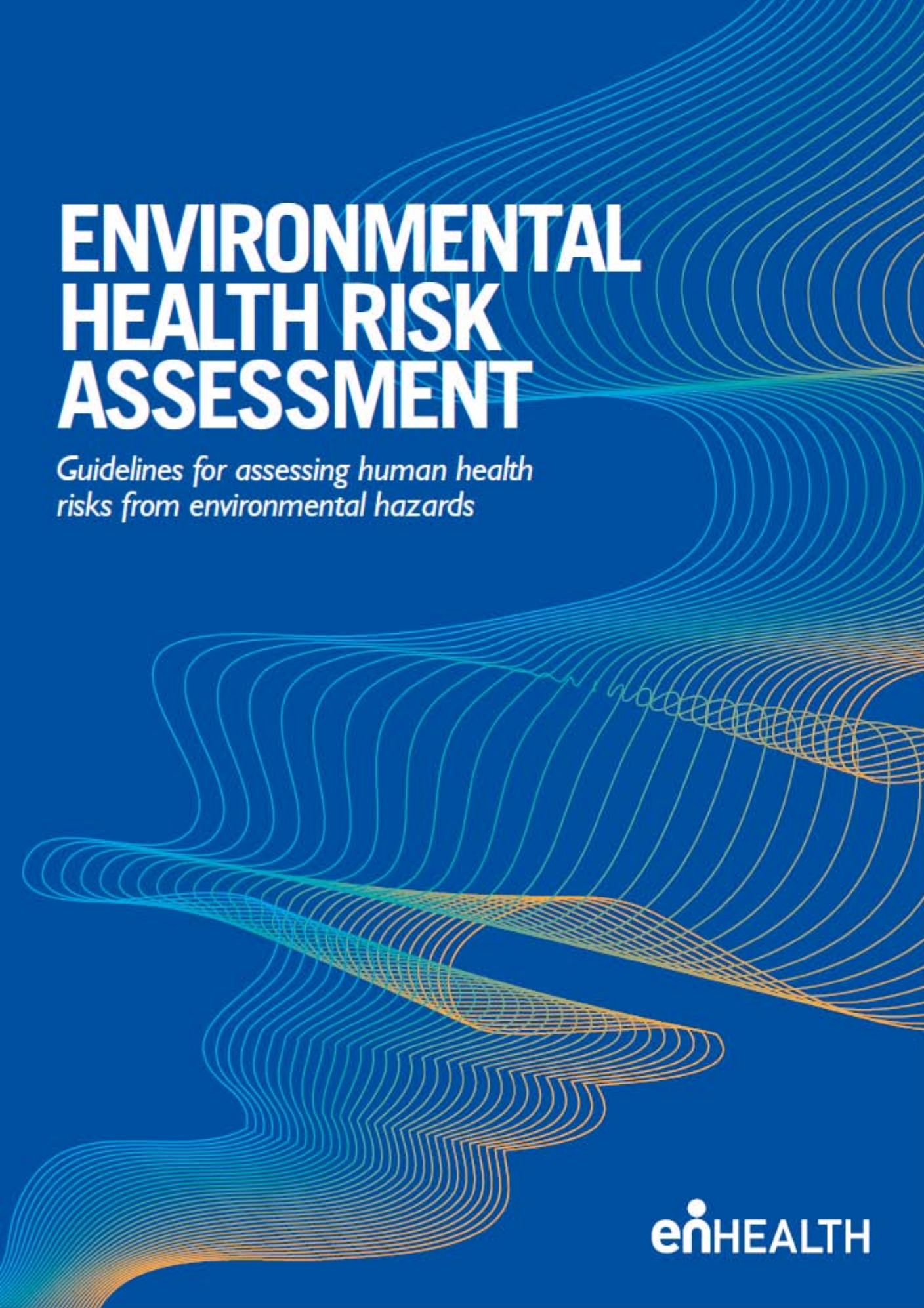
ENVIRONMENTAL HEALTH RISK ASSESSMENT

Guidelines for assessing human health risks from environmental hazards



enHEALTH

Contents

FIGURES

TABLES

OBJECTIVES

This enHealth document provides a national approach to environmental health risk assessment.

Risk assessments are being undertaken for a wide variety of projects by governments and industry. Environmental health agencies need to be able to assess their content and approach against a benchmark. The document presents a general environmental health risk assessment methodology applicable to a range of environmental health hazards. The focus is on chemical hazards in the first instance, but the core methodology can also be applied to physical (e.g. radiation, noise) and microbiological hazards. The core methodology is intended to be able to accommodate specialised 'modules' that will deal with issues such as physical and microbiological hazards and mixtures as they become available. The links to risk management and community consultation/risk communication will be identified. The document emphasises the importance of prior planning and appropriate scoping in the design phase of a risk assessment. It further notes that appropriate consultation with all stakeholders, but particularly with decision makers, is essential to ensure the conceptual models and methodologies used are adequate to address the desired outcomes.

Due to the complexity and scale of the environmental health risk assessment process a concise 'cookbook' is not practicable, although this document attempts to provide pragmatic and user-friendly advice. Similarly, the situation-specific issues are often sufficiently complex and situation-specific that a manageable and complete algorithm for decision making cannot be drafted; however, the document provides a series of guidelines and checklists to assist the decision-making process. Where possible, the document is prescriptive about certain aspects of risk assessment. Having specific requirements for the content of investigations and having them presented in uniform, coherent and logically developed reports will enable more efficient, accurate, timely and transparent decision making and a greater consistency of environmental health decision making across Australia. However, contemporary paradigms of risk assessment acknowledge that stakeholders may sometimes impose unrealistic demands on the available science of risk assessment and that data gaps and uncertainties may limit the options for establishing all the available risk management options. Such knowledge gaps should not deter decision makers from considering the range of options within an appropriate science-based framework, but this should always be done with a full and frank acknowledgement of the inherent uncertainties.

AUDIENCE

This enHealth document is primarily intended to be used by: environmental health agencies reviewing risk assessments; people preparing risk assessments for environmental health agencies; and those regulatory agencies reviewing risk assessments. It is also intended to be of assistance to a broader audience seeking information about processes of environmental risk assessment in Australia.

Risk assessors should have a basic grounding in epidemiology, toxicology and chemistry.

ACKNOWLEDGMENTS

This enHealth document is an update of the original enHealth guidance document *Environmental health risk assessment: Guidelines for assessing human health risks from environmental hazards*, first published in 2002. This 2012 update has been funded by the enHealth committee through a contract with the Australian Centre for Human Health Risk Assessment (ACHHRA) at Monash University. The principal author of the update is Professor Brian Priestly, director of ACHHRA and president of the Australasian College of Toxicology and Risk Assessment (ACTRA), with assistance in collecting and reviewing relevant literature by Dr Jennifer Ong, an ACHHRA research associate.

Toxikos Pty Ltd, under the direction of its principal toxicologist Dr Roger Drew, provided significant input into sections dealing with exposure assessment, including a substantial re-drafting of the *Australian exposure factor guidance* document initially drafted by Dr Andrew Langley. Dr Roger Drew also provided valuable assistance with planning the review of the document, including advice to ensure the document adopts a user-friendly approach to health risk assessment and addresses the pragmatic needs of risk assessment professionals. Mr John Frangos and Ms Tarah Hagen provided significant input into the revised *Australian exposure factor guidance* document.

The original 2002 version of the enHealth guidance was prepared by Dr Andrew Langley with assistance from various colleagues, including Mr Jack Dempsey, Dr Les Davies and Dr Jim Fitzgerald, who prepared the hazard identification section. Professor Steve Hrudey, Dr Andrew Langley, Dr Hugh Davis and Dr Les Davies provided extensive written editorial advice augmenting the comments made by people at the two workshops that reviewed the original document.

The assistance of other contributors to the 2002 edition was acknowledged in that edition. Since some of the text was carried through into the 2012 revision, in a revised format, their contributions are once again acknowledged with thanks.

The steering committee that oversaw the development of the 2012 update comprised:

Ms Jan Bowman (chair), Dr Kevin Buckett, Dr Peter Di Marco, Dr Liz Hanna, Dr Andrew Langley, Mr John Dempsey and Dr Roscoe Taylor, with secretariat assistance provided by the Environmental Health Unit of the Department of Health, Victoria.

All these people provided editorial comment and assistance through the preparation of the draft revisions, with further comments from staff of SA Health and Victorian Department of Health (Dr David Cunliffe, Dr Kateryna Babina, Dr Richard Evans and Mr Robert Mckenzie).

International peer review on the draft document was provided by Professor Steve Hrudey, University of Alberta, and ChemRisk LLC, San Francisco (Ellen Donovan and Hugh Scobie). Box 1 in Section 5.16 describing the regulatory experience with chloroform was drafted by Professor Hrudey.

Part 1: Basic concepts in environmental health risk assessment

Chapter 1: Introduction to environmental health risk assessment

Virtually all aspects of life involve exposure to risks (National Research Council – NRC 2008). Understanding the nature of risk, including the way people perceive threats to their health and the rational and emotive factors that govern that perception, is vital to developing appropriate ways to manage environmental health risks. Risk assessment can be a useful tool in managing environmental health risks.

1.1 WHAT IS RISK ASSESSMENT?

Risk assessment is the process of estimating the potential impact of a chemical, physical, microbiological or psychosocial hazard on a specified human population or ecological system under a specific set of conditions and for a certain time frame.

The scope of environmental health risk assessment (EHRA) can cover health impacts of:

- chemical pollutants and contaminants in air, water, soil and food
- pathogenic microbiological contaminants in food and water
- radiation sources
- electromagnetic fields (EMFs)
- climate and climate change

In all cases of the above impacts, priority is attached to evaluating the potential human health impacts. This update of enHealth guidance on EHRA focuses primarily on hazardous chemicals (and to a lesser extent, microbiological hazards). Risk assessment relating to radiation hazards, EMFs and climate change are covered elsewhere.

Risk assessment is intended 'to provide complete information to risk managers, specifically policymakers and regulators, so that the best possible decisions are made' (Paustenbach 1989 p. 28).

There are uncertainties related to risk assessment and it is important to make the best possible use of available information. It is equally, if not more important, to be able to explain to stakeholders in the EHRA processes how these uncertainties have been identified and managed.

Risk assessment gathers and organises information and enables:

- risks at a point in time (including baseline risks) and changes in risk over time to be estimated and to establish whether action is necessary
- assessments of new and different types of risk
- the identification and comparison of different factors that affect the nature and magnitude of the risk
- issues to be prioritised according to their levels of risk
- health guidance values (GVs) to be estimated for environmental hazards that can be used and will adequately protect public health, as a preface to setting risk-based standards for regulatory exposure limits as well as clean-up standards
- a comparison of the potential health impacts of various environmental health interventions (thus enabling cost-effectiveness estimates)
- risk-based policy making and consistent, transparent appraisal and recording of public health risks
- questionable theories, methods and data to be challenged and addressed by providing a clearly documented and open process (Covello & Merkhofer 1993).

Risk assessment is significantly influenced by science policy considerations (see NRC 2008 for an outline of American EHRA policies). Science policy on EHRA in Australia is somewhat fragmented, with various Commonwealth and state or territory authorities applying risk assessment policies and default approaches, which are often not explicitly laid out in legislation or regulations. The objective of this enHealth document is not to enunciate specific science policy relating to EHRA but to provide information to risk assessors on different approaches to EHRA methodology, and to provide guidance on how to use default values at various stages of an EHRA. The difficulties in establishing such defaults within a science policy context are discussed in some detail in Section 5.16, where there is a discussion on the selection of 'target risk' in the EHRA of carcinogens.

Risk assessment may be done as a relatively rapid 'desktop' study or 'screening' study for simple issues, or may be a large and complex process where there are significant health concerns. These processes may be designated as Tier 1, 2 or 3 processes (see Section 1.9). There are numerous models of risk assessment to suit the many contexts in which risk assessments are undertaken. Even limited measures of the level of risk can be valuable for identifying complex cause-and-effect processes and the most efficient means of addressing the risks.

