Chapter 2
Assessing Risks to Human Health from Chemicals in the Environment

SUSAN P. FELTER, MICHAEL L. DOURSON AND JACQUELINE PATTERSON

2.1 INTRODUCTION

Humans are exposed to a multitude of potentially hazardous chemicals in their indoor and outdoor air, food, soil and ambient and drinking waters. Risk assessment is the process whereby scientists evaluate the toxicity data for chemicals to which humans are, or may be, exposed, and attempt to identify and quantify potential risks to health. The risk assessment process is also used to estimate levels of intake via the various media that are expected to be 'safe'. These values are then used in conjunction with information on exposure in order to determine acceptable levels for concentrations of hazardous chemicals in environmental media.

The process of human health risk assessment was first described as a four-component paradigm by the National Research Council (NRC) of the National Academy of Sciences (NAS) in 1983 and was subsequently updated in 1994. This chapter follows the NAS paradigm and introduces each component with an excerpt from the NAS (1994) publication, *Science and Judgment in Risk Assessment*. The focus is primarily on the risk-assessment methods used by national or health agencies, such as the International Programme on Chemical Safety (IPCS) or the US Environmental Protection Agency (USEPA). However, scientists from other groups have made contributions to the field, especially in the area of research to improve the standard methods.

2.2 HAZARD IDENTIFICATION

> *Hazard Identification* entails identification of the contaminants that are suspected to pose health hazards, quantification of the concentrations at which they are present in the environment, a description of the specific forms of toxicity (neurotoxicity, carcinogenicity, etc.) that can be caused by the contaminants of concern, and an evaluation of the conditions under which these forms of toxicity might be expressed in exposed humans ...
>
> NAS, 1994

2.2.1 Hazard identification of non-cancer end-points

Hazard identification is generally the first step of the risk assessment process, in which it is determined if there is a potential cause for concern over human exposure to an agent. This involves an evaluation of the appropriateness, nature, quality and relevance of scientific data on the specific chemical; the characteristics and relevance of the experimental routes of exposure; and the nature and significance to human health of the effects observed. The USEPA, for example, has developed hazard identification guidelines for developmental and reproductive toxicity that carefully address these issues (USEPA 1991, 1994a). Table 2.1 gives a brief list of considerations.

Much of the process of hazard identification for non-cancer end-points depends on professional judgement as to whether or not an observed effect, or collection of effects (or syndrome), constitutes an adverse response. This is not always easy, and often requires the views of experts in the subject area, because although many effects are clearly adverse (e.g. fatty infiltration of the liver), many

Table 2.1 Considerations in characterizing hazard. (Information from USEPA 1995b.)

What are the key toxicological studies and of what quality? Are the data from laboratory or field studies? Single or multiple species? For cancer: Was there a single or multiple tumour site(s)? Benign or malignant? Was the maximum tolerated dose achieved? For other-than-cancer: What end-points were observed and what is the basis for the critical effect? Other supporting studies? Conflicting?
Besides for the critical effect, are there other end-points of concern?
What are the significant data gaps?
What are the available epidemiological or clinical data? • What types of studies were used (i.e. ecological, case–control, cohort)? • Describe the degree to which exposures were described adequately, to which confounding factors were accounted for adequately, and to which other causal factors were excluded
Were there non-positive animal or human data?
How much is known about the biological mechanism of action and how does this aid in the interpretation of the data?
Summarize the hazard identification and discuss the confidence in the conclusions, alternative conclusions that are also supported by the data, significant data gaps and highlights of any major assumptions

others are of uncertain toxicological consequence (e.g. decrease in body weight gain).

Because toxic chemicals often elicit more than one adverse effect, the process of hazard identification for non-cancer toxicity includes an evaluation of the target organ or 'critical' effect; i.e. the first adverse effect or its known precursor that occurs as the dose rate increases. This is shown hypothetically in Fig. 2.1, where several effects are evoked from chemical exposure:

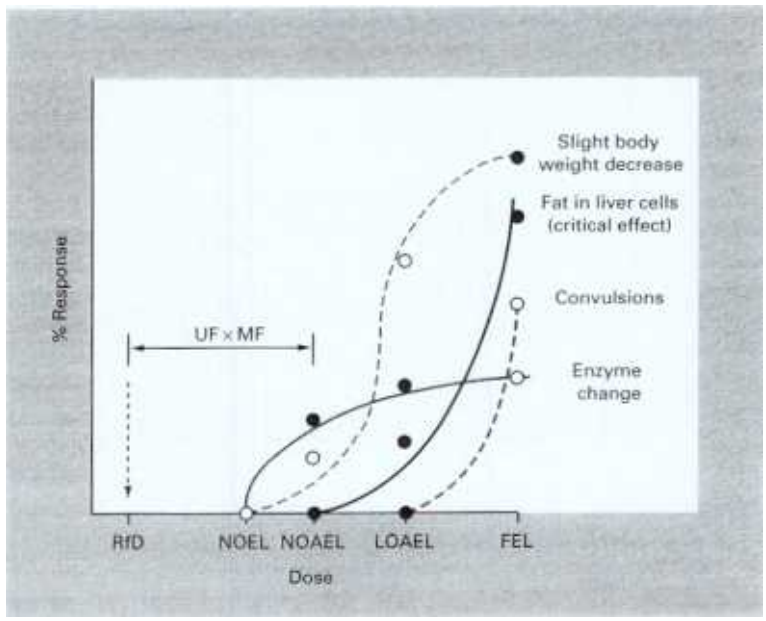

Fig. 2.1 The judgement of the critical effect and its NOAEL (no-observed-adverse-effect level), along with the appropriate uncertainty factor (UF) and modifying factor (MF), leads to the estimation of the RfD (reference dose). NOEL, no-observed-effect level; LOAEL, lowest-observed-adverse-effect level; FEL, frank effect level.

enzyme change, slight decrease in body weight, fatty infiltration of the liver and convulsions. Enzyme change and slight decrease in body weight are judged not to be adverse effects. Fatty infiltration of the liver is judged to be the critical effect.

The judgement of whether an effect is adverse or critical may change among toxicity studies of different durations, and may be influenced by toxicity in other organs or by toxicokinetics. A good example of this is increased liver weight due to a proliferation of smooth endoplasmic reticulum with chemical exposure. Such an effect may be judged as not adverse if the parent chemical is the toxic moiety and such an increase is likely to quicken its metabolism, or may be judged to be adverse if a metabolite is the toxic moiety (Farland & Dourson 1992). The distinction of adverse compared with non-adverse effects and the choice of critical effects in the hazard identification component of the paradigm is then used as a basis for the dose–response assessment.

