. This is mostly because the direct method is less accurate in taking into account the mortality from other causes of death. The life-table method calculates and accumulates the ELR for each life year during the lifetime of a virtual cohort (see Goldbohm et al 2006 for an example). It gives a more accurate estimate and can incorporate specific requirements, such as changing exposure patterns over a lifetime, competing risks due to effects of exposure on other endpoints, etc. The life-table method may be used if there is a need to calculate the risk more accurately. A life table should also be used if age-dependent RRs are indicated (see example on breast cancer above). Although the life-table method is the preferred option, the direct method can be used if the background rate of the disease and the potency of the substance are low or if the age for which the risk is considered relevant is relatively young (< 70 years) (Goldbohm et al 2006).

At the end, the ELR estimate linked to the known level of exposure is used to extrapolate the exposure level that corresponds to a given risk level considered from a societal point of view to be of very low concern. i.e. the DMEL. A review of carcinogenicity risk levels used or discussed by different organisations, countries and committees is given in APPENDIX R.8-14. If the RR was calculated from a linear relative risk model, the derived ELR for a given exposure can directly be converted to a DMEL with a linear extrapolation.

If the RR value was based on models other than the linear model, low dose extrapolation should be performed according to one of the following two options. If there is additional evidence (e.g. based upon available experimental data of good quality) that the dose response outside the observable range is non-linear, a non-linear model may be used to assess the risks associated with these lower exposure levels. Otherwise, if there is no information on the shape of the dose-response in the low dose range, as a default, linear extrapolation should be applied. The application of a non-linear model to low dose extrapolation should be performed on a case-by-case basis, and should be extensively documented and justified.

8.B.2 Application of the assessment factors

The exposure level that corresponds to the chosen level of low concern (obtained as described above) is divided by the overall assessment factor (obtained in Phase 7-B) in order to calculate the DMEL.

The derived DMELs are collected in a table (see Table R.8-16 in APPENDIX R.8-1). In case there are more than one DMEL per endpoint, all of them are taken to Phase 9.

Result of Phase 8-B (Linearised approach): DMELs are calculated using human studies of sufficient quality and including adequate exposure data. This is done by extrapolating the study data to a risk level of low concern and dividing this by the overall assessment factor. DMELs are collected in a table (see Table R.8-16 in APPENDIX R.8-1) and taken to Phase 9.

The 'Large Assessment Factor' Approach

The 'Large Assessment Factor' approach was recently presented by the Scientific Committee of the European Food Safety Authority (EFSA SC) when providing guidance for managing risks posed by contaminants in food (EFSA, 2005). When applied to animal data, the approach uses a large assessment factor (10 000 or higher) in order to derive an exposure level of low public health concern from a BMDL10 from an animal study. EFSA SC notes that the benchmark dose approach can also be applied to human data when available. It is considered that presently there is not enough experience on the use of human data and on this approach in the hazard assessment of industrial chemicals and further guidance on such an application is difficult to give.

The method relies on a large assessment factor and the lower confidence limit of a BMDL10, the critical points in applying this approach to human data concerning a non-threshold carcinogen would be the selection of the dose descriptor equivalent to BMDL10 (the lower limit of the confidence interval of the dose related to absolute effect frequency of 10%), and the adjustments of the large assessment factor. Human data are usually described in terms of doses related to relative effect estimates (RR or similar). Apart from very rare situations (e.g. heavily asbestos-exposed worker groups) absolute effect frequencies of 10% are not observed. Therefore the approach would first necessitate an upward extrapolation to the 10% effect level and then a second downward extrapolation to the level of low concern. It is also not clear whether in the absence of the interspecies AF, the "large assessment factor" would anymore ensure a high level of protection. Therefore the use of this approach in the derivation of DMELs from human data needs to be well justified and special attention should be paid to ensure that a sufficient level of protection is reached.

(Phase 9) Integration of human and animal data and selection of the critical DNEL/DMEL to be taken to the risk characterisation

At the start of this phase, the DNELs and DMELs derived from human and animal data have been collected in Table R.8-16 (see Chapter R.8, APPENDIX R.8-1). Data in this table are then integrated in order to arrive at the specific entries of Table R.8-9. Please note that when human data are used for obtaining DNEL(s)/DMEL(s), the guidance given in this phase should be followed instead of that of Section R.8.7.1. Thus, this phase addresses the selection of the leading health effects and the critical DNELs/DMELs, which are subsequently used in the risk characterisation. Integration is based on the quality, relevance, completeness and level of the DNELs derived from different studies as explained below. The same principles and criteria are used, in case there are more than one human study at this phase (for certain endpoint), but no animal data. The selection of the critical DNELs/DMELs to be taken to the risk characterisation should be justified/documented..

The decision on which dose descriptor to use to derive a DNEL based on human or animal data is not straightforward and should be seen in the context of Mode of Action Framework (IPCS 2007a and b). Even when human data are not of adequate quality to derive the DNEL, consideration should be given on their potential use together with animal data for refinement of the risk characterisation. This can be the case for example in the development of PBPK modeling (see Section R.8.4.3.2 and APPENDIX R.8-4). IPCS is about to finalise guidance on these issues (see http://www.who.int/ipcs/methods/harmonization/areas/pbpk/en/index.html for the status of the project).