In this context, the methods used in EHRA are inherently conservative[1] and highly protective of public health. This is especially true of 'screening' type risk assessments, which tend to use the most conservative assumptions about exposure and risk. These are generally termed Tier 1 risk assessments. A conservative approach is also taken when the EHRA is used as a basis for establishing environmental guidelines or standards. Conservatism is often built into an EHRA by using exposure estimates that represent 'worst case' or at least the upper percentiles of parameter distributions, rather than mean, average or typical values. Furthermore, exposure is usually considered to be constant over a substantial period of time (sometimes

1 In this context, 'conservative' is intended to imply a cautious approach to evaluating and managing the uncertainties inherent in a risk assessment, which reduces the probability of harm occurring.

an entire lifetime), whereas many environmental exposures are episodic, and may decline over time due to loss or degradation of the contaminant.

The conservatism in EHRA can sometimes lead to the development of risk-based GVs that are so far below the capacity of contemporary analytical techniques that compliance monitoring becomes impossible or impractical. In some cases, conservative risk-based GVs may be driven to levels below background concentrations, casting doubt on the credibility of the process.

It is important that assessors, users, regulators and members of the public recognise risk assessment may not always provide a compelling or definitive outcome. Some of the criticisms of risk assessment are as follows:

- Default values and assumptions are not realistic – a series of such unrealistic values or assumptions compounds the inaccuracy so that risks may be seriously overstated or understated if the default values are too conservative or insufficiently conservative, respectively.
- Interactions between agents (i.e. mixtures of agents) and the variability of response between individuals are commonly unknown and may be insufficiently taken into account.
- The use of default values and assumptions may become too rigid so that situation-specific data is not applied.
- The nature of the population to whom the risk assessment is to be applied regarding its exposure characterisation or susceptibility is often poorly defined.
- The uncertainties of risk assessment are often inadequately described, for example, specific point estimates are given that do not recognise uncertainty, or simplistic upper-bound estimates of uncertainty are used.
- There is an emphasis on cancer risk to the possible neglect of other adverse effects, for example, reproductive and developmental outcomes.
- In some situations, there may be insufficient scientific knowledge to be able to perform credible risk assessments.
- Risk assessment can be perceived to be tailored to provide a desired or predetermined outcome (NRC 1994).
- Excessive emphasis is given to the process of risk assessment rather than its content.
- The risk assessment process can become so 'bogged down' (NRC 2008) that it takes far too long to achieve useful or timely outcomes.
- The risk assessment process is used as a 'whitewash' or used to justify the continuation or increase of polluting activities.
- The efforts in risk assessment may be inappropriately distributed in cases where enormous effort is spent on complex modelling in cases where some targeted data collection could provide much more relevant and credible evidence.

Tal (1997) indicates that environmental groups identify a number of problems with the way risk assessments have been practised, including disempowerment and potential regulatory delays. Risk assessments should be designed and undertaken in ways that minimise these pitfalls.

1.2 WHEN TO UNDERTAKE RISK ASSESSMENT

The issues identification phase (see Chapter 2) will determine when to undertake a risk assessment. The need to undertake a risk assessment will be influenced by situation-specific factors. As such, the following list is indicative and not exhaustive. In general, risk assessments will be needed for products, processes, situations and activities where there is a plausible case that there could be an increased risk of significant health consequences for the human population from the product, process, situation or activity. A risk assessment can also be used to inform the selection of the safest option when making decisions about how to achieve a particular aim. A screening level comparative risk assessment could be used to compare the risks associated with various options when, for example, formulating a particular product or controlling pests.

Examples are:

- new additives to food or potable or recreational waters
- introduction of a new chemical under the NICNAS (National Industrial Chemicals Notification and Assessment Scheme) program (see Section 17.2)
- assessment of a contaminated site
- assessment of a major planning development, especially where hazards are anticipated
- assessment of pollution impacts at existing facilities
- changes to climate, landform, geography or demography that may impact on disease vectors and parasites
- situations where environmental standards or guidelines are unavailable
- environmental changes that will increase traffic flow and may increase the risk of injury or air pollution, such as new traffic corridors
- changes where impacts on environmental health factors may be permanent and irreversible
- changes that may impact on the microbiological or chemical safety of food chains and food supplies
- situations where there is a high level of public interest in or concern about environmental health issues
- situations where vulnerable populations may be affected by environmental health issues such as the location of schools
- legislative or policy changes
- designating housing setbacks from industry and transport corridors
- where health impact assessments are undertaken.

Risk assessment is inappropriate when it is a ritual rather than a meaningful process and should not be undertaken when:

- there is no data or an insufficient amount of data
- it is clear that the proposal, situation or activity is seen by health and other experts as having few potential risks to health
- risks may be likely, but the evidence is already well documented and it may be possible to develop evidence-based recommendations without the need for a comprehensive assessment
- there is an inability to take action or it is too late to take action
- there are insufficient resources
- the proposal is clearly politically or socially unacceptable.

Of relevance to risk assessment is Bardwell's reference (cited in Thornton & Paulsen 1998 p. 799) to a study that indicates that 'about 90 per cent of real world problem solving is spent:

- solving the wrong problem;
- stating the question so that it cannot be answered;
- solving a solution;
- stating questions too generically; or
- trying to get agreement on the answer before there is agreement on the question'.

1.3 TYPES OF RISK ASSESSMENT

1.3.1 Individual and population risk assessments

Risk assessments generally make risk estimates for defined groups or populations. The term 'receptors' is often used to designate people who may be exposed to an environmental hazard, and to whom the EHRA would be directed. Identification of 'receptor' locations and pathways by which they might be exposed is an integral part of any EHRA.

Individual risks are usually estimated for a hypothetical person with assumed characteristics for various durations of exposure (e.g. per year or per lifetime) or for different locations. The hypothetical individual is designed to represent the average person in the situation or the maximally exposed person. However, such risk estimates cannot be targeted to a specific person. The distinction between 'there is a risk' and 'I am at risk' is often difficult to explain to both the public and by regulators, especially when discussing very small probability estimates and this can lead to serious misunderstanding among stakeholders about the meaning of quantitative risk estimates (McAuley & Hrudey 2006. In the case of a lottery, a winner may be found, despite the small odds of winning, whereas in most quantitative risk assessments the probability of anyone being at risk is small and the probability of a specific individual being at risk is very much smaller.

Population risk may relate to the number of adverse health effects (e.g. fatalities, cancers or illnesses) in a population over a specified period of time or the rate of adverse effects for a given location or sub-population (Covello & Merkhofer 1993).

1.3.2 Qualitative and quantitative risk assessments

The level of risk can be described either qualitatively (i.e. by putting risks into categories such as 'high', 'medium' or 'low') or quantitatively (with a numerical estimate). Practical guidance on how to manage risks is the approach taken in AS/NZS ISO 31000:2009 (Standards Australia, 2009) and in the *Risk analysis framework* used by the Office of the Gene Technology Regulator to manage risks associated with genetically modified organisms (GMOs) (OGTR 2009). (See Sections 5.3, 17.6 and 17.7.)

Current risk assessment methods do not enable accurate quantitative estimates of risk for low levels of exposure to environmental hazards. Numerical estimates of risk can be presented, but caution must be exercised in assigning strict meaning to the numbers:

> *... a number is a number is a number ... and yet exactitude should not be confused with accuracy.*
>
> (Langley 2003 p. 166)

Complexity of the exposure conditions, variability in the environmental agents and exposed populations, and any inherent limitations in toxicological data may limit the accuracy of numerical risk estimates. While a degree of quantification may be possible for some components, such as data collection and exposure assessment, it is important that all uncertainties are reflected in the EHRA outcomes. Further discussion of qualitative and quantitative risk assessment appears in Chapter 5.

1.4 THE DISTINCTION BETWEEN RISK ASSESSMENT AND RISK MANAGEMENT

Risk assessment is a process that informs the risk management process. Risk assessors and risk managers should be sensitive to the distinctions between risk assessment and risk management. The enHealth framework for EHRA (see Figure 1) clearly differentiates risk assessment and management as separate but interlinked processes, with risk management following the risk characterisation phase of a formal risk assessment.

The development of risk management plans is outlined in detail in AS/NZS ISO 31000:2009. The important elements of a risk management framework are whether it:

- evaluates the external and internal contexts of the organisation tasked with implementing new or existing risk management plans or policies
- provides for accountability and transparency in the decision-making process
- ensures that resources are made available to measure and report on risk and risk mitigation procedures
- establishes internal and external communication and reporting mechanisms
- ensures that there are audit processes appropriate to the evaluation of the risk management strategies
- provides effective processes for collecting feedback and information for continuous improvement
- develops monitoring and review processes at the implementation stage of all risk management plans and strategies.

Risk assessors should generally strive to:

- generate a credible, objective, realistic and scientifically balanced analysis
- present information on the separate components of the risk assessment
- explain the confidence in each assessment by clearly delineating strengths, uncertainties and assumptions, along with the impacts of these factors (e.g. confidence limits, use of conservative/non-conservative assumptions) on the overall assessment.

The risk assessors should do this without considering issues such as cost of remediation, feasibility or how the scientific analysis might influence the regulatory or site-specific decision (United States Environment Protection Agency – US EPA 1995a). However, it is likely that a more thorough EHRA process (e.g. moving to a Tier 2 or Tier 3 analysis) may provide the risk manager with a suite of options for managing the identified risks. This should assist in determining the most cost-effective set of actions.

Risk assessment processes should be coherent and transparent. It is important that the basis of the decision making is clearly documented. This formal record should be clear, comprehensive and concise, and include a summary of the key data that has influenced the risk assessment and an appraisal of its quality (Advisory Committee on Dangerous Pathogens – ACDP 1996). Further guidance on compiling an EHRA report is in Chapter 7.

Risk assessment information is only one of several kinds of information used for decision making. The risk management decision will be determined not only by the risk assessment but a range of other factors, including 'technical feasibility (e.g. treatability, detection limits), economic, social, political,' and legislation when determining whether to regulate and, if so, to what extent (US EPA 1995a p. 2).

Consultation with the community to identify their concerns is clearly an important component of both risk assessment and risk management. See Chapter 6.

Scientific judgements and policies must be clearly identified. Inevitable gaps in knowledge will be filled by scientific judgements and policies. These must be clearly identified so that others may understand the role of judgement in interpreting the evidence.

1.5 EVALUATING RISK ASSESSMENT METHODS

Criteria for evaluating risk assessment methods (Covello & Merkhofer 1993) include:

- the logical soundness of the method (e.g. its justification based on theoretical arguments or scientific knowledge, and the validity of the underlying methodological assumptions)
- completeness (e.g. whether it can address all aspects of the problem and the degree to which it excludes issues because they are hard to accommodate)
- precision and accuracy (e.g. reflected in the confidence level associated with the results or the biases resulting from undue weight being given to specific interests or considerations and the sensitivity of results to untested or untestable assumptions)
- acceptability (e.g. compatibility with existing processes; whether it is viewed as rational and fair; the level of understanding for all parties affected by it; and the confidence and familiarity of those who will use it)
- practicality (e.g. the level of expertise, time and input data required)
- effectiveness (e.g. usefulness of results; range of applicability across different risks and problem areas; the generalisability of the conclusion to other problem areas; and effectiveness and efficiency of linkage with other types of methods).

1.6 RISK ASSESSMENT MODELS AND FORMATS

A variety of models are used for risk assessment in Australia by government agencies and consultants. Many of these models are based on paradigms for risk assessment first outlined by the US National Academy of Sciences in 1983 in the seminal work *Risk assessment in the federal government: managing the process* (NRC 1983). This document laid the foundation for contemporary risk assessment processes (including quantitative methodologies) and established different approaches to assessing carcinogenic risks versus non-carcinogenic health effects.

Through the 1980s, 1990s and 2000s these risk assessment paradigms were formalised (and continue to be updated) in a number of US guidance documents, which include:

- *Risk assessment and management: framework for decision making* (US EPA 1984)
- *Science and decisions: advancing risk assessment* (NRC 2008)
- *Risk assessment guidance for superfund* ('RAGS' documents), published successively from 1989, and which continue to be updated today
- *Guidelines for carcinogen risk assessment* (US EPA 2005a)
- *Exposure factors handbook* (US EPA 1997a), with an updated version released for public comment in October 2009.

Further details of the US approach to EHRA are outlined in Chapter 18.

Recent summaries of approaches to EHRA in the Australian context are:

- guidance on health risk assessment of contaminated sites in schedule B(4) of the National Environmental Protection Measure (NEPM) – currently under review and being updated (National Environment Protection Council – NEPC 2010)
- guidance on ambient air quality standards-setting (National Health and Medical Research Council – NHMRC 2006) – which includes a review of health-based approaches to hazard assessment of air pollutants. This led to a comprehensive *Methodology for setting air quality standards in Australia* (NEPC 2011), which includes detailed information regarding risk assessment in the specific context of air pollutants.