2.2.2 Hazard identification of carcinogens

Hazard identification of carcinogens refers to the process of determining if a compound has the potential to elicit a carcinogenic response in humans. Many types of information may be used to determine the overall weight-of-evidence of carcinogenicity: epidemiological information, chronic animal bioassays, mechanistic data, mutagenicity tests, other short-term tests, structure–activity relationships, metabolic and pharmacokinetic properties, toxicological effects and physical and chemical properties.

The first organization to develop a classification scheme for carcinogenicity was the International Agency for Research on Cancer (IARC) in 1978. Based on a strength-of-the-evidence approach (evidence coming from human or laboratory animal data or short-term studies), chemicals were placed in one of three categories.

Group 1: carcinogenic to humans.

*Group 2**: probably carcinogenic to humans.

Group 3: cannot be classified as to its carcinogenicity to humans.

In 1986, the USEPA published general guidelines to be used by Agency scientists in developing and evaluating risk assessments for carcinogens (USEPA 1986). Based on the weight-of-evidence from epidemiological and laboratory animal bioassays, chemicals are placed in one of six categories. Supporting data (e.g. mutagenicity data, mechanistic data) may then be used to move a chemical up or down in the ranking. These categories are modelled after those used by IARC.

Group A: carcinogenic to humans.

Group B†: probably carcinogenic to humans.

Group C: possibly carcinogenic to humans.

Group D: not classifiable as to human carcinogenicity.

Group E: evidence of non-carcinogenicity for humans.

In April 1996, the USEPA proposed revisions to the carcinogen risk assessment guidelines. In contrast to the concise alpha-numeric classification system of 1986, the guidelines proposed advocate the development of a more comprehensive characterization of the carcinogenic hazard in the form of a narrative. Within this context, a cancer hazard characterization should include all information relevant to the weight-of-evidence for carcinogenicity, not just tumour data in humans and animals. This means that mechanistic data can play an integral role in the hazard identification step for carcinogenicity, and may also influence the choice of a dose–response model. Another change is that the hazard characterization can provide specific information about the conditions under which a chemical is likely to be carcinogenic; for example, it may be likely to be carcinogenic by the route of inhalation but not by ingestion. These proposed changes in the USEPA methods reflect a general movement in the field of cancer risk assessment to include more chemical-specific data and to move away from the use of default positions wherever possible.

* Group 2 includes subgroups 2A (for chemicals having limited evidence of carcinogenicity in humans) and 2B (for chemicals having sufficient evidence of carcinogenicity in laboratory animals, and inadequate evidence in humans).

† Group B includes subgroups B1 (for chemicals having limited evidence of carcinogenicity in humans) and B2 (for chemicals having insufficient human data but sufficient animal data).

In addition to what has been described here for the USEPA, other groups have published carcinogen classification schemes along similar lines. Moolenaar (1994) has provided a summary and comparison of several international classification schemes, including eight governmental agencies and two independent organizations. These classification schemes have anywhere from two to six distinct categories with varying degrees of emphasis on mechanistic data. In addition, Ashby *et al.* (1990) have recommended an eight-category system.

Common to many of these groups, the determination of carcinogenic hazard includes a determination of whether the incidence of tumour types observed to occur in laboratory animals is statistically significantly elevated over that observed in controls. Two forms of statistical tests are used to answer this question: trend tests, which look for an overall trend of increasing tumour incidence with increasing dose; and pairwise comparison tests, which directly compare the tumour incidence in an individual dose group with that seen in controls.

Determination of the mechanism by which a chemical causes cancer in laboratory animals also provides information about the potential for human carcinogenicity relevant to the hazard identification process. The potential for the same or a related mechanism to be operative in humans provides the basis for extrapolation from other animal species to estimate the risk of cancer to humans. For some specific tumour types, or mechanisms of carcinogenicity, there are indications that tumours observed in laboratory animals may have no relevance or limited relevance to human carcinogenicity. Tumour types that are included in this group include kidney tumours in male rats that are caused by the accumulation of a male-rat-specific protein (alpha$_{2u}$-globulin); liver tumours in male B6C3F1 mice; thyroid follicular cell tumours; and bladder tumours related to the formation of silicate-containing precipitate and crystals (e.g. as seen in saccharin-induced bladder cancer in rats).

2.3 DOSE–RESPONSE ASSESSMENT

> *Dose–Response Assessment* entails a further evaluation of the conditions under which the toxic properties of a chemical might be manifested in exposed people, with particular emphasis on the quantitative relation between the dose and the toxic response. The development of this relationship may involve the use of mathematical models. This step may include an assessment of variations in response, for example, differences in susceptibility between young and old people.
>
> NAS, 1994

2.3.1 Non-cancer end-points

The 'Safe' dose approach

Dose–response assessment follows hazard identification in the risk assessment process. Dose–response assessment involves the quantitative evaluation of toxicity data to determine the likely incidence of the associated effects in humans. The information available for dose–response assessment ranges from well-conducted and well-controlled studies on human exposures, and epidemiology studies with large numbers of subjects, well-characterized exposures, and supportive studies in several animal species, to a lack of human and animal toxicity data with only structure–activity relationships to guide the evaluation. In any case, scientists should consider all pertinent studies in this process; even a single human case study can provide useful information. However, only data of sufficient quality, as judged by experts, should be used in the dose–response assessment of a chemical. Table 2.2 lists some questions to be considered in characterizing the dose–response relationship for an agent.

Most non-cancer effects resulting from exposure to toxic agents are thought to be associated with a threshold; i.e. an exposure exists below which toxicity does not occur. The dose–response component of risk assessment involves the quantitative evaluation of toxicity data to determine a level of exposure for humans that is considered by risk assessors to be below the threshold for toxicity for sensitive subgroups.

Health agencies throughout the world support the use of a 'safe' dose concept, and define terms

Table 2.2 Considerations in characterizing the dose-response assessment. (Information from USEPA 1995b.)