Within the REACH Regulation the so-called weight of evidence (WoE) approach is a component of the decision-making process on substance properties and thus an important part of the chemical safety assessment (see Chapter R.4). The term WoE neither constitutes a scientifically well-defined term nor an agreed formalised concept characterised by defined tools and procedures. It is based on assigning weights to each available piece of information either in an objective way by using a formalised procedure or by using expert judgment (see Chapter R.4.for more details). Although the use of structured frameworks can be invaluable in promoting harmonization in the assessment of chemical risks (IPCS 2007a), a commonly agreed formalised procedure is not yet available for integrating animal and human data for the various purposes of human health hazard assessment. The weight given to the available evidence will be influenced by factors such as the quality of the data, consistency of the results/data, nature and severity of the effects, relevance of the information for the given toxic endpoint. Tools like the Klimisch scores for experimental toxicity studies or Hills criteria for evaluation of epidemiological data are available for specific factors to be assessed, but the way in which the WoE is implemented to integrate all data remains case-dependent. Some guidance has been developed by IPCS, IARC and ECETOC and these are briefly described below.

The IPCS has developed a procedure, termed the IPCS Human Relevance Framework to make judgments about the relevance (to man) of findings in animal studies for both cancer endpoints and non-cancer endpoints (IPCS 2007a, IPCS 2007b). This procedure involves describing key events leading to the toxicity observed, and establishing the mode of action (MoA) in animals. Each key event in animals is then evaluated for its plausibility in man including both fundamental qualitative aspects and quantitative aspects. This procedure includes several elements that are useful when integrating human and animal data.

When evaluating the human carcinogenicity of agents, the IARC classifies the evidence into five groups: definitive carcinogen, probable carcinogen, possible carcinogen, not classifiable and probably not carcinogenic. The categorization of an agent is a matter of scientific judgment that reflects the strength of the evidence derived from studies in humans and in experimental animals and from mechanistic and other relevant data (IARC 2006). The human and animal evidence are each first classified into four categories: sufficient, limited, inadequate and evidence suggesting lack

of carcinogenicity. The information concerning a given mechanism of carcinogenicity is classified as weak, moderate or strong. A further assessment is then done on whether this particular mechanism is likely to operate in humans. When integrating the above-described categorized human, animal and mechanism information, specific rules are applied to arrive into one of the five groups of overall carcinogenicity.

A more formalised process of integrating the human and animal data would improve the utility and robustness of the risk assessment process. Such an approach has recently been proposed (ECETOC 2009). After having formulated the problem, the proposed method uses five separate category scores to characterise the available human data (for **quality**) and animal data (for **quality and relevance**). The scores are based on a collective weight of evidence assessment of the human data on the one hand and the animal data on the other hand. In case the human data are of equal or better score than the animal data, then human data takes precedence. Otherwise animal data takes precedence. Special care is taken when considering the concordance of animal and human data. In general positive data take precedence over null data. Nevertheless, negative data are also scored, with special emphasis on the confidence intervals of the negative studies. While the principles behind this method (e.g. using quality categories of human data) are appreciated, it is too early to recommend this approach as a systematic tool for the overall integration phase of human and animal data and for the specific purpose of deriving DNELs/DMELs.

9.1. Principles of integrating human and animal data

This phase of derivation of DNEL/DMEL should be transparent and rigorous and the decisions should be made based on the best data available.

The human data carried over to the integration phase include studies in which bias, confounding and chance are ruled out with reasonable confidence and in which the causal relationship between exposure to a specific chemical agent and a certain health outcome has been evaluated. These studies include both studies providing a quantitative estimate for the DNEL/DMEL and possibly studies which do not contribute directly to the quantitative assessment of the DNEL/DMEL, but are still judged to be of qualitative value in the WoE analysis. The available human studies have been assessed for their completeness as regards covering the relevant health endpoints.

The available data need to be assessed for their reliability and consistency across different studies (including available animal data) and endpoints taking into account the quality of the study protocol/methodology, size and power of the study design, biological plausibility, dose-response relationships and statistical association (adequacy of the database). When the human data are robust and of good quality, it should always be considered in this integrative step.

Where human data are inconsistent,they should not generally be used for DNEL derivation, although even in such cases the severity of the observed effect may indicate that the human findings should be considered within the WoE.

A particular point is the use of **negative human** studies. The term "negative" refers to studies where no effect was seen. A single human study can rarely, if ever, be regarded as negative in the sense that it proves an absence of relationship. If a human study is of compromised sensitivity or of compromised completeness,s it could rather be called “inconclusive” than “negative”. For example a study with a relative risk lower than unity, but with an upper limit of the confidence interval being above one, is rather inconclusive as a single study if no other human data are available.

Good quality human data that illustrate the lack of a health effect in a specific exposure range (negative overall data) should be considered in this phase of the process. Although the same quality criteria apply to negative studies as to positive studies, special care must be taken to ensure that the

negative outcome is not the result of inadequate sample size (statistical power),or design or measurement error or uncertainty of exposure or effect. In addition the completeness and adequacy of the negative human data for the purpose it is intended to be used should be ensured. In conclusion, a high level of quality is required of negative human data, especially when it is used in this phase to overrule positive evidence from one or more animal studies. The difficulty with negative human data lies in the fact that the only conclusion to be drawn is that the exposure range under investigation is below the effect threshold. Thus, the evidence does not allow estimation of the true effect threshold, but it can allow the inference that the effect threshold is higher than the exposure range investigated. This complicates the derivation of a health based DNEL in that the application of conventional AFs may result in unnecessarily strict exposure limits.