An extensive discussion of the different framework models for human health risk assessment in use in Canada and the US can be found in Jardine et al. (2005). This review compares the models, emphasising the ways in which the basic framework can be changed to accommodate different regulatory settings, and the extent to which different models emphasise the importance of consultation with stakeholders and socioeconomic analyses.

1.7 THE FIVE STAGES OF ENVIRONMENTAL HEALTH RISK ASSESSMENT (EHRA)

The historical development of formalised EHRA has resulted in the process being categorised into five distinct stages:

1. Issue identification
2. Hazard identification
3. Dose–response assessment
4. Exposure assessment for the relevant population
5. Risk characterisation.

Some of the key factors and questions that must be taken into consideration at each of these stages include the following:

1. Issue identification

- What are the true drivers for the issue being assessed? (e.g. there is no point in doing a quantitative cancer risk assessment if the real concern is cognitive impairment of children, and if the latter cannot be addressed by risk assessment, then another approach may be necessary).
- Are intervention strategies available to manage the outcomes of the EHRA (e.g. containment of contaminated soil, chlorination of water, pasteurisation of food)?
- Have transport mechanisms been adequately considered (e.g. meteorological factors affecting air pollution, vectors for communicable diseases)?
- Are there factors that could affect persistence (e.g. photolysis and volatilisation of chemicals, desiccation of micro-organisms)?
- Has the risk assessment been initiated as the result of a breakdown of public health measures (e.g. flooding affecting waste control and potable water treatment)?

2. Hazard assessment

- Have the severity and reversibility of health effects been considered?
- Is there any interaction between the identified hazards and other agents in the environment?
- Is the onset of health effects immediate or delayed? While health-based guidelines generally assume long-term continuous exposure, and are usually based on chronic (preferably lifetime)

dosing in animal studies, there may be circumstances where data from an acute or short-term toxicity test may be more appropriate to use in the risk assessment (e.g. adverse effects associated with irritancy).

- Is there is a critical window of exposure? This is often associated with chemicals that modify foetal development, either during gestation or in the early postnatal period when critical neural or organ system developmental processes are occurring. It is also likely that epigenetic and hormonal disturbance mechanisms act mainly during critical exposure windows.
- Has the carcinogenic and/or genotoxic potential of the identified hazards been addressed?

3. Dose–response

- Is appropriate dose–response data available, and has the data been appropriately scaled in translation from animal to human?
- Has the potency of the agent been determined for both acute and chronic dosing?
- Does a threshold or non-threshold model best describe the data?

4. Exposure

- What is the duration, timing, frequency and consistency of exposure?
- Are exposures continuous, intermittent or episodic, or do they show clear patterns?
- Are there are relevant past, current or future exposure patterns to consider?
- Have all exposure routes (ingestion, inhalation, dermal) have been considered?
- Are exposures intergenerational or cumulative, or should they be aggregated?

5. Risk characterisation

- Has genetic variability in the exposed population (or in the source toxicological data) been adequately accounted for?
- Are there individual host characteristics (e.g. age, gender, body weight, pre-existing poor health, immune status, nutritional status, previous exposures or reproductive status) that need to be considered?
- Are there population characteristics (e.g. herd immunity and social behaviours for communicable diseases, social mobility for exposure to air and soil contaminants, recreational patterns for exposure to contaminated recreational waters) that need to be considered?
- Has the risk estimate been expressed quantitatively or qualitatively and, if quantitative, is it a finite risk estimate based on extrapolation of the dose–response relationship, or is it an acceptable daily intake (ADI) or tolerable daily intake (TDI), based on application of safety/uncertainty/modifying factors to a no observed adverse effect level (NOAEL), lowest observed adverse effect level (LOAEL) or benchmark dose (BMD)?

All of these issues will be addressed in more depth in Chapters 2 to 16. The terminology ADI, TDI, NOAEL, LOAEL and BMD is explained in more depth in the Glossary and their derivation in Section 5.6.

One of the concerns about some stakeholder perceptions of current EHRA methodologies is that an impression may be given that the derived risk assessment number, whether based on extrapolation or an ADI/TDI approach, can be taken as a 'bright line between possible harm and safety' (NRC 2008 p. 8) or, in other words, the separation between safe and unsafe exposures. While it is important to dispel this myth in the risk communication process by explaining its inaccuracy because of variability and uncertainty, one should not lose sight of the fact that an exceedence of a standard or guideline or other indicator of 'safety' by a derived risk assessment number should always trigger further consideration of the situation being assessed. Such consideration could include refinement of the assumptions, modelling or input values, and the magnitude of safety factors.

Part of the reason behind this false perception is a lack of understanding of what the numbers generated in an EHRA really mean. This has possibly been compounded by the dichotomy that has developed in the approach to cancer and non-cancer endpoints. The recent review of EHRA methodology by the US National Research Council (NRC 2008) proposes harmonising approaches to these two types of endpoints, including the use of BMD methodology to derive a point of departure (POD) for risk estimation, and assigning a finite risk estimate to both a cancer risk estimate and the calculation of a derived reference dose (RfD) (see Section 3.9). The danger in this approach is that it introduces additional challenges in communicating the meaning of small probability risks, where previously, assessments based on a threshold approach were explained without recourse to citing finite risk numbers. The NRC recommendations have also been criticised (Goldstein 2010) on the basis that EHRA methodology has worked well and is not in need of any such 'improvements'.

At the present time, where EHRA is practised in Australia, assigning an ADI or TDI carries no such implication that it is associated with any finite level of risk. The operating definition is that an ADI or TDI represents 'an estimate of the intake of a chemical which, during a lifetime of exposure, appears to be without appreciable risk, on the basis of all facts known at the time'.

1.8 RISK ASSESSMENT FRAMEWORKS

The framework model that encompasses the five stages of EHRA and their interlinkage with stages of risk management and stakeholder consultation was first proposed for use in Australia in the 2002 enHealth document, which is revised in this document. Conceptually, the five stages are closely linked and dependent on the preceding stages.

The original model is illustrated in Figure 1. The terminology is similar to terminologies used by other major models.

Note that stakeholder consultation is considered essential at all stages of the EHRA process, and that a review/reality check should be built into the critical stages of hazard and exposure assessment, and risk characterisation, to ensure the outcomes have not been distorted by inappropriate choice of data inputs.

Various revisions to this basic model of EHRA are set out in the NRC (2008) update of approaches to risk assessment. This NRC document outlines a more holistic framework (see Figure 2), which emphasises the importance of problem formulation and planning as precursors to the formal steps of quantitative risk assessment. It further reinforces the view that the outcomes of the EHRA process are critical in better informing the risk management stages, and providing for stakeholder consultation and review to occur at all stages.

While the model described in Figure 1 has served Australia well over the past decade, it is recommended that the more holistic model (Figure 2) provides a more structured and informative framework for EHRA in Australia from now on.

Figure 1: Environmental health risk assessment model

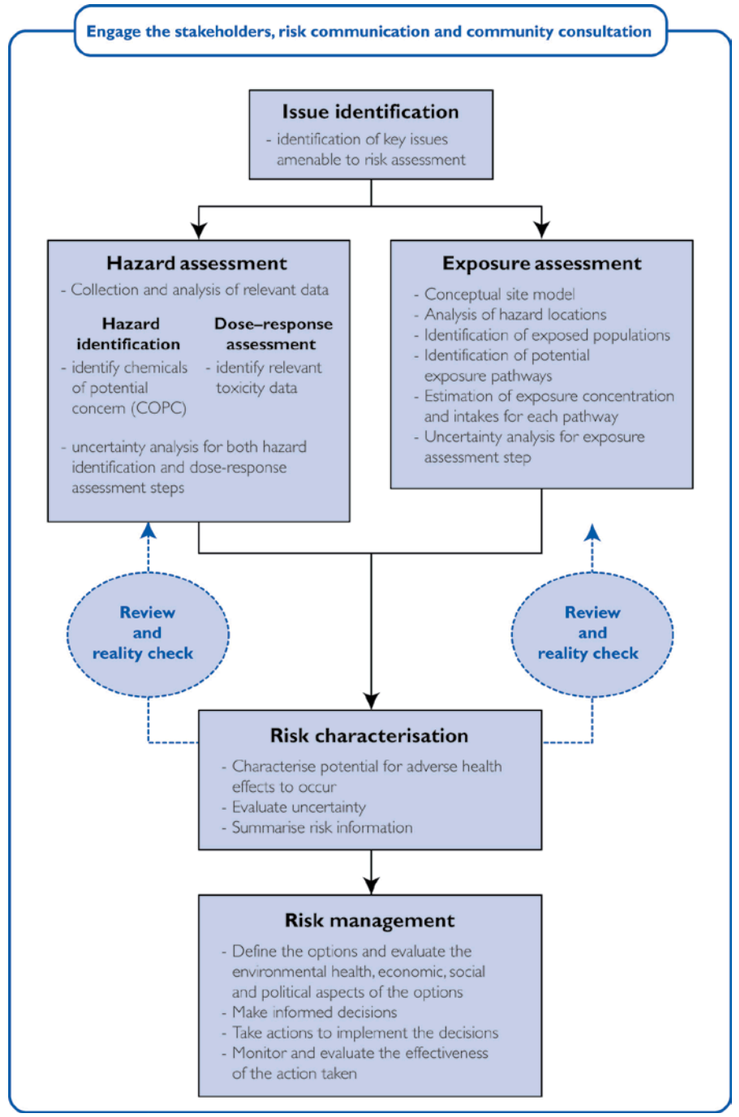
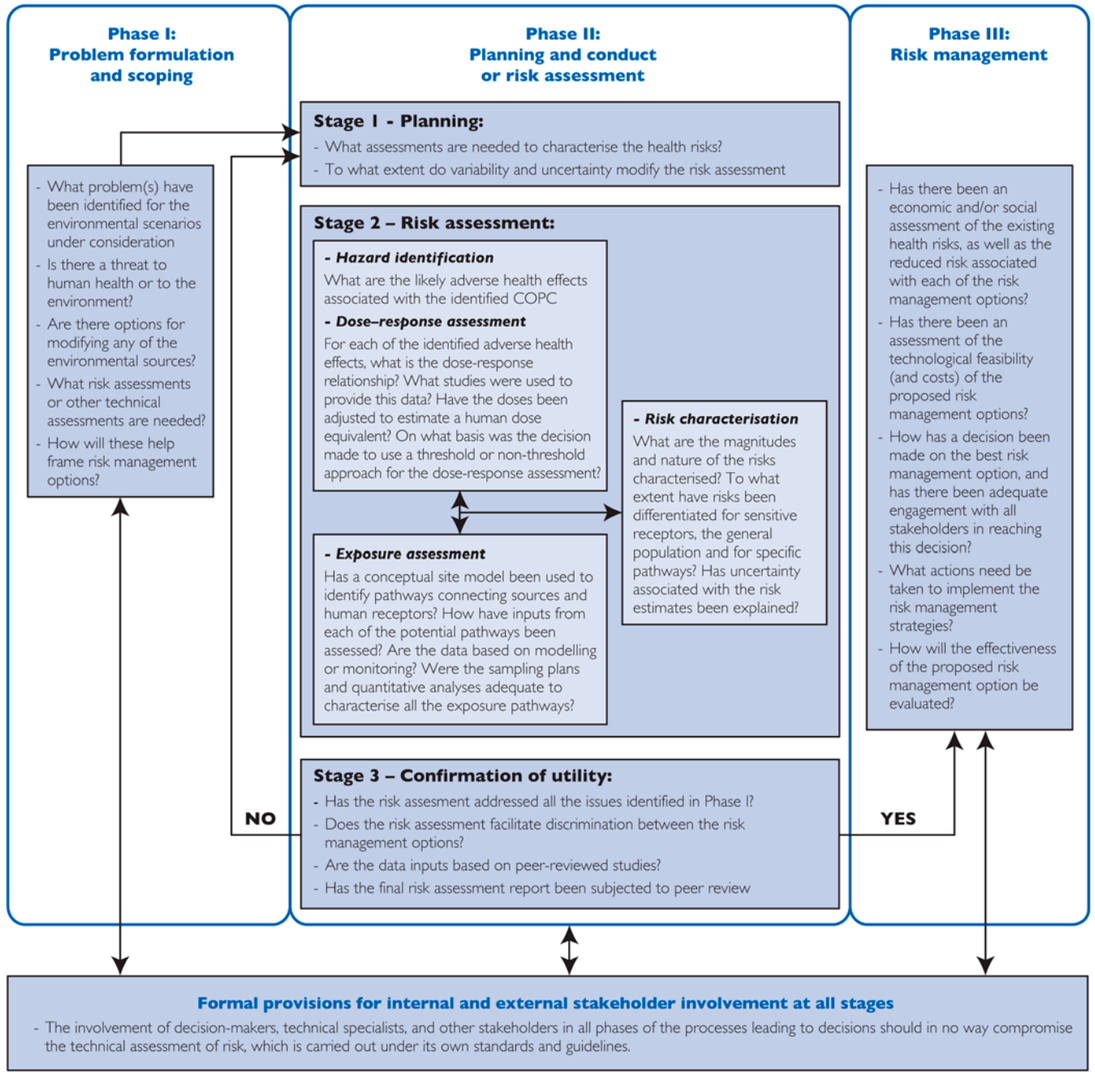


Figure 2: A revised outline of the interlinked processes of EHRA



Adapted from: NRC 2008.