Data
Which data were used to develop the dose-response curve? Would the results have been different if based on a different data set?
If animal data were used, which species were used—most sensitive, average of all species, or other? Were any studies excluded and why?
If epidemiological data were used, were they only the positive, all studies, or a combination? Were any studies excluded and why?
Was a meta-analysis performed to combine the studies? If so, what approach was used?
Models
What model was used to develop the dose-response curve? The rationale for this choice? Is chemical-specific information available to support this approach?
For non-cancer end-points how was 'safe' dose calculated? What assumptions and/or uncertainty factors were used? For benchmark doses, what model was used and why?
For cancer end-points, what dose-response model was used and why was it selected? Would other models have provided as plausible results?
Discuss the route and level of exposure observed in the data compared with anticipated human exposure. If data are from a different route, are pharmacokinetic data available to extrapolate across routes? How far is the extrapolation from the observed data to environmental exposures and what is the impact of this extrapolation?
Toxicity values
Summarize the risk value and discuss the confidence in the value. Can a range of values be provided? What are the results of different approaches or models?

and conditions for use. This 'safe' or subthreshold dose often goes by different names, such as: Health Canada's Tolerable Daily Intake or Concentration (TDI or TDC) (Meek *et al.* 1994); IPCS's Tolerable Intake (TI) (IPCS 1994); US Agency for Toxic Substances and Disease Registry's (ATSDR's) Minimum Risk Level (MRL) (Pohl & Abadin 1995); USEPA's Reference Dose (RfD) (Barnes & Dourson 1988; Dourson 1994) or Reference Concentration (RfC) (Jarabek 1994; USEPA 1994b); or the World Health Organization's Acceptable Daily Intake (ADI) (Lu 1985, 1988). Many of the underlying assumptions, judgements of critical effect, and choices of uncertainty factors (or safety factors) are similar among health agencies in estimating these subthreshold doses.

One of the best-known methods is that used by the USEPA to derive reference doses (RfDs) and reference concentrations (RfCs), which are subthreshold exposures for non-cancer toxicity. They are defined as: '... an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime' (Barnes & Dourson 1988).

The subthreshold dose approach starts with an identification of the critical effect(s), as described in Hazard Identification (2.2.2). The critical dose is then chosen. All groups rely on the experimental dose that represents the highest level tested at which the critical effects were not demonstrated as this critical dose. This dose is often called the no-observed-adverse-effect level (NOAEL), or the no-observed-effect level (NOEL). If a NOAEL is not available, the use of a lowest-observed-adverse-effect level (LOAEL) is often used as the critical dose.

Human data are preferred in the determination of an RfD or RfC. However, in the absence of these data, animal data are closely scrutinized. Risk assessment scientists seek to identify the animal

model that is most relevant to humans, based on the most defensible biological rationale. In the absence of a clearly most relevant species, the critical study and species that shows an adverse effect at the lowest administered dose are generally selected. This is based on the assumption that, in the absence of data to the contrary, humans may be as sensitive as the most sensitive experimental animal species. Uncertainty factors (UFs) are then used as divisors to this critical dose (NOAEL or LOAEL) to determine the subthreshold dose. These factors are considered as reductions in the dose rate to account for several areas of scientific uncertainty inherent in most toxicity databases. As shown in Table 2.3, these areas include interhuman variability (designated as H); extrapolation from experimental animals to humans (designated as A); extrapolation from subchronic to chronic exposure (designated as S); extrapolation from an experimental LOAEL to NOAEL (designated as L); and how to account for the lack of a complete database. In addition to these UFs, several groups also use a modifying factor that can be used to account for uncertainties not explicitly dealt with by the standard factors.

All groups occasionally use a factor less than 10 or even a factor of 1, if the existing data reduce or obviate the need to account for a particular area of uncertainty.* For example, the use of a 1-year rat study as the basis of an RfD may reduce the need for a tenfold factor for the area of subchronic-to-chronic extrapolation to threefold, because it can be demonstrated empirically that 1-year NOAELs for rat are generally closer in magnitude to chronic values than are 3-month NOAELs. Lewis *et al.* (1990) investigate this concept of variable uncertainty factors more fully through an analysis of expected values.

The choice of appropriate uncertainty and modifying factors reflects a case-by-case judgement by experts and should account for each of the applicable areas of uncertainty (described in Table 2.3) and any nuances in the available data that might change the magnitude of any factor. Several reports describe the underlying basis of uncertainty factors (Zielhuis & van der Kreek 1979; Dourson & Stara 1983) and research into this area (Calabrese 1985; Hattis *et al.* 1987; Hartley & Ohanian 1988; Lewis *et al.* 1990; Renwick 1991, 1993; Calabrese *et al.* 1992; Dourson *et al.* 1992; Calabrese & Gilbert 1993; Kroes *et al.* 1993; Abdel-Rahman & Dourson 1995).

The scientific strengths and limitations of this approach have been discussed in the literature (Munro & Krewski 1981; Lu 1983, 1985, 1988; Krewski *et al.* 1984; Crump 1984, 1986; Dourson *et al.* 1985, 1986; Barnes & Dourson 1988; Kimmel & Gaylor 1988). The scientific strengths, in brief, are that all toxicity data are reviewed in the choice of the NOAEL for the critical effects, and that uncertainties in the entire data base can be factored into the resulting value of the subthreshold dose through the use of professional judgement as to the appropriate uncertainty and modifying factors.

The limitations, in brief, are that the NOAEL is restricted by the choice of dose-spacing and the number of animals, as well as factors that influence the quality of the study. Studies with wide dose-spacing and a low number of animals per dose group can lead to a more poorly characterized subthreshold dose as compared to studies with tighter dose-spacing and more animals per dose group (see, e.g., Hattis *et al.* 1987; Leisenring & Ryan 1992). The NOAEL is also not generally influenced by the nature of the dose–response curve. Uncertainty factors, although considered necessary and perhaps to reflect accurately the potential underlying areas of uncertainty, are quite imprecise. Nor does the subthreshold-dose approach enable an estimate of risks at exposures greater than the subthreshold dose.

Scientists are developing methods that address some of these latter limitations (e.g. DeRosa *et al.* 1985; Dourson *et al.* 1985; Kimmel & Gaylor 1988; Kimmel *et al.* 1988; Hertzberg 1989; Hertzberg & Dourson 1993; Renwick & Walker 1993; Faustman *et al.* 1994; Allen *et al.* 1994a,b). Two of these methods are described briefly here.