Positive, but non-statistically significant findings should not be regarded as null findings by default. In these instances biological significance is a more appropriate criterion than statistical significance.

9.2 Pragmatic approaches to integrating human and animal data

While the formalised methods to integrate human and animal data are not fully developed and "tested", more pragmatic approaches which rely on the current experience of the use of Weigh of Evidence analysis, can be useful. Some typical and/or challenging cases of integration are therefore described below. The main challenges of integrating animal and human data concern cases where the available data are inconsistent. This is particularly the case when there are both negative and positive data on the same endpoint/health effect and an unclear mode of action.

9.2.1 Inconsistent data

In cases where **different types of effect** are seen in human and animal studies the possible reasons for inconsistency should be assessed. The WoE analysis should start by characterising the available studies for their quality and relevance. This step should already have been done in the earlier phases of analyzing the human data. As regards animal studies, Klimish score can be taken as a measure of the quality and the relevance. The animal study/studies should be assessed according to the criteria given in Chapter R.4.

In case the human data is of sufficient quality and relevant and the animal data has a Klimish score of 1-2 and is also considered relevant, both human and animal data should be included in the WoE. In that case, DNELs should be obtained from the critical studies and the lower DNEL should be used for the risk characterisation. However, in case the effect/endpoint was not adequately addressed in the human study, use of the animal data is preferable.

It is important to consider why data is inconsistent. It may be due to a different mode of action in humans versus in animals, in which case human data should usually be preferred in the assessment. The human data can be negative simply because the effect observed in animals does not have any relevance for humans as the underlying mechanism is specific to the animal species used in experiment/test. If this can be justified the negative human data is the basis of further evaluation. This has been addressed below in a specific section on "*Negative human or animal data"*. Also a low exposure level and/or compromised study power in the human study may lead to an apparent inconsistency between human and animal data.

Inconsistency between human and animal data may also be due to significantly **lower sensitivity of humans** to the toxic effect (e.g. due to interspecies difference in toxicokinetics). Deriving DNELs/DMELs from both human and animal data and selecting the lower of these values is the recommended approach. However, the human study has to be of sufficient size and there needs to be an understanding of the relevant mode of action, before lower sensitivity in humans can be established and used in the risk assessment.

There may also be inconsistencies within the human data (i.e. between two human studies) that have to be assessed. In such cases it must be determined whether an explanation can be given for the diverging results (were they caused by different kinds of study designs with different sensitivity; different kinds of effect examinations or measuring techniques; not quite comparable groups; different methods for evaluating exposure etc.?). Inconsistencies may not necessarily weaken the evidence if there are good explanations for the diverging results.

9.2.2 Incomplete human data

As a starting point it should be realized that in most instances it is not possible to obtain comparable data sets in humans and experimental animals, e.g. histopathological data are usually not available in human studies. Also, in many cases human studies have not covered as high dose levels as the animal studies and therefore it will not have been possible to observe some relevant effects in the human study. Nonetheless, even incomplete human data can be relevant and should be used at this phase when it is of sufficient quality and gives quantitative information on the exposure.

In cases where the human study did not cover some specific endpoints and animal studies did, the animal study should primarily be used for setting the DNEL/DMEL. An example of this might be an epidemiological study where only a certain type of malformations in human were examined/studied, while both the developmental toxicity study and two generation reproduction toxicity study were carried out animals. In case the DNEL (for malformations) derived from the human study is lower than the respective DNEL from animal studies, DNEL from the human study should be selected for developmental effects. In addition it would be necessary to obtain and report the DNEL concerning the other reproductive toxicity effects from the animal study i.e. from the two generation reproduction toxicity study. In practice, the lowest of these DNELs would normally be taken to the risk characterisation.

Another example could be an occupational surveillance study of limited size and only addressing some of the relevant effects. In that case the animal study might be a more appropriate and reliable starting point in obtaining the DNEL/DMEL. However, also limited occupational surveillance data should be used in the WoE and integration as supporting evidence and/or source of qualitative data.

In order to use occupational survey data instead of good quality animal studies, one should exclude the possibility of chance (i.e. low statistical power), bias (especially selection bias and healthy worker effects) confounders or measurement errors in the study. Assessment of such quality factors is crucial, since unpublished occupational surveillance data have not undergone independent scientific review. Depending on the outcome of the assessment of quality and relevance, occupational surveillance data will either be robust enough to be used instead of animal data or be taken only as supporting study.

9.2.3 Negative human or animal data

Negative and inconclusive human studies are not evaluated in isolation but are taken to WoE analysis together with other relevant human and animal studies, when these are available. Thereby, the whole database and not only individual studies are evaluated.

In case a positive human study is of sufficient quality and relevant, while the animal study is negative, obviously the human study is taken into the risk characterization. It would be useful to explain this inconsistency, when possible, since that would increase the reliability and confidence in the risk assessment outcome. It may be, for example, that the animal study is negative because some observations that were made in the human epidemiological studies are not routinely made in animal studies (e.g. decrease of bone density and increase of fractures caused by cadmium in humans) It is also possible that there is a mechanism of toxicity in humans that is not relevant in animals.