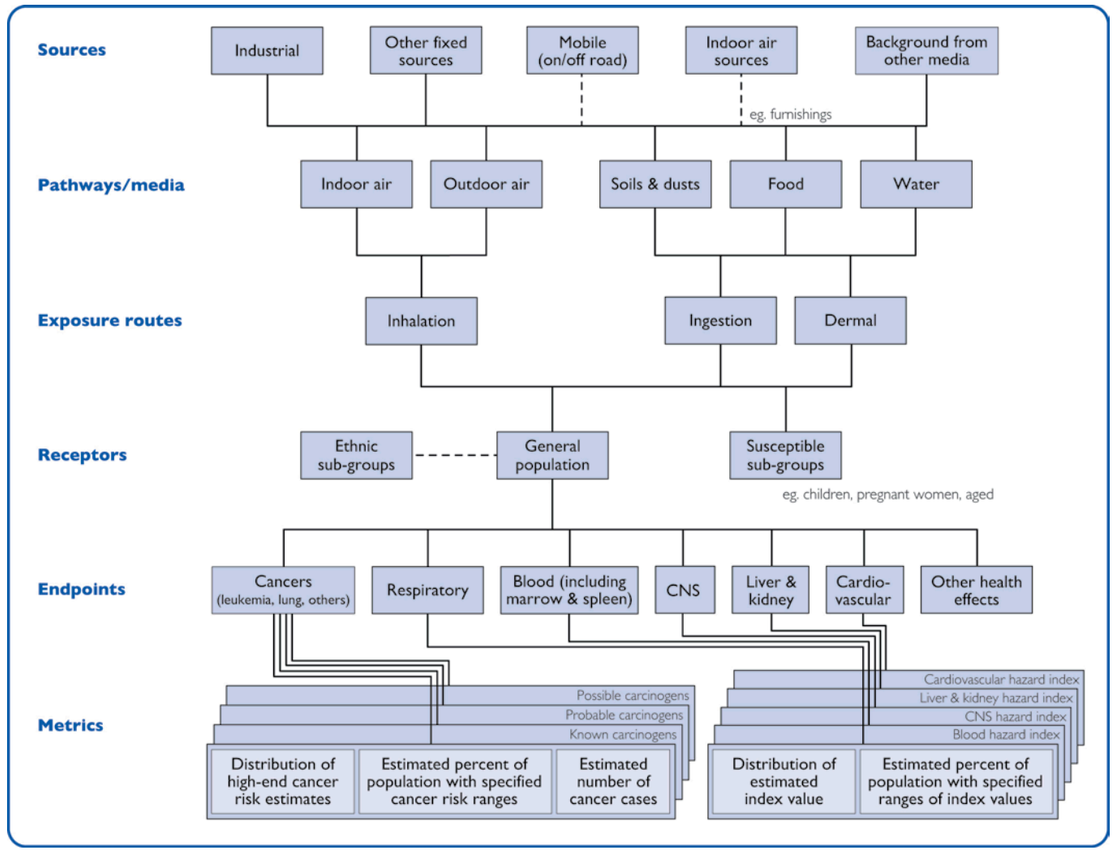
The relationship between the expanded framework and the framework first described in 2002 is that Phase I and the planning stage of Phase II are aligned with the former 'issue identification' stage, the Phase II 'Planning and conduct of risk assessment' elements (hazard identification, dose–response assessment and risk characterisation) are aligned with comparably named elements of the former framework, while the Phase III 'Risk management' elements are an expanded version of the fifth stage of the 2002 framework. All phases of the new model have more description of the key issues that need to be addressed at each stage, although it is likely that in some risk assessments, answers to some questions will be obvious, while others may need a full and detailed approach.

As in the former framework, stakeholder engagement at all phases is emphasised as a critical element.

Various elements of the framework can be expanded to illustrate the critical individual components.

For example, Figure 3 illustrates the multiplicity of exposure sources, exposure pathways, receptors, endpoints and measurements (metrics), which could be included in conducting an EHRA.

Figure 3: Expanded illustration of the major exposure pathways, potentially exposed groups leading to potential health outcomes



Solid lines indicate pathways usually considered. Other pathways shown may not be considered in conventional EHRAs.
Adapted from: NRC 2008.

1.9 TIERED APPROACHES TO EHRA

Because of the cost and complexity of contemporary formal EHRAs, circumstances may suggest a tiered approach to formulating a site- or issue-specific EHRAs. The simplest approach (Tier 1) would be an initial screening-type evaluation of risks using conservative default exposure parameter estimates and comparison with published health-based guidelines. Tier 2 and Tier 3 processes would involve collecting additional exposure data and a more detailed analysis of dose–response data, possibly including calculation of target tissue doses or translating animal doses into human-equivalent dose estimates.

The tiered approach in risk assessment is common in many jurisdictions, although the number of tiers and their precise usage may differ. For example, Health Canada uses the term Preliminary Quantitative Risk Assessment (PQRA) to refer to what would otherwise be called Tier 1, while the terms Tier 2 and 3 are allocated to site-specific risk assessment (SSRA).

Figure 4 is a schematic depiction of some of the elements that might comprise Tier 1 to Tier 3 EHRAs.

Tiered risk assessment is particularly relevant to the EHRA of contaminated sites, and is discussed in more depth in schedules B(4) and B(7) of the contaminated sites NEPM (NEPC 2010). The tiered approach allows the problem under consideration to be assessed at an appropriate level of complexity. The degree of health protection achieved is equal at each tier. As the amount of data and assessment detail increases and the conceptual understanding of site conditions (i.e. the conceptual site model) is refined, the level of uncertainty decreases. In turn, the amount of caution which must be substituted for knowledge in the risk assessment process may be reduced (NEPC 2010).

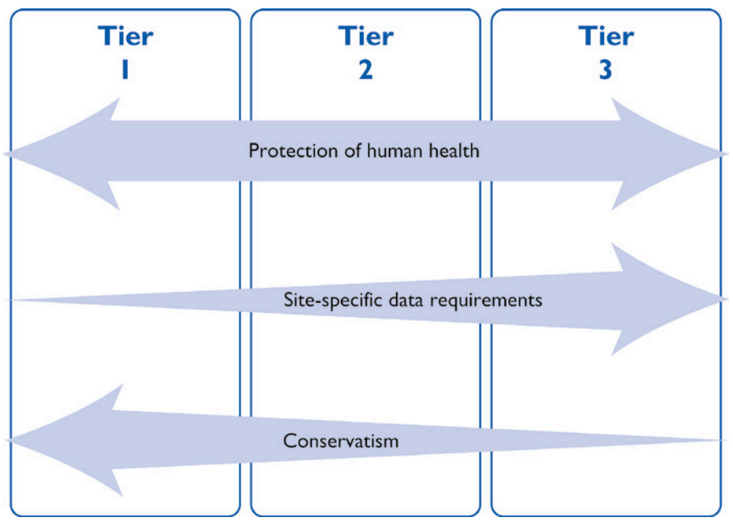
A risk assessment progresses from Tier 1 to Tier 2 when the less-refined risk estimates at Tier 1 may be unacceptable, and further assessment is needed. Progression from Tier 2 to Tier 3 is driven by potentially unacceptable risks at Tier 2. Tier 3 provides more detailed and specific focus on risk-driving factors.

In Australia, there is often no clear break between the different tiers. The investigations and risk assessment proceed until the level of information is appropriate for the decision making required. It is common for most risk assessments, regardless of which tier, to have a screening step and a detailed assessment step.

In the screening step, usually the maximum concentration of each chemical in the full list of chemicals or other agents that might pose a risk at the site are compared to relevant national or international guidelines. This is conservative as the maximum concentration is presumed to be present at all times in all situations for this step. If a chemical or other agent is found to exceed the guideline value then that chemical is classified as a chemical of potential concern and a detailed assessment should be triggered.

In the detailed assessment step, the chemicals are assessed more fully, and this may include exposure scenario modelling, fate and transport modelling or further investigations to better understand the situation under investigation and to refine the assumptions to make them more realistic. Such assessments usually include consideration of the maximum case and an average case.

Figure 4: Elements of a tiered approach to EHRA



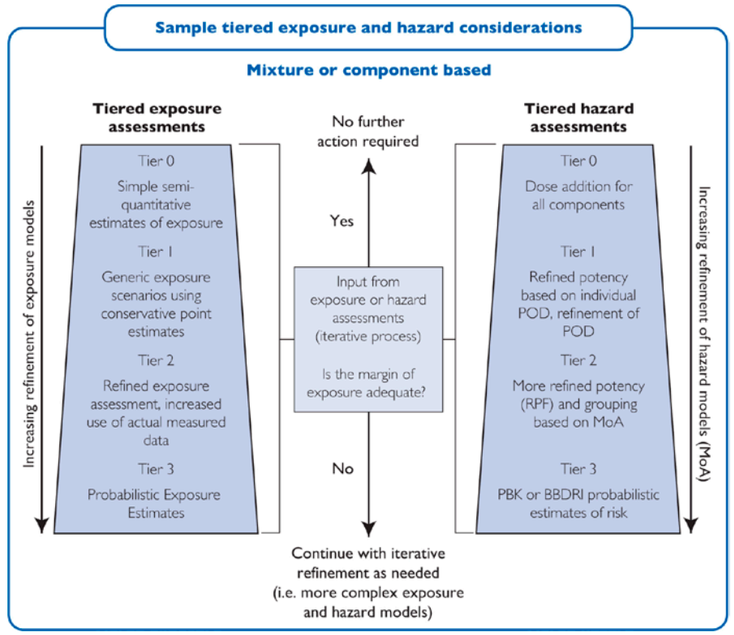
Adapted from: NEPC 2010.

There is a danger that the tiered structure common to some regulatory risk assessments can lead to a negative impression that the only possible outcome contemplated by going to a higher tier is a relaxation of remediation requirements rather than an unbiased approach allowing for more evidence leading to more onerous remediation requirements. This perception needs to be offset by clearly explaining that the only purpose of going to a higher tier is to reduce uncertainty by including more realistic estimates of exposure. Going to a higher tier may also aid risk management by better targeting the risk management options.

Conservative exposure settings and assumptions (as in Tier 1 assessments) may not be realistic for the site under consideration as they are based on generic assumptions and parameters that are not likely to be realistic. A Tier 2 assessment may be used to produce more appropriate values by amending the assumptions to reflect actual site conditions. Where available, data on biodegradation of contaminants and bioavailability of chemicals should be considered (see Section 4.2.1). Exposure factors (and assumptions) should reflect the scenarios under consideration.

The tiered approach is expanded and outlined in more detail in the International Programme on Chemical Safety (IPCS) framework for risk assessment of combined exposures to multiple chemicals (IPCS 2009b). A flow chart describing the IPCS framework (Figure 5) emphasises integration of information on mode of action (MoA) and gradual refinement of exposure assessment throughout the tiers. It applies a separate tiered approach to both hazard and exposure assessment. It also notes that the hazard assessment component could be based on individual components or incorporate dose-additive approaches for mixtures (see Chapter 12 for discussion of toxicological assessment of chemical mixtures).