* The usual intermediate factor used is 3 because it is the approximate logarithmic mean of 1 and 10. The choice of 3, instead of 5 for example, reflects both the expected precision of the UFs (about 1 digit, log base 10) and the view that it is not generally possible to be more precise in considering the nuances of these areas of uncertainty than about half-way.

Table 2.3 Description of typical uncertainty and modifying factors in the development of subthreshold doses for several groups*

Uncertainty factors (UFs)†	Guidelines ‡	UF value				
		Health Canada Agency§	IPCS	RIVM	USATSDR	USEPA
Interhuman (or intraspecies)	Generally use when extrapolating from valid results from studies of prolonged exposure to average healthy humans. This factor accounts for the variation in sensitivity among humans and is composed of toxicokinetic and toxicodynamic components	1–100	10 (3.16 × 3.16)	10	10	10
Experimental animal to human	Generally use when extrapolating from valid results of long-term studies on experimental animals when results of studies of human exposure are not available or are inadequate. This factor accounts for the uncertainty in extrapolating animal data to humans and is also composed of toxicokinetic and toxicodynamic uncertainties		10 (2.5 × 4.0)	10	10	10
Subchronic to chronic	Generally use when extrapolating from less than chronic results on experimental animals or humans. This factor accounts for the uncertainty in extrapolating from less than chronic NOAELs or LOAELs to chronic NOAELs or LOAELs	1–100	~100	10	NA¶	≤10
LOAEL to NOAEL	Generally use when extrapolating a LOAEL to a NOAEL. This factor accounts for the experimental uncertainty in developing a subthreshold dose from a LOAEL, rather than a NOAEL			10	10	≤10
Incomplete database	Generally use when extrapolating from valid results in experimental animals when the data is 'incomplete'. This factor accounts for the inability of any single study to adequately address all possible adverse outcomes			NA	NA	≤10
Modifying factor	Generally use upon a professional assessment of scientific uncertainties of the study and data base not explicitly treated above (e.g. the number of animals tested)	~10	1–10	NA	NA	0< to ≤10

* Source: Dourson (1994), Jarabek (1994), IPCS (1994), Meek *et al.* (1994) and Rademaker & Linders (1994).
† Note: The maximum uncertainty factor used with the minimum confidence data base is generally 10 000. LOAEL, lowest-observed-adverse-effect level; NOAEL, no-observed-adverse-effect level.
‡ Professional judgement is required to determine the appropriate value to use for any given UF. The values listed in this table are nominal values that are used frequently by these agencies.
§ Abbreviations used: IPCS (International Programme on Chemical Safety); RIVM (The Netherlands National Institute of Public Health and Environmental Protection); USATSDR (US Agency for Toxic Substances and Disease Registry); USEPA (US Environmental Protection Agency).
¶ ATSDR develops minimum risk levels (MRLs) for specified durations of exposure, and generally does not extrapolate among durations. Therefore, an uncertainty factor for extrapolation between subchronic and chronic exposures is not used.

Benchmark dose

Another form of quantitative risk assessment of non-cancer end-points is the benchmark dose (BMD) method. The USEPA (1995a) has defined the BMD as: 'a statistical lower confidence limit for a dose that produces a predetermined change in response rate of an adverse effect ... compared to background.'

This method, which was first described by Crump (1984) and Dourson *et al.* (1985), was developed in an attempt to remedy some notable shortcomings of the use of a NOAEL in the subthreshold-dose approach described above. For example, the NOAEL is limited by the experimental doses chosen by the investigators in the toxicity studies. The larger the dose spacing, the less accurate the experimental NOAEL (or LOAEL) is apt to be. Also, the slope of the dose–response curve provides valuable information that is not used explicitly in this approach (although it may influence the choice of uncertainty factors). The BMD method attempts to use more of the available dose–response information by fitting a mathematical model to the data and then determining the dose associated with a specified incidence of adverse effect. In this way, the BMD is not limited to the experimental doses chosen by the investigators.

Although the BMD method offers some advantages over the NOAEL, it can be used only in cases where data are available that are suitable for modelling. It is not, therefore, a replacement for the NOAEL, but should be considered as an additional method that may offer advantages for some risk assessments.

There are a number of decisions to be made in applying the BMD method, for example: which mathematical model to use; what degree of confidence limit to use; what incidence rate to predetermine as the benchmark response (e.g. a 1%, 5% or 10% incidence of an effect). For more information, the reader is referred to a guidance document on the use of the benchmark-dose approach in risk assessment that was issued by USEPA's Risk Assessment Forum (USEPA 1995a).

Categorical regression

Another method that has been proposed for quantitative dose–response analysis for non-cancer effects is that of categorical regression. This involves statistical regression on severity categories of overall toxicity (Hertzberg & Miller 1985; Hertzberg 1991; Hertzberg & Wymer 1991). By assigning severity categories, all adverse effects may be taken into account rather than just the critical effect. Categorical regression also allows use of group data (i.e. at the dose-group level, not individual animals) as well as toxicity data from multiple studies. The results of the regression can then be used to develop a subthreshold dose much as the BMD is used.

Categorical regression can also provide information about relative risks from exposures exceeding the RfD. The NOAEL and BMD approaches are limited in that they are focused on the determination of 'safe' and 'acceptable' intake levels (i.e. they are point estimates designed to be below the population threshold for toxicity). In situations where exposures may exceed these levels, however, information is needed to help determine the urgency of a situation. Herein lies one of the advantages of categorical regression because it provides information about increasing toxicity with increasing dose rate. If human data are available, categorical regression can be used to actually estimate potential risk above the RfD or RfC. With only animal data, categorical regression can help prioritize risks based on how quickly the toxicity severity changes with dose.

2.3.2 Cancer end-points

The elicitation of a carcinogenic response traditionally has been presumed to occur without a threshold. Because of this, it has often been assumed by several regulatory agencies (e.g. USEPA 1986; Rademaker & Linders 1994) that any dose of a carcinogen is associated with some increased risk. As a result, dose–response assessment for carcinogens often focused on determining a *de minimis* risk level, frequently expressed as the risk of one-in-a-million, by using a linear model to extrapolate risks down to low-dose levels. Other groups, such as Health Canada (Meek

et al. 1994), do not advocate extrapolation beyond the range of observable data, but rather use a margin-of-exposure approach (described more fully below). In EPA's 1996 proposed revisions to the carcinogen risk assessment guidelines, an option of using a margin-of-exposure (MOE) approach is also described.