If the animal data are positive and the human data are negative (see also Section 9.1) and both are of good quality the relevance of the animal data becomes crucial. There are two cases:

1. When the **human relevance of the animal data cannot be excluded**, the animal data will be the basis of the dose descriptor, provided the human data do not reasonably exclude the effect shown in the animal data. For example, human study design can be such that all relevant endpoints/health effects have been covered. In a case where, for example, only effects on haematology parameters were covered in the human study, but effects were seen in clinical chemistry or histopathology parameters in the animal study at a lower dose level, the negative human data is not conclusive, and therefore, the animal data has to be used when DNEL is obtained. Also the exposure levels in the human study and study power should be considered. If exposure levels in the human study were low, the effects may not have been observed in the human study and the animal data remains valid.

2. There are a number of **mode of action considerations** that are crucial when integrating negative human and positive animal data. These may lead to a conclusion **that the animal data is not relevant for humans**. The basic concept is that in case the mechanism of toxicity in the animal study is characterised and has been shown not to be relevant in humans, the negative human data have a stronger weight in the analysis made in the integration. An example are the renal neoplasms in male rats, developed coincidentally with α_{2u}-globulin nephropathy, due to accumulation of a specific protein, which are not considered predictive of risk to humans (Doi et al 2007). In cases similar to those described above, the negative human data are regarded as conclusive and no DNEL/DMEL needs to be obtained for this endpoint. It is to be underlined that it needs to be justified with data that the mode of action of the positive animal data does not apply to humans. Finally one needs to verify that the data concerning other non-threshold effects are negative.

9.2.4 Consistent data

In case data of good or sufficient quality both from human and from animal studies are consistent, i.e. essentially the same effects are seen, DNELs/DMELs from both sources should be considered in the WoE and integration. The lowest of the DNELs /DMELs should be used in the risk characterisation. However, in case the human data are adequate and complete they would take precedence over animal There may be cases where e.g. a biomarker information of an early effect (or e.g. hyperplasia) is obtained from a human study but only gross histo-pathological changes, which take place at higher level of exposure, are seen in the animal study. In that case the most sensitive study could be used for setting the DNEL/DMEL for that effect category based on expert judgment.

For ”data rich” substances there are often human data from several studies. When consistent qualitative and quantitative data come from several independent studies, low intraspecies assessment factors may be applied as described in Phase 7. Furthermore, in risk characterisation an effect seen in several studies should be preferably used instead of another effect, which is poorly characterised or only anecdotally described in literature.

9.2.5 Qualitative data in the integration

As instructed above, qualitative data, i.e. a study that has no dose descriptor (specified levels of exposure) but has valuable and reliable data on the relevant health effects should be considered in the phase of integration, while it cannot serve as a basis of setting the DNEL/DMEL.

Qualitative data should be regarded as a potentially supporting element in the WoE analysis/integration that aims at obtaining DNELs/DMELs. This could in practice imply that when

there are more than one DNEL/DMEL or dose descriptor either from human or animal studies, the supporting information concerning the modes of action, should be considered. This may be relevant e.g. in assessment of negative data as described above.

It is important that even after having derived DNELs either on the basis of an animal or human datasets, an assessment be made to verify if the proposed DNELs would be protective for all other endpoints for which a health effect was identified. Similarly in case of DMEL derivation, the proposed DMEL should be assessed to make sure that no other health effect presents a higher risk level.

Result of Phase 9: Having Table R.8-16 as a starting point the available human and animal data are integrated in order to select the critical DNEL (or DMEL) value for the exposure patterns of Table R.8-9. The decisions taken in the selection of the critical DNEL/DMEL have been documented in a transparent way. It has also been verified that the critical DNEL/DMEL values protect from all the other identified health effects of the substance. Table R.8-9 will be taken to the risk characterisation.

APPENDIX R.8-16 Examples of modification or deviation from the default intraspecies assessment factors

Below, three examples are given of the use of human data in chemical risk assessment. These cases are based on Risk Assessment Reports prepared by EU Member States within the implementation of the Existing Substances Regulation (ESR). In the guidance, which was applied in the ESR assessments, the instruction to use minimal Margins of Safety (MOSmin) was given. The basic elements of this guidance were the same as these described in Chapter R.8., i.e. the purpose of using MOSmin was the same and also their numerical values were the same as the standard Assessment Factors given in APPENDIX R.8-15.

The following examples are given to illustrate cases where deviation from MOSmins has been adequate and justified. In some of these cases the human studies have been large, heterogeneous and well characterised for human variability. Furthermore, there are examples of both sufficient and insufficient data base of human data for derivation of the DNEL. It is noteworthy that even if human data are insufficient for quantitative derivation of the DNEL, it may be relevant within the Weight of Evidence analysis where all available animal and human data are considered.

Thus, the examples show how in the past in certain cases the human data have been weighed against the animal data and how the default intraspecies factor was adjusted. Examples on how DNEL/DMEL would be derived in concrete cases will be developed and provided in the next phases of this guidance development process.

Cadmium

A relevant example of a situation where deviation from the default AFs is justified, is given in the Risk Assessment Report on cadmium. It is acknowledged that cadmium is a carcinogen; most of the evidence derived from studies where the exposure took place via inhalation and from occupational epidemiology. While deriving a DMEL for carcinogenicity of cadmium is a relevant topic, it is not the item of the example below. Instead the example given below addressed the effect of cadmium on kidney and on bone density (threshold effects) due to long-term exposures mainly via oral route, but also due combined exposure. This example is limited to those target organ effects which have a threshold in order to illustrate how intraspecies variation can be covered.