Figure 5: Conceptual representation of the proposed IPCS risk assessment framework



Reproduced from IPCS 2009b with permission from WHO.

Note that the IPCS framework includes an additional tier (Tier 0) that is not included in the Australian guidance. In the IPCS context, Tier 0 would encompass initial crude exposure estimates that are less well defined than those used in an Australian Tier 1 assessment.

1.10 DETERMINISTIC VERSUS PROBABILISTIC ESTIMATES IN EHRA

A deterministic approach means that input values in an exposure model are expressed as single values or point estimates. These are intended to be 'best estimates' of the value of the input variables. The advantage of this approach is that it is simple, easily understood, and therefore widely applied in EHRA practice. However, an inadvertent consequence is that many of the point estimates used are chosen at the upper end of their likely ranges. This can lead to compounding of the conservatism in a model, and consequently in the EHRA outcome. Sensitivity and uncertainty analysis are used to overcome this disadvantage, and provide increased understanding and clarity on which values are risk-driving; this is itself a useful part of the risk assessment (NEPC 2010).

Probabilistic techniques can overcome the potential for compounding conservatism, and may provide for a better descriptor for the uncertainty associated with the various input parameters, and also provide estimates of the statistical limits of underlying parameter distributions.

This can be useful for the risk manager to decide on the extent to which 'outliers' may influence the EHRA process, and provide a basis for deciding the limits applicable to protecting the extremes of the population distributions to which the EHRA applies.

Probabilistic risk assessment methodologies have been reviewed by Bogen et al. (2009). These methods can be used to assess and manage uncertainty, inter-individual heterogeneity and other sources of variability.

Monte Carlo analysis is one probabilistic tool that has been promoted for use in EHRA because it replaces deterministic estimates of individual parameter inputs with probability distribution functions describing the variability of those input parameters. The probabilistic exposure model can be run through thousands of iterations, with values for each parameter selected randomly on the basis of their occurrence frequency. The ultimate output is a probability distribution function, which describes the calculated parameter (usually an estimate of exposure). Further discussion of the strengths and weaknesses of the Monte Carlo approach is in Section 13.2.

1.11 ENVIRONMENTAL HEALTH RISK ASSESSMENT AND HEALTH IMPACT ASSESSMENT

Although they are related processes, health impact assessment (HIA) and EHRA address different issues. HIA is defined by different agencies in different ways. The consensus definition is that of the 1999 *Gothenburg consensus paper* by the WHO Regional Office for Europe, as described in the enHealth *Health impact assessment guidelines* (enHealth 2001).

> *... a combination of procedures or methods by which a policy, program or project may be judged as to the effects it may have on the health of a population.*

In other words, HIA is a systematic process to assess the actual or potential, and direct or indirect, effects on the health of individuals, groups or communities arising from environmental conditions or hazards arising from policies, objectives, programs, plans or activities. It looks at both potential health benefits and health impacts from an activity or situation. It is usually a process undertaken as part of an environmental impact assessment for a significant project and looks at both positive and negative impacts on health. HIA is generally undertaken in the early stages of project planning in order to predict and facilitate avoidance of potentially negative health impacts, to promote more positive health impacts and to promote sustainable development. It takes into consideration the social and socioeconomic factors.

The definition of 'health' is taken to be 'a complete state of physical, mental and social wellbeing and not merely the absence of disease or infirmity' (WHO Constitution). This definition has not been altered since it was promulgated in 1948. In this context, EHRA is simply a tool for appraising health risks (i.e. adverse health impacts) from contaminant exposures in the broader process of health impact assessment.

1.12 ENVIRONMENTAL HEALTH RISK ASSESSMENT AND THE 'PRECAUTIONARY PRINCIPLE'

The precautionary principle (PP) was first formalised in 1984 at the First International Conference on the Protection of the North Sea and has since been integrated into several international conventions and agreements. The UN Conference on Environment and Development (UNCED – 1992 Rio declaration from Agenda 21) interpreted a precautionary approach to chemicals management in the following statement:

> *'In order to protect the environment, the precautionary approach shall be widely applied by the States according to their capabilities. Where there are threats to of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation.'*

Some key points relating to the precautionary principle from a 2002 paper from the UK Inter-Departmental Liaison Group on Risk Assessment (ILGRA 2002) were:

- 'The purpose of the precautionary principle is to create an impetus to take a decision, notwithstanding scientific uncertainty about the nature and extent of risk.
- The precautionary principle should be invoked when there is good reason to believe that harmful effects may occur to human, animal or plant health or to the environment.
- Action in response to the precautionary principle should be in accord with the principles of good regulation, i.e. be proportionate, consistent, targeted, transparent and accountable.
- Applying the precautionary principle is essentially a matter of making assumptions about consequences and likelihoods, and then using standard procedures of risk assessment and management to inform decisions on how to address the hazard or threat.
- Invoking the precautionary principle shifts the burden of proof in demonstrating presence of risk or degree of safety towards the hazard creator. The presumption is that the hazard creator should provide, as a minimum, the information needed for decision making.
- Decisions reached by invoking and applying the precautionary principle should be actively reviewed, and revisited when further information that reduces uncertainty becomes available.'

In Australia, the 1992 Inter Governmental Agreement on the Environment (IGAE) <http://www.environment.gov.au/about/esd/publications/igae/index.html> defines the precautionary principle as follows:

> 'Where there are threats of serious or irreversible damage, lack of scientific certainty should not be used as a reason for postponing measures to prevent environmental degradation. In the application of the precautionary principle, public and private decisions should be guided by:
>
> i. careful evaluation to avoid, where practicable, serious or irreversible damage to the environment; and
>
> ii. an assessment of the risk-weighted consequences of various options.'

There are clearly some common elements to these various espousals of the precautionary principle, which are also relevant to an alternative term that is frequently used, 'precautionary approach'. These common elements include:

- It is aimed at preventing serious or irreversible damage to the environment.
- It provides for taking appropriate actions in the face of uncertainty or lack of complete knowledge of the potential risks.
- There needs to be consideration of the economic and practical feasibility of implementing the risk management measures that may be suggested by invoking the precautionary principle.
- Application of the precautionary principle is closely intertwined with social equity issues. Protection of more vulnerable, socioeconomically disadvantaged or ethnic segments of the community who may be at greater risk from environmental hazards is an important element of risk management.

In EHRA, consideration of the precautionary principle is particularly relevant during the risk management stage. Risk assessment provides a process for applying the precautionary principle by providing information about the nature and magnitude of the threats of serious or irreversible environmental damage' associated with various risk management options.

There is a detailed description of the precautionary principle and its application to setting air quality standards in a recent consultation paper associated with the review of the ambient air quality NEPM (NEPC 2009).

However, it is quite difficult to find any specific reference to the 'precautionary principle' in chemicals management legislation or regulation in Australia. A search of the Chemicals Gateway, an Australian Government website with extensive references to chemicals risk management gives a 'no match' when searched for the 'precautionary principle' and there is no formal reference to it in any other chemicals risk management site. However, the principle is embedded in various public health and environmental legislation around Australia and there are abundant references to the fact that Australian agencies use a precautionary approach to managing uncertainty in the risk assessment and management of hazardous chemicals. Advocacy to adopt a precautionary approach to EHRA is reflected throughout this enHealth document.

Chapter 2: Problem formulation and scope

The purpose of this stage of EHRA is to formulate the problems to be considered by the risk assessment and to clarify the proposed scope. It corresponds with Phase I of the framework outlined in Figure 2 and with 'issue identification', the first stage of the original risk assessment framework depicted in Figure 1.

Essentially, this means addressing the following points:

- What is the concern?
- Why is it a concern?
- How urgent is the concern?
- How do stakeholders perceive the concern?

This will include identifying and describing:

- issues associated with existing environmental conditions
- susceptible and/or vulnerable populations likely to be exposed
- potential exposure pathways
- potential management options that may mitigate exposure
- the risk and other technical assessments necessary to evaluate risk and discriminate between potential risk management options.

2.1 IDENTIFYING AND DESCRIBING ISSUES WITHIN EXISTING ENVIRONMENTAL CONDITIONS

'Hazards' need to be distinguished from 'issues'. The determination of the issues is necessary to establish a context for the risk assessment, and assists the process of risk management. Issues have dimensions related to perceptions, science, economics and social factors.

Examples of issues are:

- community concerns over emissions from a smelter or other industrial facility
- community outrage over the proposed development of a communications tower
- how contaminated sites are managed
- development of new standards for water quality, including use of a new water treatment chemical or new uses for recycled water
- changes to a food standard that permit higher levels of exposure or introduce new chemicals into the food chain.

'Hazards' relate to the capacity of a specific agent to produce a particular type of adverse health or environmental effect. The environmental agents of concern may include physical, chemical, biological or social factors.

- Physical factors include heat, cold, noise, mechanical hazards, solar radiation, ionising radiation (e.g. X-rays) and non-ionising radiation (e.g. microwaves), noise and vibration.
- Chemical factors include synthetic and naturally occurring substances.
- Biological factors include viruses, prions, bacteria, parasites and vermin.
- Social factors include poverty, unemployment cultural values and effects on access or amenity.

Examples of hazards include the capacity of:

- benzene to cause leukaemia
- solar radiation to cause skin cancer
- salmonella to cause vomiting and diarrhoea.

Hazardous agents may be identified from the range of data sources, including:

- environmental monitoring (e.g. of food, air, water and soil)
- emissions inventories (e.g. the National Pollutant Inventory)
- biological monitoring (e.g. of children's blood lead levels or Ross River virus antibody levels)
- disease surveillance (e.g. of salmonella types for food poisoning, skin cancer rates, pregnancy outcomes)
- health monitoring (e.g. of lung function testing to detect the onset of environmentally caused asthma)
- epidemiological studies (e.g. of particular disease rates in certain populations such as workers) to identify previously unknown hazards
- information about analogous hazards.

2.1.1 Phases of issue identification

Issue identification comprises several phases:

1. Identification of environmental health issues (or an individual issue) and determining whether there are hazards amenable to risk assessment – this will involve demarcating 'hazards' from 'issues' and may require environmental sampling.
2. Putting the hazards into their environmental health context (clarification and prioritising of problems and hazards).
3. Identifying all the chemicals of potential concern (COPC) (i.e. prioritising those chemicals that need to be fully considered in a quantitative risk assessment)
4. identification of potential interactions between agents.

At this stage it often becomes apparent that the setting for the risk assessment is a situation where:

- there are multiple, interacting hazards rather than an isolated hazard – perhaps the contaminant affects multiple environmental media (e.g. lead smelter emissions contaminating soil, air, water and food)
- the hazard may have single or multiple sources (e.g. atrazine contamination of a drinking-water supply from a chemical spill versus particulates in ambient air arising from diesel engines, wood stoves and environmental tobacco smoke)
- there are concerns about a range of potential health effects from various hazards
- there is variable and often superficial information on exposure and the level of health problem
- there is a context of public anxiety, anger and impatience
- different stakeholders may have different perceptions of the issues for example, a stakeholder group comprised of workers at a smelter who are also nearby residents may have complex perceptions
- the hazards may be compared with other environmental hazards affecting the community; this component of the appraisal will be affected by objective data (e.g. of different disease rates) and subjective perceptions by the stakeholders (Presidential/ Congressional Commission on Risk Assessment and Risk Management – P/CCRARM 1997).

In relation to assessment of multiple exposure routes and sources, US regulations define two types of exposure that need to be considered in a risk assessment:

- **aggregate exposure:** the analysis of exposure to a chemical by multiple pathways and routes of exposure
- **cumulative exposure:** the combined risk estimate where exposure occurs simultaneously or consecutively to multiple chemicals that exert toxicity through a common mechanism.

In the case of 'aggregate' exposure, the requirement is no more than would be normally done in a conventional EHRA, where all potential exposure pathways should be considered in the risk assessment. The methodologies for 'aggregate' risk assessment are set out in Chapter 4.

In the case of 'cumulative' exposure, the methodologies are more complicated, particularly since cumulative exposure risk assessments are intended to address the interactions of multiple agents or stressors, not all of which are necessarily chemical agents. It could include biological or physical agents that could modify the toxicity of the environmental chemicals under consideration.

Further guidance on assessing multiple chemical exposures (i.e. mixtures of chemicals) is outlined in Chapter 12.