Exposure to a carcinogenic agent often causes more than one tumour type. Similar to the process used for the evaluation of non-cancer toxicity, the risk assessor must evaluate the data to determine which end-point(s) occur(s) at the lowest dose. Unless there are data to support otherwise, it is generally assumed that humans may be as sensitive as the most sensitive animal model. After identifying the study(ies) that is(are) most appropriate for developing a quantitative risk estimate, the next step is to transform the doses to which the animals were exposed into human equivalent doses. For example, in the absence of a chemical-specific model, the USEPA (1996) recommends the use of a cross-species scaling factor of (body weight)$^{3/4}$ for oral exposures, and default methodology (USEPA, 1994a) for estimating respiratory deposition and absorption of particles and gases for inhalation exposures. Finally, the dose–response data are modelled to determine the carcinogenic potency of the chemical at low doses.

Estimating risk with mathematical models

Dose–response assessment for carcinogens is concerned with estimating the central estimate and/or the upper confidence bound for carcinogenic risk associated with environmental exposures. Alternatively, risk managers may be interested in setting standards for exposures by various media (e.g. air, drinking water) based on a carcinogenic risk level that is considered to be *de minimis* (e.g. one-in-a-million excess risk). The cancer bioassays generally used in the dose–response assessment, however, are performed in laboratory animals at very high doses relative to levels at which humans may actually be exposed. These high doses are necessary in order to produce a statistically measurable effect given the relatively small number of animals used. The nature of the curve at levels of exposure below the lowest experimental dose is not known. Many models have been developed to estimate cancer risk in this low-dose region.

A common model used to perform this extrapolation is adapted from the multistage model. This model assumes that cancer is the result of a sequence of changes in a cell or organ and that exposure to a carcinogen can increase the transition rate between these stages, resulting in malignancy (Armitage & Doll 1954, 1961; Crump *et al.* 1976). The 95% upper confidence limit of the linear component of this model (often referred to as the q_1^*) has been used by the USEPA (1986) as an upper bound estimate of cancer potency because it is numerically more stable than a central estimate and also is in keeping with the low-dose linear approach adopted for cancer-risk assessments. Although there are no data to demonstrate that the linearized multistage model is more appropriate than any other model, it has been used as the default because it provides a plausible and stable upper bound estimate that is not likely to underestimate the cancer risk. The USEPA and others recognize that at very low doses the response could be as low as zero.

In the USEPA's 1996 revised guidelines, it is proposed that the dose–response assessment be considered as a two-step process. The first step is to fit a model to data in the observed range only. If sufficient data are available, a biologically based model is the preferred approach. Also, there may be cases in which data other than tumour incidence (e.g. information on DNA adducts) can be used to extend dose–response below the observable range. The outcome of the first step is the estimation of an ED_{10} or LED_{10}. The ED_{10} (effective dose at the 10% level) is the dose associated with a 10% increase over background in the end-point being measured (e.g. tumour incidence or other). The LED_{10} is the lower 95% confidence limit on this dose.

The second step under the USEPA's 1996 revised guidelines is to use an extrapolation procedure to estimate risk in the low-dose region (i.e the range of human exposure) if it is appropriate to do so. For cases where data do not support the development of a biologically based model, and for which the dose–response relationship is thought to be linear, the proposed guidelines suggest terminating the model in the range of experimental data, and

drawing a straight line to the origin. Whereas the linearized multistage model may still be used in modelling the data in the experimental range, justification for using this (or any other) model needs to be provided.

For chemicals that have a non-linear dose–response relationship, the proposed guidelines advocate the use of an MOE analysis. The MOE is the LED_{10} (or another predetermined starting point within the range of observation) divided by the exposure of concern. The risk manager then decides whether the margin of exposure is large enough to satisfy management policy criteria. The proposed guidelines suggest that a factor of 100 be used as a science policy default position to reflect allowances for intra- and interspecies variability. Chemical-specific data can then be used to adjust this factor upward or downward as appropriate.

This type of MOE analysis is similar to that used currently by Health Canada (Meek *et al.* 1994). Main differences between the MOE approaches of the agencies are twofold: Health Canada determines a TD_{05} (the dose associated with a 5% increase over background of tumour incidence), whereas the USEPA is proposing determination of an ED_{10}. Secondly, the USEPA uses the 95% lower confidence bound on the ED_{10} whereas Health Canada uses the central estimate.

Approaches with uncertainty factors

A number of organizations have developed other methods for quantitative dose–response assessment for carcinogens. Moolenaar (1994) has described the similarities and differences between approaches used by the USEPA, the UK, Denmark, the European Union (EU), The Netherlands, and Norway. He points out that the USEPA is the only organization to have described carcinogenic risk in terms of an 'upper bound' (i.e. the 95th percentile of the slope of the dose–response curve in the low-dose region).

For example, each of the aforementioned groups has developed separate methods for dealing with genotoxic versus non-genotoxic carcinogens. Norway does not perform low-dose extrapolation for any carcinogens, but rather uses the TD_{50} to determine a potency classification for a carcinogen. For non-genotoxic carcinogens, the UK, the EU and The Netherlands use a subthreshold dose approach: they set ADIs using the method described above for non-cancer toxicity. For genotoxic carcinogens thought to have no threshold, The Netherlands extrapolates linearly from the lowest experimental dose having an increased incidence of tumours. Clearly, there are many variations in the ways that dose–response assessment for carcinogens can be, and are, performed. A common theme among all of these groups is that the mechanism by which the agent is believed to cause cancer is playing a greater role in the way in which the dose–response assessment is approached.

2.4 EXPOSURE ASSESSMENT

> *Exposure assessment* involves specifying the population that might be exposed to the agent of concern, identifying the routes through which exposures can occur, and estimating the magnitude, duration, and timing of the doses that people might receive as a result of their exposure.
>
> NAS, 1994

Environmentally relevant routes of exposure for humans are inhalation, oral, and dermal. An exposure assessment may include a component for each, such as an assessor would conduct when investigating the potential impact of a point source of pollution. In such a multimedia investigation, an exposure assessment is initiated by estimating the amount and rate at which a toxic agent is released from a given source. Fate and transport models are then used to estimate the movement of the agent through environmental media to which humans may be exposed. A number of models is available for use in estimating transport and fate; many of these are described in the USEPA's exposure assessment guidelines (USEPA 1992). Table 2.4 provides guidance in characterizing the exposure assessment step.