Among the industrial chemicals, cadmium has one of the largest toxicological data bases, of which human data is a significant part. The two most relevant human studies (Buchet et al. 1990 and Järup et al 2000) had a total sample size of 2720 individuals (workers and general population) in Belgium and in Sweden. Age of the subjects of the studies in the samples was 16-80. In addition, the following independent variables were considered in the analysis: sex, renal disease, diabetes, use of medication, body mass index and renal diseases. There are also several other relevant studies, which represent other European populations, e.g. from the Netherlands and from UK. In conclusion, deviation from the standard AFs in deriving a DNEL for cadmium risk assessment would be justified, because most, if not all, of the factors causing human variability are covered. In the risk characterization, the critical level of 2 µg urinary Cd/g creatinine is used, and this value is taken directly from the most recent, representative and good quality epidemiological studies, in which the renal effects and effects on bone density were seen approximately at this level of excreted cadmium. Depending on the type of effect (kidney or bone) the mode of calculation and relevant study, several critical urinary levels are given in the risk assessment of cadmium ranging from 0.5 to 5.0 µg urinary Cd/g creatinine.

The text for the "compromise" LOAEL is illustrative, as it says: "Trying to aggregate all these data, a LOAEL of 2 µg urinary Cd/g creatinine is proposed. This figure should be understood as a composite level, based on the association between Cd and not only low molecular weight proteins in urine but also calcium excretion in urine and its possible relationship with bone effects" Margin of Safety of 3 is used in the RC to account the conversion from LOAEL to NOAEL, **but no intraspecies assessment factor** (or MOS accounting for human variability) was used for cadmium.

This example furthermore illustrates that since biomonitoring values are used instead of exposure data, the toxicokinetic factor of intraspecies variation has largely been covered and does not need to be additionally accounted for. The example also suggests that with large heterogeneous populations and the known relevant parameters covered, which cause the intraspecies variation, there appeared to be no reason to use an intraspecies assessment factor.

Hydrogen peroxide

Another (opposite) example is provided in the risk assessment report of hydrogen peroxide (H_2O_2), a strong oxidising agent, which acts as an irritant or corrosive agent, depending on its concentration. A health monitoring study (occupational surveillance) of six aseptic packaging workers was conducted. It involved a 10-month period of high exposure (2-3 mg/m^3 8-hour (time weighted average), peaks up to 11 mg/m^3) due to machine malfunction and, after repairs, a one-year follow up at a reasonably low and stable exposure (0.5-0.7 mg/m^3 8-hour TWA). The results indicated that three of the workers experienced eye and airway irritation, headache, and a uniform course of recurring bronchitis-sinusitis which coincided with the high exposure (Riihimäki et al., 2002). The study did not include specific examinations of the lungs. It was concluded that further data, including human observations, are helpful to characterise and confirm the repeated dose toxicity of hydrogen peroxide by inhalation.

Furthermore, industrial experience from health surveillance of H_2O_2 production workers suggested no exposure-related effects on simple respiratory functions at airborne levels of up to 0.8 mg/m^3 (CEFIC, 1996b) or less than 1.4 mg/m^3 with short-term peaks of up to about 5 mg/m^3 (Degussa-Hüls, 1999). Since these observations were not derived from properly conducted studies, the health data cannot be used as solid evidence for the absence of adverse pulmonary effects.

Because of the uncertainties and/or preliminary nature of the human data, they were not taken into account in the risk characterisation in that risk assessment report. Instead, more robust animal data were used to characterise the repeated dose inhalation toxicity of hydrogen peroxide. Interestingly, the effect concentrations in animal compared to human studies are rather consistent. Whether human data was dealt with in the Weight of Evidence analysis is not explained in the risk assessment report. This example illustrates that a study with small sample size, where all relevant parameters/observations are not covered is not a valid basis for obtaining a NOAEL or dose descriptor.

Toluene

The third example is about human data on reproductive toxicity of toluene. Two studies suggest an increased risk of spontaneous abortions associated with exposure to toluene in the workplace.

Spontaneous abortions among women working in laboratories, (together with congenital malformations and low birth weights of the children) were examined in a retrospective case-control

study (Taskinen et al., 1994). The exposure to toluene was assessed on the basis of the reported frequency of the use of the chemical and classified as frequent if the chemical was handled at least 3 days a week and rare if the toluene was handled 1 or 2 days a week. Significant associations with spontaneous abortions were found for frequent exposure to toluene (odds ratio 4.7, confidence interval 1.4 to 15.9) after adjustment for various covariates (206 cases and 329 referents). This study suggests an association between exposure to toluene during early pregnancy and increased risk of spontaneous abortion. The result should be interpreted cautiously because the women were often exposed to several solvents and other chemicals simultaneously. Furthermore, no information on exposure levels is presented. In conclusion the rapporteur (under ESR) considered that the results are of limited use for the risk assessment of toluene.