2.1.2 Identifying chemicals of potential concern

It is quite likely that the issue and/ or hazard identification stages will find a large number of chemicals whose presence in the environment and toxicity may give rise to adverse health outcomes. The issue then becomes which of them must be designated as chemicals of potential concern (COPC) (which must be addressed in the formal EHRA) or whether any of them can be readily eliminated from further consideration.

The tiered risk assessment screening process (see Section 1.9) may enable COPC to be discriminated from among a much larger number of environmental contaminants, and may point to the need for a more advanced (Tier 2 or 3) risk assessment.

It may also be possible to establish that chemical concentrations in the environment are so low that exposures are unlikely to exceed a generic threshold of toxicological concern (TTC) and can therefore be discounted. The application of a TTC approach to risk assessment is discussed in Section 5.13.

2.2 IDENTIFYING AND DESCRIBING SUSCEPTIBLE POPULATIONS

Within the general population there may be sub-groups potentially more susceptible to the effects of environmental chemicals than others in the population. Human variability (also called intra-species, i.e. inter-individual, variability) may arise through toxicokinetic or toxicodynamic variability (Dybing & Soderlund 1999). Both these areas of variability may be due to acquired and/or inherent factors that may make a person more susceptible to environmental pollution.

Sensitivity of individuals is also likely to be affected by age, sex, nutritional and pregnancy status, and combinations of these (IEH 1999c). It is therefore imperative that the issue identification stage considers whether any of these factors could influence the outcome of the EHRA.

A distinction needs to be drawn between the use of the terms 'susceptibility', 'vulnerability' and 'sensitive groups' within a population of 'receptors' under consideration in an EHRA.

Susceptibility: refers to intrinsic biological factors that can increase the health risk of an individual at a given exposure level. Examples of susceptibility factors include: genetic factors, late age and early life, and prior or existing disease.

Vulnerability: refers to human populations at higher risk due to environmental factors. Examples of vulnerability factors include poverty, malnutrition, poor sanitation, climate change and stress associated with mental health diseases.

Sensitive groups: refers to populations with both susceptibility and vulnerability factors.

2.2.1
Epidemiological principles

A sensitive sub-population is one where an adverse response to an environmental pollutant occurs at concentrations substantially lower than that affecting the majority of the population. Another sub-population that may be considered to be 'sensitive' is one where the consequences of exposure are more significant than in the majority of the population. For example, children may be considered a sensitive population because any irreversible adverse effects may influence their health throughout their life. The elderly, especially those with specific comorbid effects such as cardiac or respiratory failure, may also constitute a sensitive group because the secondary consequences (e.g. pneumonia, worsening cardiac failure) may be more serious than in the remainder of the population (NHMRC 2006).

The identification of sensitive sub-populations may be guided by:

- clinical history (e.g. the presence of diseases such as asthma, cardiac failure, chronic bronchitis or cystic fibrosis, which may exacerbate sensitivity to environmental pollutants)
- clinical evidence of hyper-responsiveness (e.g. using methacholine or more specific challenge tests to assess susceptibility to irritant air pollutants)
- demographic factors (e.g. the elderly or very young)
- genetic factors (e.g. cystic fibrosis).

From a physiological standpoint, any person who has decreased functional reserve in an organ system is theoretically less able to cope with additional environmental stressors, be they 'non-chemical' or 'chemical' in nature. As with other areas of medicine and toxicology, whether a particular individual will respond adversely to a certain environmental stressor depends upon the relative balance between the extent of physiological compromise (in some cases this is proportional to the degree of disease severity) and extent of exposure. In many situations, acquired susceptibility (e.g. illness or old age) shifts a person towards the 'sensitive' tail of the population dose–response curve for the pollutant. These individuals nonetheless remain part of the continuum of the overall population dose–response and experience effects similar to others but at lower exposures to pollutants, or more intense effects at equivalent exposures (NHMRC 2006).

Genetic variability can make an important contribution to human variability, such as in the form of polymorphic genes for metabolism or tissue repair from toxic insult. Although it has long been recognised that genetic polymorphism plays an important role in driving the variability in xenobiotic metabolism, and genetic polymorphisms have been used as biomarkers of potential effect (Scherer 2005), this awareness has typically not translated into quantitative use of the data in risk assessment or standard-setting (Haber et al. 2002; US EPA 2002a). This is likely due to data gaps in our understanding such as:

- prevalence of polymorphism
- lack of a defined link between genetic polymorphism and an adverse effect
- extent of induction/inhibition through co-exposure with other substances, lifestyle or diet
- relative contribution of multiple enzyme systems
- allelic frequencies for major ethnic groups
- large numbers of low-frequency alleles
- absence of chemical-specific phenotype data.

Guideline values used in risk assessments should normally have been developed in such a way that most sensitive sub-populations are protected. Some older guideline values do not specifically address this issue. Risk assessors should check that sensitive sub-populations are covered in the guidelines being used. Particularly, each assessment should consider whether early childhood exposure to carcinogens is relevant for the site or activity being investigated and, if so, whether it is covered in the guideline proposed for use.

2.2.2
Risk assessment and children

Children may differ from adults in a range of behavioural and physiological parameters that may need to be taken into account in the risk characterisation phase of risk assessments.

The principal factors causing these potential differences are:

- growth, development and maturational rates
- children's greater potential future durations of life, which is relevant to the potential for accumulation or exceeding latency periods
- dietary differences – children can eat much greater quantities of particular foods (particularly dairy products, soft drinks and some fruit and vegetables) than adults on a body weight basis (Rees 1999)
- placental transfer of contaminants and accumulation in breast milk can result in exposures which are unique to the prenatal and postnatal states
- behavioural factors, for example, children's play activities may put into more frequent contact with soil contaminants and they are also more likely to indulge in soil eating behaviours (pica)
- available parameters for toxicity assessment, for example, techniques for assessing dizziness, intelligence and hearing impairment are different between children and adults
- biochemical and physiological responses, for example, children have a higher metabolic rate, more limited ability to control body temperature, more rapid growth rate and a higher percentage of water in the lean body tissue
- disposition of the agent within the body, for example, transit time, pH and enzyme activity in the gut are different for children as are tissue–chemical bindings
- liver function related to detoxification matures after birth, as does the renal excretion of foreign compounds
- differences in gut microflora
- the immaturity of children's immune systems
- differences in the clearance of chemicals – the higher clearance of certain chemicals from the body in children compensates in part for the greater sensitivity for their developing organ systems (Renwick 1999) but for some other chemicals, clearance may be lower
- exposure factors – the surface area to body mass ratio will change markedly with ageing. In the newborn, the ratio is typically 0.067 (m^2/kg) decreasing to 0.025 in an adult. While the respiratory volume remains fairly constant at 10 ml/kg/breath, the surface area of the alveoli increases from 3 m^2 in an infant to approximately 75 m^2 in an adult and the respiration rate drops from 40 breaths per minute to 15 breaths per minute (Snodgrass 1992).

A general discussion of the issues relating to risk assessment for children may be found in Calder (2008), Roberts (1992) and US EPA (2005b; 2006a), with a more extensive review of the physiological, pharmacokinetic, behavioural, genetic and exposure factors that may alter the sensitivity of children to environmental hazards (Hines et al. 2010).

The potential impact of these differences highlights the need for agent-by-agent appraisal. However, it should be recognised that the derivation of some types of toxicological reference doses (e.g. ADI or TDI) envisage whole-of-life exposure and may therefore be relevant for assessing risks associated with early-life exposure stages.

Special consideration has been advocated by the US EPA for assessing risks associated with early-life exposure to mutagenic carcinogens (see Section 5.8). US legislation mandates the application of an additional 10x safety/uncertainty factor in the derivation of an RfD for pesticides where studies indicate developmental neurotoxicity or other toxic effects that could be associated with early-life susceptibility.

While Australian environmental health authorities have not enunciated specific policies relating to applying these US early-life risk assessment strategies, additional precaution tends to be applied on a case-by-case basis when justified by relevant data. While the US early-life risk assessment policies are not automatically adopted in Australia, they have been incorporated into the development of health investigation levels (HILs) for benzo(a)pyrene (BaP) in the revision of the contaminated sites NEPM (see Section 16.1) and they should be considered where there is good evidence that such an approach is relevant.

2.2.3
Risk assessment and older persons

For the ageing, there is a decrease of functional reserve in the physiological and psychological systems. Distribution of chemical agents is affected by changes in body composition with age: body fat increases and body water decreases with age. The clearance of renally eliminated compounds is reduced because of changes in renal function. Liver function can be reduced in the elderly affecting biotransformation of chemical agents. Increased sensitivity to the central nervous system in the ageing population from many drugs has been reported (Crome 1999). Changes will occur to the immunological system often resulting in reduced immunocompetence.

Ageing populations are very heterogeneous in terms of their general health. For those with impaired health, there may be a variety of conditions present. Cognitive impairment is common in the very old or those with age-related pathology (e.g. Alzheimer's) and affects their abilities to recognise, interpret and react to acute and chronic environmental hazards. They are higher consumers of pharmaceuticals and there is a potential interaction with these pharmaceuticals and other agents.

2.2.4
Risk assessment and gender

Gender differences may need to be taken into consideration when identifying potential exposure pathways in the exposure assessment phase and characterising potential adverse health effects in the risk characterisation phase of the risk assessment process.

There are anthropometric (e.g. height, weight, body surface area) and body composition differences (e.g. fat content, muscle mass) between males and females that may affect exposure concentrations of agents from different pathways. These differences may also influence the absorption, distribution, metabolism and elimination of xenobiotics and have a significant influence on toxicity (Silvaggio & Mattison 1994). Some of the factors that influence these processes are summarised in Tables 1–4.

Men and women differ in many lifestyle and occupational exposure factors (e.g. alcohol drinking and cigarette smoking) dietary patterns and how they spend their time. These factors may influence the exposure and effect of an agent on the individual.

For many chemical toxicants there are important differences between males and females in experimental studies. Calabrese (1985) identified 200 toxicants where toxicological analysis of animal studies suggest there are important differences between males and females in the expression of toxicity.

Table 1: Factors influencing the absorption of chemicals

Parameter	Physiological difference	Effects on toxicokinetics
Gastric juice pH	M < F < pregnant F	Absorption of acids/bases modified by change in pH
Gastric juice flow	M > F > pregnant F	Absorption modified by decreasing flow
Intestinal motility	M > F > pregnant F	Absorption increases with decreasing motility
Gastric emptying	M > F > pregnant F	Absorption and gastric metabolism increase with decreasing gastric emptying
Dermal hydration	Pregnant F > M, F	Altered absorption in pregnant F
Dermal thickness	M > F	Absorption decreases with increasing dermal thickness
Body surface area	M > pregnant F > F	Absorption increases with increasing body surface area
Skin blood flow	Pregnant F > M, F	Absorption increases with increasing skin blood flow
Pulmonary function	M > pregnant F > F	Pulmonary exposure increases with increasing minute volume
Cardiac output	M > pregnant F > F	Absorption increases with increasing cardiac output

Table 2: Factors influencing the distribution of chemicals in the body

Parameter	Physiological difference	Effects on toxicokinetics
Plasma volume	Pregnant F > M > F	Concentration increases with increasing volume
Total body water	M > pregnant F > F	Concentration decreases with increasing body water
Plasma proteins	M, F > pregnant F	Concentration fluctuates with changes in plasma proteins and protein binding
Body fat	Pregnant F > F > M	Body burden of lipid-soluble chemicals increases with increasing body fat
Cardiac output	M > pregnant F > F	Distribution rate increases with increasing cardiac output

Table 3: Factors influencing the rate of metabolism of chemicals

Parameter	Physiological difference	Effects on toxicokinetics
Hepatic metabolism	Higher BMR in M, fluctuating hepatic metabolism in pregnant F	Metabolism generally increases with BMR
Extra-hepatic metabolism	Metabolism by foetus/placenta	Metabolism fluctuates
Plasma proteins	Decreased in pregnant F	Elimination fluctuates with changes in plasma proteins and protein binding

BMR = Basal metabolic rate.