An exposure assessment may also be focused on one particular medium and one route of exposure, for example, the oral intake of a chemical from drinking water. This type of exposure assessment may be used, for example, to determine whether there is sufficient exposure of humans to a chemical in a given medium to warrant regulation.

Exposure can be determined directly, through

Table 2.4 Considerations in characterizing exposure. (Information from USEPA 1995b.)

What are the most significant sources and pathways of environmental exposure, presently and in the future (if appropriate)?
Are there data on other sources of exposure and what is the relative contribution of different sources of exposure?
Describe the populations that were assessed, including highly exposed groups and highly susceptible groups
Describe the basis for the exposure assessment, including any monitoring, modelling, or other analyses of exposure distributions such as Monte Carlo.
Describe the range of exposures to 'average' and 'high-end' individuals, the general population, high exposure group(s), children, susceptible populations
How was the central tendency estimate developed? What factors and/or methods were used in developing this estimate?
Are there highly exposed subgroups and how are they accounted for?
Is there reason to be concerned about cumulative or multiple exposures?
What are the results of different approaches, i.e. modelling, monitoring, probability distributions?
What are the limitations of each and the range of most reasonable values?
What is the confidence in the results obtained and the limitations to the results?

personal monitoring devices, or indirectly, through environmental monitoring. If environmental monitoring is used, then the assessor must estimate the extent to which individuals may be exposed to the media for which monitoring data are available. Risk assessment scientists often use default values for these assessments (e.g. assuming an inhalation rate of 20 m^3 per day or consumption of 2 l of water daily).

A need exists to estimate the distribution of exposures that may result to individuals and populations. For example, the USEPA (USEPA 1992) recommends assessing exposure to the total population, and also for assessing the upper end of the exposure distribution; i.e. a 'high-end exposure estimate' and a 'theoretical upper bounding estimate'.

2.5 RISK CHARACTERIZATION

> *Risk characterization* involves integration of information from the first three steps to develop a qualitative or quantitative estimate of the likelihood that any of the hazards associated with the agent of concern will be realized in exposed people. This is the step in which risk-assessment results are expressed. Risk characterization should also include a full discussion of the uncertainties associated with the estimates of risk.
>
> NAS, 1994

Risk characterization is the final step of the risk assessment process, in which information from the hazard identification, dose–response and exposure steps are considered together to determine and communicate the actual likelihood of risk to exposed populations. The risk characterization discussion includes an evaluation of the overall quality of the data, the specific assumptions and uncertainties associated with each step, and the level of confidence in the resulting estimates.

Specific key qualities, or attributes, of risk characterizations have been identified (AIHC 1992; USEPA 1995b). These attributes include transparency in decision making, clarity in communication, consistency and reasonableness. Exercising transparency and clarity result in scientific conclusions being identified separately from policy judgements. In addition, default values, assumptions and uncertainties are disclosed so that the end-user can better identify what is based on data and what is assumed. Greater consistency in the terminology used, along with definitions and assumptions, will provide for better comparability across assessments. In addition, use of standard

descriptors, such as those outlined in the USEPA's Exposure Assessment Guidelines (USEPA 1992), reduce confusion and lead to greater understanding. Lastly, the risk characterization should be reasonable and balanced in its presentation. The information and conclusions should be presented in such a fashion that they are clearly understood by the intended audience. The ultimate goal of risk characterization is to provide the decision makers with enough information, presented in a comprehensible fashion, that they understand what is known and unknown about the risk to human health from the situation being evaluated, thereby leading to the best possible risk-based decisions.

In order for risk assessors to meet this goal, they must understand the need for the risk assessment and its intended end-use. Risk assessors should meet with the decision makers and engage them in the process throughout. Involving the risk managers and decision makers will help the risk assessor to meet a level of detail and analysis appropriate for the situation (e.g. initial screening versus national regulation). By communicating with the end-user, risk assessors can ensure that the risk manager will comprehend the results of the analysis.

Involvement of the end-user supports taking an iterative approach to the risk assessment. For example, if the risk assessment is for a contaminated site, it would be very useful to have the exposure assessment scientists involved in developing a monitoring plan and in reviewing initial results so that the monitoring could be refined to collect the most useful data. Only if a first, conservative screening indicates that some level for concern is warranted would a more in-depth analysis be pursued. The iterative approach to risk assessment assures better use of limited resources to address problems.

2.5.1 Uncertainty and variability

The field of risk assessment is increasingly utilizing uncertainty and sensitivity analyses to better assess risks to human health. Critical to a complete risk characterization is a full discussion of the uncertainty within each analysis and that related to the overall assessment. Uncertainty discussions are important because they form the basis for the overall judgement as to the adequacy of the data and conclusions drawn from it. In addition, highlighting of uncertainties can identify areas where the collection of additional data may reduce the uncertainty and strengthen the risk assessment. An uncertainty discussion includes the quality and quantity of data available (toxicity and exposure), identification of data gaps, use of default assumptions and parameter values, and the uncertainties in the models used.

Crucial to a discussion of uncertainty is maintaining a clear distinction between uncertainty and variability within each step of the process. The USEPA distinguishes between these two concepts in its risk characterization guidance (USEPA 1995b). Variability describes inter-individual, spatial or temporal differences within an animal or human population or within monitoring data. It reflects the heterogeneity of the data. Uncertainty, on the other hand, applies to areas for which data are unknown. There are uncertainties associated with both dose–response or fate and transport models; an uncertainty analysis would evaluate the basis for the model and validation of the model.

Given the extensive use of modelling to estimate exposure in particular, the uncertainty related to the chosen parameters can have a great impact on the resulting risk estimates. Risk assessors must be careful to identify the parameter values and their sources so that others can evaluate their appropriateness and impact on the final results. Risk assessors can use probability density functions and/or likelihood distributions to characterize uncertainty quantitatively. Monte Carlo analysis is one statistical procedure used.

To summarize a risk characterization, the risk assessor should consider questions such as those listed in Table 2.5. These questions can be used to help outline a discussion of risk conclusions and comparisons. These questions build upon those in Tables 2.2–2.4 relating to hazard identification, dose–response assessment, and exposure assessment, respectively.

2.6 SUMMARY

This chapter has outlined a process to assess the

Table 2.5 Questions to assist in developing a risk characterization summary. (Information from USEPA 1995b.)