In another study, rates of late spontaneous abortions were determined using a questionnaire addressing reproductive effects in 55 women with 105 pregnancies exposed to toluene (mean 88 ppm, range 50-150 ppm), 31 women (68 pregnancies) working in the same factory in departments where little or no exposure to toluene occurred (0-25 ppm), and an external community control group of 190 working class women with 444 pregnancies (Ng et al., 1992b). Significantly higher rates for late spontaneous abortions defined as ‘pregnancy loss’ between weeks 12 to 28 were noted in the toluene-exposed women compared with those in the internal and external control groups (12.9% vs. 2.9-4.5%). The differences in the rates of late spontaneous abortions between groups were not likely to be confounded by classical risk factors such as maternal age, gravidity, smoking, or alcohol, which were taken into account both in the study design and the analysis. Information on pregnancy outcomes might be biased by questionnaire interview. The pregnancies and abortions in the factory were not validated by access to medical records or with biological methods. However, relatively unequivocal endpoints were used in the questionnaire, thus excluding doubtful pregnancies and abortions.

In the course of this risk assessment the authors of the study clarified that, since the study was a cross-sectional observational study that relied on a questionnaire, with information obtained by the subject's recall of her recent pregnancy(ies), it was difficult to determine with absolute certainty whether a spontaneous abortion had indeed occurred, especially in the first two months after conception. It is known that foetal loss is a lot more common than is generally supposed especially in the first month immediately after conception. When it occurs, it is often disregarded as a 'missed period', when menstruation resumes a month or two later.

It was concluded in the risk assessment report that the second study cannot be used to establish definitively a causal relationship between late spontaneous abortions and toluene exposure. (To establish a definite relationship, a prospective study of pregnant women exposed to toluene at similar exposure levels (mean 88 ppm, range 50-150 ppm) with individually monitored data on toluene exposure and foetal loss would be needed.) However, based on the current evidence suggesting an increased risk for late spontaneous abortions, exposing pregnant women to such exposure levels would raise serious ethical concerns. Consequently, the results of the second study are used as a basis for the risk characterisation of developmental toxicity in humans.

It is noteworthy that animal inhalation studies provided strong evidence of developmental toxicity (lower birth weight and long-lasting developmental neurotoxicity) in the absence of maternal toxicity. Furthermore, there were also in total about 45 human cases reported in the literature of so-called toluene embryopathy as result of sniffing toluene. (These cases resembled foetal alcohol syndrome, and there might be a common mechanism.)

The human LOAEC of 88 ppm (330 mg/m^3) and the rat NOAEC of 600 ppm (2,250 mg/m^3) were taken forward to the risk characterisation. Different MOSs were applied to these data. Risk characterisation for workers was made for “toxicity to reproduction (fertility and development)”

where the animal data was used with a MOSmin of **30**. Another risk characterisation was on "toxicity to reproduction (spontaneous abortions)", where a lower MOSmin of **5** was applied because "the NOAEC for this endpoint is derived from human data". (To correct an error in the report, in fact LOAEC and not the NOAEC was used when the MOSs were calculated.)

The human data was not used for risk characterisation of consumer exposure.

The case of toluene shows that a human study made with only a few hundred individuals may be very important element in the Weight of Evidence evaluation. In this case, the better of the two human studies with relevant exposure data led to use of a specific "dose-descriptor" in the risk characterisation. In fact, the animal NOAEC and human LOAEC are used in parallel, with the respective MOSs and they have led to exactly the same risk assessment result in terms of formal conclusions under ESR, thereby showing the relevance of the human data and increasing the overall robustness of the assessment.

Furthermore, the example shows how important it is to consider the confounding factors. Also the weaknesses of the human study were adequately reported, i.e. the lack of a definite causal relationship and the potential reporting bias.

Assuming that the MOSmin of 5 (for workers, derived from human data) actually covers (i) some of the intraspecies variation and (ii) the step from LOAEC to NOAEL, the case of toluene also demonstrates how human data can be used to modify the intraspecies assessment factor.

References

(Please note that the references in the Appendix R.8-16 have not been listed here; they can be found in the respective Risk Assessment Reports included in the ECB site ***http://ecb.jrc.ec.europa.eu/esis/index.php?PGM=ora****)*

Burin G J, and Saunders D R (1999) Addressing Human Variability in Risk Assessment – The Robustness of the Intraspecies Uncertainty Factor. *Regul Toxicol Pharmacol* **30**: 209-216

Checkoway H, Pearce N, Kriebel D. Research Methods in Occupational Epidemiology. Oxford: Oxford University Press, 2004. ISBN: 0-19-509242-2.

Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. Brit Med Journal online. http://www.bmj.com/epidem/epid.html

Danish EPA (2004). Principper for sundhedsmæssig vurdering af kemiske stoffer med henblik på fastsættelse af kvalitetskriter for luft, jord og drikkevand. [Principles for health based evaluations of chemicals for the setting of quality criteria for ambient air, soil and drinking water] in Danish with English summary. Miljøprojekt nr. 974, pp 167.

Doi A, Hill G, Seely J, Hailey J R, Kissling G, Buchler J R (2007) α2u-Globulin Nephropathy and Renal Tumors in National Toxicology Program Studies. Toxicol Pathol **35**, 533-540.

Dourson N L, Andersen M E, Erdreich L S, MacGregor J A (20019 Using Human Data to Protect the Public's Health *Regul Toxicol Pharmacol* **33**: 234-256.

Dutch Expert Committee on Occupational Standards (2003), Formaldehyde, helath based recommended occupational exposure limits. No. 2003/02OSH, The Hague, 27 January 2003.

Dutch Health Council, developed jointly by the members of the committee on the evaluation of carcinogenic substances (1989) Carcinogenic risk assessment of benzene in outdoor air. Regul. Toxicol Pharmacol 9: 175-185

Ecetoc (2009). Framework for the Integration of Human and Animal Data in Chemical Risk Assessment. Technical Report No. 104. Ecetoc, Brussels, January 2009.