Table 4: Factors influencing the elimination of chemicals from the body

Parameter	Physiological difference	Effects on toxicokinetics
Renal blood flow, GFR	Pregnant F > M > F	Renal elimination increases with increasing GFR
Pulmonary function	M > pregnant F > F	Pulmonary elimination increases with increasing minute volume
Plasma proteins	Decreased in pregnant F	Elimination fluctuates with changes in plasma protein and protein binding

GFR = glomerular filtration rate
(Tables 1–4 adapted from Government/Research Councils Initiative on Risk Assessment and Toxicology, 1999)

There have been reports of differences when comparing men and non-pregnant women in their response to toxic levels of lead, beryllium and benzene. Gender differences have also been reported to occur from exposure to ionising radiation, noise and vibration, and extreme temperature changes (i.e. heat and cold stress) (Hunt 1982).

2.2.5 Risk assessment and reproductive status

The human reproductive system is susceptible to environmental factors that can produce a variety of adverse effects during the production of ova (oocytogenesis) and viable sperm (spermatogenesis) on fertilisation, on implantation within the uterus, and growth and development of the embryo and foetus.

Reproductive status is also influenced by the extent of exposure and adverse effects from occupational and environmental agents. Teratogenesis (abnormal development of the embryo and foetus) is a risk for the foetus that may be exposed to environmental agents. The principal factors that determine an agent's risk of teratogenicity and which need to be considered in risk assessment include (Goldfrank et al. 1990):

- the nature of the agent
- access of the agent to the foetus
- the onset and duration of exposure
- the level and duration of dosage
- the genetic constitution of the foetus
- the timing of the exposure in relation to the stage of foetal development.

Substances that inhibit mitosis (e.g. antineoplastic agents such as vincristine and vinblastine) are also of particular risk to pregnant women and exposure to such agents may lead to teratogenicity and embryotoxicity. The female foetus is sensitive to toxic chemicals or to other agents affecting gametogenesis, which in humans finishes by the seventh month.

Access of an agent to the foetus is determined by its lipophilicity, molecular weight or ionic nature. Generally the more lipophilic a chemical is the more likely it is to cross the placental barrier. For large molecules like polymers, size generally prevents their passage across the placental barrier. Most teratogenic effects are also dose-related; that is, the larger the dose, the more likely and severe the effect. High-dose exposures to polychlorinated biphenyls (PCBs) have been associated with foetal abnormality.

Timing of exposure is particularly important. The critical period for organogenesis is in the first trimester (between days 18 and 55 of gestation). This is the time of greatest cell differentiation, and environmental agents may have a profound effect on development at this stage.

The extent of the toxicity effect will also depend on the genetically determined detoxification mechanisms (i.e. enzyme systems) of individuals.

Exposure of environmental or occupational agents can also occur at the postnatal stage. The production of milk during nursing and breastfeeding is one pathway for the excretion of contaminants such as lead, mercury, PCBs and organochlorine pesticides (e.g. DDT) stored in other body tissues. Kinetic processes, such as absorption, distribution and elimination, will influence the passage of agents into breast milk. Milk has a high fat and protein concentration and lipid-soluble or protein-bound contaminants pass readily to milk and are dissolved in or bound to the milk fat and protein (Hunt 1982).

2.2.6 Risk assessment and lifestyle factors

Lifestyle factors may have an impact on individual risk assessments and population risk assessments if the activity is widespread. For this reason, where the influence of these factors can be distinguished, the potential influence of lifestyle factors should be clearly identified in risk assessments. Specific lifestyle factors that may have an effect on risk assessment include:

- tobacco smoking
- diet
- hobbies
- recreational drug use
- excessive use of prescribed medications.

Tobacco smoking will affect the exposure assessment component of the risk assessment process because there will be an increase in background exposure to substances found in smoke, such as cadmium, cyanide and polycyclic aromatic hydrocarbons (PAHs).

Tobacco smoking also affects the toxicity assessment component. Maternal cigarette smoking and passive smoking have been associated with respiratory illness, acute toxicity and cardiotoxicity among newborns. Furthermore epidemiological studies have shown evidence of synergistic interaction between human carcinogens and long-term cigarette smoking. The best studied interactions have included joint exposure to tobacco and radon and tobacco and asbestos. Results from epidemiological studies of joint exposure to radon and cigarette smoke have shown an additive or possibly a multiplicative increase in the number of cancers induced and a synergistic decrease in the latency period for tumour induction. Epidemiological studies have shown that asbestos and tobacco administered together can produce an increased incidence in lung cancer that is greater than from the administration of either agent alone and the interaction is considered to be multiplicative by most investigators (NRC 1994).

Diet will also influence the stages of the risk assessment process, particularly the toxicity and exposure assessment stages. Interactions between toxic metals and essential metals from the

diet have been known to affect the risk of toxicity. Absorption of toxic metals from the lung and gastrointestinal tract may be influenced by the presence of an essential metal or trace element if the toxic metal shares the same homeostatic mechanism. Examples are lead and calcium, and cadmium and iron. Other dietary interactions include an inverse relationship between protein content of the diet and cadmium and lead toxicity. Vitamin C in the diet also reduces lead and cadmium absorption.

Different types of food will have different amounts of agents and hence cause a range of toxic effects depending on dietary habits. For example, the major pathway of exposure to many toxic metals in children is food and children consume more joules per kilogram of body weight than adults do. Furthermore, children have a higher gastrointestinal absorption of metals, particularly lead.

Alcohol ingestion may influence toxicity indirectly by altering diet and reducing essential mineral intake. The ingestion of alcoholic beverages (ethanol), fats, protein, calories and aflatoxins has been implicated in carcinogenesis (Klaassen 1996).

Home-grown produce (e.g. vegetables) has been associated with contamination of heavy metals such as lead, arsenic and cadmium. This may be of particular concern when assessing contaminated sites. The NEPM processes used to develop health investigation levels (HILs) incorporate four different exposure scenarios, one of which includes potential for home-grown vegetables as an exposure source. The potential for lipophilic chemicals to be stored and/or bioaccumulate in meat and poultry tissue (e.g. meat, fat, skin) and eggs (egg yolk) is an important consideration in regulating pesticide residues and environmental contaminants in foods (see Chapter 17).

The type of diet can also influence the risk of exposure to hazardous agents. Individuals who are vegetarians will have a reduced exposure to zinc. Individuals who consume barbecued foods can be exposed to relatively large amounts of PAHs from the pyrolysis of fats and other food components during the cooking process. Populations (e.g. general population and fishermen) who consume seafood may be exposed to heavy metals such as mercury in fish and zinc in shellfish (e.g. oysters).

The exposure to a hazard may also be influenced by lifestyle and hobbies. For example, the amount of time spent indoors (e.g. in the home, work environment/office, factory), outdoors or travelling in a car, bus, aeroplane or train will also influence the amount of exposure of agents and the risk to health (e.g. lead, benzene levels in the car, cosmic radiation in aeroplanes, etc.). Hobbies such as pistol shooting in indoor shooting ranges, antique furniture restoration, lead soldering, boat building and lead lighting can also result in an increased exposure to lead (Lead Safe 1997). House renovating can result to an increase exposure to hazardous agents such as lead and asbestos. Other hobbies involving paint stripping using methylene chloride can cause exposure to its metabolic breakdown product (carbon monoxide), and car maintenance can also result in an increase in exposure to hydrocarbons and heavy metals.

2.3 IDENTIFICATION OF POTENTIAL EXPOSURE PATHWAYS

When issues have been identified, a preliminary qualitative risk assessment may be carried out to prioritise issues for a more detailed study. This will consider the likelihood of exposure and the possible consequences, taking into account things such as biological plausibility, evidence of exposure and community concerns. There may be multiple iterations of hazard appraisal as the risk assessment proceeds and new information and perspectives emerge.

The issue identification processes relating to exposure may be aided by the development of appropriate conceptual site models (CSMs), which delineate the exposure sources and potential pathways leading to human exposures. Identification of exposure pathways using CSMs and the modelling and quantitative description of them are discussed in more detail in Chapters 4 and 13.

Where multiple exposure pathways may include background exposure not specifically associated with the source under consideration in the EHRA, consideration needs to be given to allocating a permissible component of the exposure under consideration. Allocating specific proportions of the ADI/TDI to account for background or other sources of exposure is discussed in more detail in Section 5.11.1.

Other issues relating to identifying exposure pathways and quantitation of pathway inputs are discussed further in Chapter 4.

2.4 POTENTIAL INTERACTIONS BETWEEN AGENTS

There may be interactions between the physical, chemical, biological and social hazards that need to be identified and considered as part of the risk assessment. For example, malnutrition may increase the absorption of cadmium and hence the risk of renal dysfunction. A high zinc intake may reduce the gastrointestinal absorption of cadmium, reducing the risk from high environmental levels. People who carry the sickle cell anaemia gene have a reduced risk of malaria, while people with the genetic condition of Wilson's disease have a greatly increased risk from environmental copper.

There are several potential types of interaction between hazardous agents:

- additive, where the combined effect of two or more agents is equal to the sum of the individual effects (e.g. 2 + 3 = 5) – an example is cholinesterase inhibition from simultaneous exposure to two organophosphorus insecticides
- synergistic, where the combined effect of two or more agents is much greater than the sum of the individual effects (e.g. 2 + 2 = 20) – examples are risk of lung cancer from asbestos and smoking and the hepatotoxicity of carbon tetrachloride and ethanol
- potentiation, where one agent alone does not have a toxic effect but, when given with another agent, results in a much greater toxic effect from the other agent (e.g. 3 + 0 = 8) – an example is risk of cancer from an initiator and a promoter (tobacco smoke contains both)
- antagonistic, where the combined effect of two or more agents is less than the sum of the individual effects – an example is risk of cyanide toxicity from cyanide after receiving an antidote such as Kelocyanor (Klaassen 1996).

The potential hazards from interactions between chemicals are widely discussed but there are no generally accepted methods for predictive appraisal of interactions as part of the risk assessment process. Some contemporary approaches to the health risk assessment (HRA) of chemical mixtures are discussed in Chapter 12.

2.5 IDENTIFYING POTENTIAL MANAGEMENT OPTIONS THAT MAY MITIGATE EXPOSURE

While risk management should be considered to be a process separate from or dissociated from risk assessment, a formal EHRA is likely to identify those hazards and exposure pathways that make the greatest contributions to the overall risk. This may include assessment of possible future risks associated with continuing or expansion of existing operations. Information derived from an EHRA will be useful to risk managers in formulating and prioritising risk mitigation options. These may include:

- closing down/ceasing use of the hazard source altogether or substituting with a less hazardous material where minimisation of any further environmental contamination is required
- cleaning up a contaminated site using the EHRA outcomes and/or statutory instruments to guide the level of clean-up required – this may require a combination of *in situ* hazard treatment or containment, or removing the hazardous material to another site
- sealing off the contaminated environment to prevent further access by human receptors
- preventing ongoing release to the environment or denying development plans that could increase such release.

In some instances, the hazard and need for action will be so obvious to all stakeholders that risk assessment will be undertaken only to determine the effect and cost-effectiveness of the various management options. In this situation, the costs of undertaking a risk assessment to determine whether action is necessary are considerable. In other instances, risk assessment will be inappropriate because the solutions to the problem will not be based on addressing risk but on addressing other factors such as social and political concerns.

Risk management also needs to be understood within the inevitable constraints that it will operate when risks are found to be small. Risk management inevitably involves trying to steer a sensible course between making a Type 1 (false positive) or Type 2 (false negative) error (Hrudey & Leiss 2003). In other words, risk management for small and uncertain risks involves trying to decide between taking action when none is required or failing to take action when it is required.

2.6 WHAT RISK AND OTHER TECHNICAL ASSESSMENTS ARE NECESSARY TO EVALUATE RISK AND DISCRIMINATE BETWEEN POTENTIAL RISK MANAGEMENT OPTIONS?

A formal EHRA will generally provide sufficient information to identify the major contributors to risk and enable the risk manager to prioritise the needs for risk mitigation. Having then identified technically feasible risk management options, the next phase would be to undertake an economic analysis of the costs and benefits associated with each of these risk management options. The economic assessment processes are beyond the scope of this enHealth document, but some guidance may be found in enHealth monographs on economic evaluation of environmental health issues (enHealth 2003).

Chapter 3: Hazard identification and dose–response assessment

3.1 INTRODUCTION

The two elements of risk assessment discussed in this chapter are:

- hazard identification (using toxicity test data)
- dose–response assessment.

These elements are identified as part of Phase II of the expanded framework for EHRA outlined in Figure 2.