Risk conclusions
What is the overall picture of risk and the specific risk estimates and/or ranges?
For the hazard identification, dose–response and exposure assessment steps:
What are the major conclusions and strengths?
What are the major limitations and uncertainties?
What are the science policy options, and what other alternatives were considered?
Risk context
What are the qualitative characteristics of the hazard (e.g. voluntary versus involuntary, one population segment versus another)? Comment on any risk perception studies related to this type of hazard
What are the alternatives to this hazard and how do the risks compare?
How does this risk compare with other risks?
Are there significant community concerns which influence public perception of risk? Are there perceived or actual inequities in distribution of risks and benefits?
Other information
Have other risk assessments been done on this chemical and were there significantly different conclusions?

human health risks from exposure to chemicals in the environment. The methods used to assess and characterize these risks are being improved and expanded upon. Recently, the area of quantifying uncertainty has greatly expanded, along with developing better procedures to use the available data more fully. Improved methods to characterize and communicate the results of the risk assessment process are also being explored. Although significant research has been done to develop methods to characterize exposure and health effects, much more is needed to provide risk decision makers with the accurate estimates they need to make reasonable and cost-effective decisions. Risk assessors will always be faced with gaps in data and scientific understanding. How they deal with these uncertainties will determine how useful risk assessment will be to decision makers and the public.

2.7 REFERENCES

Abdel-Rahman, M.S. & Dourson, M.L. (Guest Eds) (1995) *Proceedings of EPA Uncertainty Factor Workshop. Journal of Human and Ecological Risk Assessment*, 1(5).

AIHC (1992) *Improving Risk Characterization. Summary of Workshop held September 26 and 27, 1991.* American Industrial Health Council.

Allen, B.C., Kavlock, R.J., Kimmel, C.A. & Faustman, E.M. (1994a) Dose–response assessment for developmental toxicity. II. Comparison of generic benchmark dose estimates with no observed adverse effect levels. *Fundamental and Applied Toxicology*, 23(4), 487–495.

Allen, B.C., Kavlock, R.J., Kimmel, C.A. & Faustman, E.M. (1994b) Dose–response assessment for developmental toxicity. III. Statistical Models. *Fundamental and Applied Toxicology*, 23(4), 496–509.

Armitage, P. & Doll, R. (1954) The age distribution of cancer and a multistage theory of carcinogenesis. *British Journal of Cancer*, 8, 1–12.

Armitage, P. & Doll, R. (1961) Stochastic models for carcinogenesis. In: *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4, pp. 19–38. University of California Press, Berkeley.

Ashby, J., Doerrer, N.G., Flamm, F.G. *et al.* (1990) A scheme for classifying carcinogens. *Regulatory Toxicology and Pharmacology*, 12, 270–295.

Barnes, D.G. & Dourson M.L. (1988) Reference dose (RfD): description and use in health risk assessments. *Regulatory Toxicology and Pharmacology*, 8, 471–486.

Calabrese, E.J. (1985) Uncertainty factors and interindividual variation. *Regulatory Toxicology and Pharmacology*, 5, 190–196.

Calabrese, E.J. & Gilbert, C.E. (1993) Lack of total independence of uncertainty factors (UFs): implications for the size of the total uncertainty factor. *Regulatory Toxicology and Pharmacology*, 17, 44–51.

Calabrese, E.J., Beck, B.D. & Chappell, W.R. (1992) Does the animal to human uncertainty factor incorporate interspecies differences in surface area? *Regulatory Toxicology and Pharmacology*, 15(2), 172–179.

Crump, K.S. (1984) A new method for determining allowable daily intakes. *Fundamental and Applied Toxicology*, **4**, 854–871.

Crump, K.S. (1986) Letter to the Editor: reply. *Fundamental and Applied Toxicology*, **6**, 183–184.

Crump, K.S., Hoel, D.G., Langley, C. & Peto R. (1976) Fundamental carcinogenic processes and their implications for low dose risk assessment. *Cancer Research*, **36**, 2973–2979.

DeRosa, C.T., Stara, J.F. & Durkin, P.R. (1985) Ranking of chemicals based upon chronic toxicity data. *Toxicology and Industrial Health*, **1**(4), 177–192.

Dourson, M.L. & Stara, J.F. (1983) Regulatory history and experimental support of uncertainty (safety) factors. *Regulatory Toxicology and Pharmacology*, **3**, 224–238.

Dourson, M.L., Hertzberg, R.C., Hartung, R. & Blackburn, K. (1985) Novel approaches for the estimation of acceptable daily intake. *Toxicology and Industrial Health* **1**(4), 23–41.

Dourson, M.L., Hertzberg, R.C. & Stara, J.F. (1986) Letter to the Editor. *Fundamental and Applied Toxicology*, **6**, 182–183.

Dourson, M.L., Knauf, L.A. & Swartout, J.C. (1992) On Reference Dose (RfD) and its underlying toxicity data base. *Toxicology and Industrial Health*, **8**(3), 171–189.

Dourson, M.L. (1994) Methods for establishing oral reference doses (RfDs). In: *Risk Assessment of Essential Elements* (eds W. Mertz, C.O. Abernathy & S.S. Olin), pp. 51–61. ILSI Press, Washington DC.

Farland, W. & Dourson, M.L. (1992) Noncancer health endpoints: approaches to quantitative risk assessment. In: *Comparative Environmental Risk Assessment* (ed. R. Cothern, pp. 87–106. Lewis Publishers, Boca Raton. FL.

Faustman, E.M. Allen, B.C., Kavlock, R.J. & Kimmel, C.A. (1994) Dose–response assessment for developmental toxicity. I. Characterization of database and determination of no observed adverse effect levels. *Fundamental and Applied Toxicology*, **23**(4), 478–486.

Hartley, W.R. & Ohanian, E.V. (1988) The use of short-term toxicity data for prediction of long-term health effects. In: *Trace Substances in Environmental Health—XXII*, University of Missouri, May 23–26 (ed. D.D. Hemphil), pp. 3–12.

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**(4), 415–426.

Hertzberg, R.C. (1989) Fitting a model to categorical response data with application to species extrapolation of toxicity. *Health Physics*, **57**, 405–409.

Hertzberg, R.C. (1991) Quantitative extrapolation of toxicological findings. In: *Statistics in Toxicology* (eds D. Krewski & C. Franklin). Gordon and Breach Science Publishers, New York.