Ecetoc (2006). Toxicological Modes of Action: Relevance for Human Risk Assessment. Technical Report No. 99. Brussels, July 2006.

Ecetoc (2004) Workshop on the Use of Human Data in risk Assessment, 23-24 February 2004, Gardiff, Workshop Report No. 3. Ecetoc, Brussels, November 2004.

Ecetoc (2003) Derivation of Assessment Factors for Human Health Risk Assessment, Technical report No. 86. Ecetoc, Brussels, February 2003.

EFSA SC (2005). Opinion of the Scientific Committee on a requat from EFSA related to A Harmonised Approach for Risk Assessment of Substances Which are both Genotoxic and Carcinogenic. The EFSA Journal (2005) 282 :1-31.

Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 ;315(7109):629–624.

Von Elm E, Altman DG, Egger M et al (2007). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453-57.

European Commission, Joint Research Centre, European Chemicals Bureau (2007) European Union Risk Assessment Report, Cadmium oxide and cadmium metal, 3rd Priority List, Volume 72.

European Commission, Joint Research Centre, European Chemicals Bureau (2003) European Union Risk Assessment Report, Toluene, 2rd Priority List, Volume 30.

European Commission, Joint Research Centre, European Chemicals Bureau (2003) European Union Risk Assessment Report, Hydrogen Peroxide, 2rd Priority List, Volume 38.

Goldbohm RA, Tielemans EL, Heederik D, Rubingh CM, Dekkers S, Willems MI, Kroese ED (2006) Risk estimation for carcinogens based on epidemiological data: A structured approach, illustrated by an example on chromium. *Regul Toxicol Pharmacol* **44**: 294-310

Hattis D, Erdreich L, Ballew M (1987) Human variability in Susceptibility to Toxic Chemicals A Preliminary Analysis of Pharmacokinetic Data from Normal Volunteers. Risk Analysis **7**, 415-426

Hernberg S (1992). Introduction to Occupational Epidemiology. Lewis Publishers, Inc. Chelsea, MICH 1992.

Hill AB (1965). The environment and disease: Association or causation? Proc of the Royal Soc Med 58:295-300.

IARC (2006) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Preamble. Word Health Organisation, International Agency for Research on Cancer, Lyon, January 2006.

IARC (2009). A review of human carcinogens—Part F: Chemical agents and related occupations. The Lancet Oncology 10,1143-44.

IPCS. 2005a. Workshop on poison centres and the use of human data in consumer product risk assessment. Report. WHO in collaboration with EC JRC, EAPCCT, ECETOC and the Federal Institute of Risk Assessment, Germany held at the Federal Institute for Risk Assessment, Berlin, Germany, 9 May 2005. International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland.

IPCS. 2005b. Harmonization Project Document No. 3. Principles of characterizing and applying human exposure models. International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland.

IPCS. 2005c. Harmonization Project Document No. 2. Chemical-specific adjustment factors for interspecies differences and human variability: Guidance document for use of data in dose/concentration–response assessment. International Programme on Chemical Safety, World Health Organization, Geneva, Switzerland.

IPCS 2007a. IPCS framework for analyzing the relevance of a cancer mode of action for humans. In Harmonization Project Document No. 4 -Part 1: IPCS framework for analysing the Relevance of a cancer mode of action for Humans and Case-studies. IPCS, WHO, Geneva, Switzerland 2007a. (authors Boobis AR, Cohen SM, Dellarco V, Mcgregor D, Meek ME, Vickers C, Wilcocks D, Farland W, also published in Crit. Rev. Toxicol. 2006;36:781-92).

IPCS 2007b. IPCS framework for analysing the relevance of a non-cancer mode of action for humans. In Harmonization Project Document No. 4 -Part 1: IPCS framework for analysing the Relevance of a cancer mode of action for Humans and Case-studies. IPCS, WHO, Geneva, Switzerland 2007b. (Authors: Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, Vickers C, also published in Crit. Rev. Toxicol. 2008;38:87-96).

Kilimisch H, Andreae M, Tillman U (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological information. Regul Toxicl Pharm 1997;25:1-5.

Lehman A.J and O.G. FITZHUGH. 1954. 100-fold margin of safety. Assoc. Food Drug Off. U.S. Q. Bull. 18: 33–35.

Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DE. A framework for human relevance analysis of information on carcinogenic modes of action. Crit Rev Toxicol. 2003;33:591-653.

Money CD, Margary SA (2002) Improved Use of Workplace Exposure Data in the regulatory Risk Assessment of Chemicals within Europe. Ann Occup Hyg 46: 279-285

Money CD (2007). The use of Human Experience Data in the EU Risk Assessment Process. Risk Analysis 27, 387-396.

OECD (2007). Manual for investigation of HPV chemicals. Organisation for Economic Co-operation and Development, Paris, France [www.oecd.org/document/7/0,3343,en_2649_34379_1947463_1_1_1_1,00.html].

Renwick A G, Dorne J L, Walton K (2000) An Analysis of the Need for an Additional Uncertainly Factor for Infants and Children. *Regul Toxicol Pharmacol* **31**: 286-296.

Rothman KJ, Greenland S. (1998). Modern Epidemiology. Lipincott-Raven Publishers Philadelphia USA. ISBN:0-316-75780-2

Steenland K, Greenland S (2004) Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. *Am J Epidemiol* **160**: 384-392

Sackett DL (1979). Bias in analytic research. *J Chron Dis* **32**:51-63

Sanco (2006) European Commission, health & consumer protection directorate-general. (Working document) Draft guidance for the setting and application of acceptable operator exposure levels (AOELs). SANCO 7531 - rev.10, 7 July 2006

Steenland K, Spaeth S, Cassinelli R, Laber P, Chang L, Koch K (1998) NIOSH life table program for personal computers. Am J Ind Med 34: 517-518

Swaen GMH. A framework for using epidemiological data for risk assessment. Hum Exp Toxicol 2006;25:147-155.

Swaen G, van Amelsvoort L (2009). A weight of evidence approach to causal inference. J of Clin Epidemiol 2009;62:270-7.

US EPA (2005). Guidelines for Carcinogen Risk Assessment. Risk Assessment Forum U.S. Environmental Protection Agency Washington, DC. EPA/630/P-03/001F, March 2005

Vlaanderen J, Vermeulen R, Heedick D, Kromhout H (2008) Guidelines to Evaluate Human Observational Studies for Quantitative Risk Assessment. Environm Health Perspect.,**116**: 1700-1705.

van Wijngaarden E, Hertz-Picciotto I (2004) A simple approach to performing quantitative cancer risk assessment using published results from occupational epidemiology studies. *Sci Total Environ* **332**: 81-87

Vineis P, McMichael AJ (1998). Bias and confounding in molecular epidemiological studies: special considerations. *Carcinogenesis* **19**:2063-2067.

WHO Working Group report (2000) Evaluation and use of epidemiological evidence for environmental health risk assessment: WHO Guideline Document. Environm Health Perspect.,108: 997-1002.

World Medical Association (2000). Declaration of Helsinki.: Ethical Principles for Medical Research Involving Human Subjects. Adopted by the 18th World Medical Association (WMA) General Assembly Helsinki, Finland, June 1964 and amended by the 29th WMA General Assembly, Tokyo, Japan October 1975, 35th WMA General Assembly, Venice, Italy, October 1983, 41st WMA General Assembly, Hong Kong, September 1989, 48th WMA General Assembly, Somerset West, Republic of South Africa, October 1996 and 52nd WMA General Assembly, Edinburgh, Scotland, October 2000.

ABBREVIATIONS

Term	Definition
ABS	Absorption
AF	Assessment Factor
AS	Allometric Scaling
AUC	Area Under the Curve; Area under the blood/plasma concentration curve vs. time curve, representing the total amount of substance reaching the blood/plasma
BMD	Benchmark dose; The BMD concept involves fitting a mathematical model to dose-response data. The BMD is defined as the dose causing a predetermined change in response.
BMD10	The Benchmark-dose associated with a 10% response (for tumours upon lifetime exposure after correction for spontaneous incidence, for other effects in a specified study)
BMDL10	The lower 95% confidence interval of a Benchmark-dose representing a 10% response (e.g., tumour response upon lifetime exposure), i.e. the lower 95% confidence interval of a BMD10
Bw	Body weight
Cmax	Peak plasma Concentration
CNS	Central Nervous System
DMEL	Derived Minimal-Effect Level; For non-threshold effects, the underlying assumption is that a no-effect-level cannot be established and a DMEL therefore expresses an exposure level corresponding to a low, possibly theoretical, risk, which should be seen as a tolerable risk.
Dose descriptor	A value obtained from a toxicity/ecotoxicity test or from other relevant data, usually the dose needed to induce a specified adverse effect (e.g., 50% lethality) or the highest dose not causing adverse effects (e.g., NOAEL). The dose descriptor is a basis for determining/setting the DNEL.
EC3	Effect concentration 3; The amount of chemical required to induce a Stimulation Index (SI) of 3.
ED10	Effective dose 10 %; a dose representing an increased incidence of 10 % due to a specific exposure (e.g., to a chemical).
EFSA	European Food Safety Authority
ELR	Excess Lifetime Risk; additional lifetime risk over the background normal risk (or incidence of disease)
HBMD10	Human BMD10

HT25	Human T25
HtLF	High to Low dose risk extrapolation Factor
LC50	Median lethal concentration. The concentration causing 50 % lethality
LD50	Median lethal dose. The dose causing 50 % lethality
LED10	Lowest confidence limit of the ED10
LMS	Linear multistage model
MTD	Maximum Tolerated Dose
NAEC	No Adverse Effect Concentration
NAEL	No Adverse Effect Level
OR	Odds Ratio; the ratio of the odds of an event occurring in one group to the odds of it occurring in another group
ORL	Lowest confidence limit of the OR
PBPK	Physiologically-based pharmacokinetic modelling
RC	Risk Characterisation
RCR	Risk Characterisation Ratio
RR	Relative Risk
RRL	Lower bound exposure value associated with the RR-value of 1.1
SI	Stimulation Index
SMR	Standardised Mortality Ratio
SMRL	Lower bound exposure value associated with the SMR-value of 1.1
sRV	Standard Respiratory Volume
T25	The chronic dose rate that will give 25% of the animals' tumours at a specific tissue site after correction for spontaneous incidence, within the standard life time of that species
TTC	Threshold of Toxicological Concern
TWA	Time-Weighted Average exposure
wRV	Worker Respiratory Volume