Hertzberg, R.C. & Miller, M. (1985) A statistical model for species extrapolation using categorical response data. *Toxicology and Industrial Health*, **1**, 43–57.

Hertzberg, R.C. & Wymer, L. (1991) Modeling the severity of toxic effects. In: *Proceedings' Papers from the 84th Annual Meeting and Exhibition*, June 16–21, British Columbia. Air and Waste Management Association.

Hertzberg, R.H. & Dourson, M.L. (1993) Using categorical regression instead on a NOAEL to characterize a toxicologist's judgment in noncancer risk assessment. In: *Proceedings, Second International Symposium on Uncertainty Modeling and Analysis* (ed. B.M. Ayyub), pp. 254–261. IEEE Computer Society Press, Los Alamitos, CA.

IARC (1978) *Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*, Vol. 17. International Agency for Research on Cancer, Lyon.

International Programme on Chemical Safety (IPCS) (1994) *Environmental Health Criteria No. 170: Assessing Human Health Risks of Chemicals. Derivation of Guidance Values for Health-based Exposure Limits*. World Health Organization, Geneva.

Jarabek, A.M. (1994) Inhalation RfC methodology: dosimetric adjustments and dose–response estimation of noncancer toxicity in the upper respiratory tract. *Inhalation Toxicology*, **6**(Suppl.), 301–325.

Kimmel, C.A. & Gaylor, D.W. (1988) Issues in qualitative and quantitative risk analysis for developmental toxicity. *Risk Analysis*, **8**(1), 15–20.

Kimmel, C.A. Wellington, D.G., Farland, W. *et al.* (1988) Overview of a workshop on quantitative models for developmental toxicity risk assessment. *Environmental Health Perspectives*, **79**, 209–215.

Kroes, R., Munro, I. Poulsen, E. (eds) (1993) Scientific Evaluation of the Safety Factor for the Acceptable Daily Intake. *Food Additives and Contaminants*, **10**(3), 269–373.

Krewski, D., Brown, C. & Murdock, D. (1984) Determining 'safe' levels of exposure: safety factors or mathematical models? *Fundamental and Applied Toxicology*, **4**, 5383–5394.

Leisenring, W. & Ryan, L. (1992) Statistical properties of the NOAEL. *Regulatory Toxicology and Pharmacology*, **15**(2), 161–171.

Lewis, S.C., Lynch J.R. & Nikiforov, A.I. (1990) A new approach to deriving community exposure guidelines from no-observed-adverse-effect levels. *Regulatory Toxicology and Pharmacology*, **11**, 314–330.

Lu, F. 1983. Toxicological evaluations of carcinogens and noncarcinogens: pros and cons of different approaches. *Regulatory Toxicology and Pharmacology*, **3**, 121–132.

Lu, F. 1985. Safety assessments of chemicals with thresholded effects. *Regulatory Toxicology and Pharmacology*, **5**, 460–464.

Lu, F. (1988) Acceptable Daily Intake: inception,

evolution, and application. *Regulatory Toxicology and Pharmacology*, **8**, 45–60.

Meek, M.E., Newhook, R., Liteplo, R.G. & Armstrong, V.C. (1994) Approach to assessment of risk to human health for priority substances under the Canadian Environmental Protection Act. *Environmental Carcinogenesis and Ecotoxicology Reviews*, C**12**(2), 105–134.

Moolenaar, R.J. (1994) Carcinogen risk assessment: international comparison. *Regulatory Toxicology and Pharmacology*, **20**, 302–336.

Munro, I.C. & Krewski, D.R. (1981) Risk assessment and regulatory decision making. *Food and Cosmetics Toxicology*, **19**, 549–560.

NAS (National Academy of Sciences) (1983) *Risk Assessment in the Federal Government: Managing the Process*. National Academy Press, Washington, DC.

NAS (National Academy of Sciences) (1994) *Science and Judgment in Risk Assessment*. National Academy Press, Washington, DC.

Pohl, H.R. & Abadin, H.G. (1995) Utilizing uncertainty factors in minimal risk levels derivation. *Regulatory Toxicology and Pharmacology*, **22**, 180–188.

Rademaker, B.C. & Linders, J.B.H.J. (1994) *Progress Report 3: Estimated-concentrations-of-no-concern of polluting Agents in Drinking Water and Air for Humans*. National Institute of Public Health and Environmental Protection, Bilthoven, The Netherlands.

Renwick, A.G. (1991) Safety factors and establishment of acceptable daily intake. *Food Additives and Contaminants*, **8**(2), 135–150.

Renwick, A.G. (1993) Data derived safety factors for the evaluation of food additives and environmental contaminants. *Food Additives and Contaminants* **10**(3), 275–305.

USEPA (1986) *Guidelines for Carcinogen Risk Assessment*. US Environmental Protection Agency, 51 Federal Register 33992, pp. 1–17.

USEPA (1991) *Guidelines for Developmental Toxicity Risk Assessment*. Federal Register, Vol. 56, No. 234, pp. 63798–63826.

USEPA (1992) *Guidelines for Exposure Assessment*. US Environmental Protection Agency, 57 Federal Register, 22888–22938.

USEPA (1994a) *Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry*. US Environmental Protection Agency, Office of Health and Environmental Assessment, Washington, DC, EPA/600/8-90-066F.

USEPA (1994b) *Draft Revisions to the Guidelines for Carcinogen Risk Assessment*. US Environmental Protection Agency, Office of Research and Development, Washington, DC, External Review Draft.

USEPA (1995a) *The Use of the Benchmark Dose Approach in Health Risk Assessment*. Risk Assessment Forum, February, US Environmental Protection Agency, Washington, DC, EPA/630/R-94/007.

USEPA (1995b) *Policy and Guidance for Risk Characterization*. Memo from Carol M. Browner, Administrator to Assistant Administrators, March 21. US Environmental Protection Agency, Washington, DC.

USEPA (1996) *Proposed Guidelines for Carcinogen Risk Assessment*. US Environmental Protection Agency, Office of Research and Development, Washington, DC, EPA/600/P-92/003C.

Zielhuis, R.L. & van der Kreek, F.W. (1979) The use of a safety factor in setting health based permissible levels for occupational exposure. *International Archives of Occupational Environmental Health*, **42**, 191–201.