There are essentially two levels of hazard identification commonly undertaken in risk assessments in Australia. For many risk assessments developed as part of environmental protection licensing, planning processes or contaminated sites assessments, the hazard identification component may simply identify the relevant national or international guideline values for each chemical that may be present. For risk assessments undertaken by national chemicals regulators or those setting national guidelines, the assessment will generally involve a full investigation of the international toxicity literature relevant to the chemical, including an appraisal of the dose–response relationships that underpin any derived guideline values. This chapter focuses on an understanding of dose–response relationships that provide insight into the development of health-based guideline values, while Chapter 5 outlines the sources of guideline values and other information that can be used in the risk characterisation process.

Additional detail on the design and interpretation of animal-based toxicity tests is included in Chapter 9, along with a discussion of some of the newer techniques for hazard assessment (*in vitro* and *in silico* techniques, genomics, structure–activity analysis and 'read-across' from comparable substances) where the toxicity database for a new industrial chemical may be less comprehensive than other types of regulated chemicals.

3.2 HAZARD IDENTIFICATION

Hazard identification examines the capacity of an agent to cause adverse health effects in humans and other animals (US EPA 1995a). It is a qualitative description based on the type and quality of the data, complementary information (e.g. structure–activity analysis, genetic toxicity, pharmacokinetic) and the weight of evidence from these various sources.

Hazard identification uses:

- animal data – this is usually assessed by toxicological methods
- human data – this is usually assessed by epidemiological methods when groups of people are involved, or by toxicological methods when using case studies and acute chamber studies (both qualitative and quantitative toxicity information is evaluated in assessing the incidence of adverse effects occurring in humans at different exposure levels)
- other data – this includes data such as structure–activity data or *in vitro* data assessed by toxicologists.

The data may come from a range of sources such as ad hoc data, anecdotal data, case-report data and data collected from epidemiological registries (including cancer or pregnancy outcome data). In each instance, the quality of the study design and methodology, as well as the resulting data, will need to be rigorously assessed.

Section 5.12 provides guidance on sources of toxicological information and where to find health-based guideline values. There is also guidance on what to do when no suitable toxicological data appears to be available (see Section 5.13).

In the case of data derived from experimental studies in animals (see Chapter 9), a comprehensive data package will generally consist of:

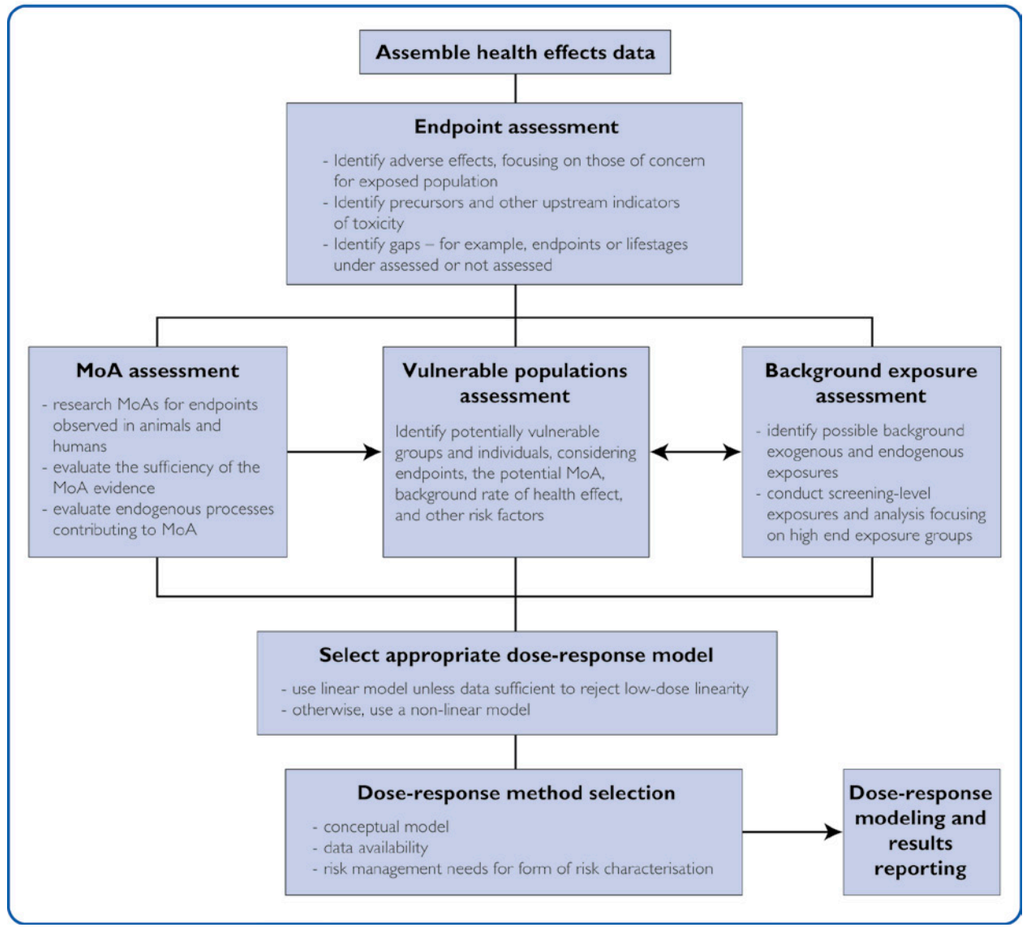
- **Acute toxicity:** Studies investigating the effects of single doses of a substance. The LD_{50} test, or medium lethal dose test are typical examples. The standard acute toxicity studies also include tests for: acute oral, dermal and inhalational toxicity; eye irritation; skin irritation; and skin sensitisation.
- **Sub-chronic toxicity:** Short-term, repeat-dose studies, generally having an exposure duration up to 90 days in rodents. The main purpose of sub-chronic testing is to identify any target organs and to establish dose levels for chronic exposure studies.
- **Chronic toxicity:** Studies lasting for the greater part of the life span of the test animals, usually 18 months in mice and 2 years in rats. Chronic studies are particularly important for assessing potential carcinogenicity.
- **Reproductive toxicity:** Studies designed to provide general information about the effects of a test substance on reproductive performance in both male and female animals.
- **Developmental toxicity:** Studies in pregnant animals that examine the spectrum of possible *in utero* outcomes for the conceptus, including death, malformations, functional deficits and developmental delays. More recent developments extend the period of dosing and/or observation into the neonatal period, to assess potential neurobehavioural effects and other potential post-partum toxicity.
- **Genotoxicity:** Studies designed to determine whether test chemicals can perturb genetic material to cause gene or chromosomal mutations.
- **Other tests:** Specific tests developed for endpoints such as neurobehavioural toxicity, developmental neurotoxicity and various *in vitro* tests (e.g. skin absorption, irritancy potential and endocrine-related endpoints), which aim to reduce or eliminate the *in vivo* use of animals, on the grounds of addressing animal welfare issues.

Key issues include:

- nature, reliability and consistency of human and animal studies
- the availability of information about the mechanistic basis for activity
- the relevance of the selected animal studies to humans
- whether the mode of toxic action is well understood – knowledge of the mode of action (MoA) is becoming increasingly important in interpreting carcinogenic responses (see Chapter 11) and assessing the risk of chemical mixtures (see Chapter 12).

Various sources of information are needed to identify and characterise environmental hazards (Figure 6). Integrating information on MoA, exposures (including background exposures) and identifying susceptible populations are all important factors in determining the correct use of dose–response information.

Figure 6: Potential sources of information used to identify and characterise environmental hazards, leading to characterisation of mode of action (MoA), dose–response models and susceptible populations



NRC 2008. Reprinted with permission from the National Academies Press, Copyright 2009, National Academy of Sciences

3.3 DOSE–RESPONSE ASSESSMENT

Dose–response assessment examines the quantitative relationships between exposure and the effects of concern. The determination of whether there is a hazard is often dependent on whether a dose–response relationship is present.

Important issues include:

- the relationship between the extrapolation models selected and available information on biological mechanisms
- how appropriate datasets were selected from those that show the range of possible potencies both in laboratory animals and humans
- the basis for selecting inter-species scaling factors to account for scaling doses from experimental animals to humans
- relevance of the exposure routes used in the studies to a particular assessment and the interrelationships of potential effects from different exposure routes
- the relevance to the assessment of the expected duration of exposure and the exposure durations in the studies forming the basis of the dose–response assessment
- the potential for differing susceptibilities in population sub-groups
- dose averaging/averaging exposure.

3.4 THE PRIME ROLE OF ANIMAL STUDIES IN RISK ASSESSMENT

While a risk assessment may be able to access data from controlled exposure studies using either animals or human (including epidemiological studies; see Chapter 10 for more detail), animal studies have several advantages that may be exploited in the risk assessment process (see Chapter 9):

- Exposures can be controlled so that groups are consistently exposed to known amounts of chemicals.
- The studies may be supplemented by useful data on bioavailability, target tissue dose and possibly other toxicokinetic parameters.
- The incidence of disease in groups treated with defined dose levels can be ascertained using a combination of observational, clinical measurements (blood and urine analysis growth) and post-mortem tissue pathology assessment.

These advantages are, to some extent, offset by a lack of knowledge of the following.

- The adverse effect observed may or may not be relevant to be extrapolated to humans (i.e. is the effect species-specific because of some basic difference in physiology or metabolic function).
- The dose–response relationship in animals is not so relevant for low-dose extrapolation (i.e. since such studies use relatively high doses in order to be able to demonstrate dose-related toxicity with an animal sample size kept to practical limits, the extrapolation to lower doses may be compromised if such responses are only likely to be seen at very high doses).
- The results have been obtained in healthy and genetically homogeneous strains of animals (usually rodents) and are therefore less likely to be representative of the human population with its genetic variance and variable background health status.

Studies in which humans have been exposed to chemicals, either in occupational settings or in controlled laboratory experiments (e.g. inhalation chambers) may provide data whose relevance certainly exceeds that derived from animals. However, ethical constraints limit the levels of exposure that can be used in such tests, and they would certainly be precluded where there is an expectation of an adverse outcome.

3.5 DOSE SCALING

The dose administered is a critical component of the dose–response relationship. To calculate a dose relevant to EHRA, it may be necessary to undertake dose scaling to convert the doses used in animals to a human-equivalent dose (HED) or human-equivalent concentration (HEC).

In the animal studies that form the mainstay of much toxicological information used in EHRA, the dose is generally known and controlled. The route of administration in animal studies and the frequency of dosing may not be directly relevant to the exposure conditions associated with environmental exposures.

For example, many animal studies use dietary admixture to provide a continuous intake of test chemical and it is then necessary to measure food intake (or use standardised food intake conversion factors) to convert the dietary concentration to 'dose'. Volatile chemicals or dusts administered by the inhalational route (nose-only, head/nose-only or in an exposure chamber) may have intermittent exposure schedules (e.g. 6 hours/day 5 days/week) that are different from those associated with environmental exposures.

These factors can be taken into consideration in using the data to calculate a dose equivalent to human exposure (see Section 4.11 for further discussion of time scaling).

3.5.1 Inter-species scaling of doses

Where animal studies provide the dose–response data used in EHRA, scaling the dose to provide an HED is a critical step in the process. For many years, the simplest form of conversion or dose scaling has been expression of the dose on an equivalent mass basis – e.g. as mg/kg body weight. This simplistic scaling approach is still very widely used in EHRA, although other forms of dose conversion may provide better estimates of the HED.

Scaling doses on the basis of body surface area (SA) has been used for animal–human dose conversion by many pharmaceutical regulatory agencies, although a direct conversion based on pharmacokinetic data is usually available for toxicological assessment of therapeutic agents. Where toxicokinetic data is available for scaling conversion for other types of chemicals, this is also the preferred approach, although such data may be less available to risk assessors engaged in EHRA.

The 1986 US EPA cancer guidelines (and many other US EPA guidance notes) recommended a SA-based scaling factor of $bw^{2/3}$ for converting oral doses in animal–human-equivalent doses. In 2005, the US EPA guidance on scaling proposed altering the SA-based scaling factor to $bw^{3/4}$, although many of the dose-scaling conversions found in the literature will have used the earlier $bw^{2/3}$ factor. The recent update of US guidance on dose–response assessment in risk assessment (NRC 2008) also recommends a dose-scaling factor of $bw^{3/4}$. Further detailed guidance on the use of body weight and SA conversion factors to produce HED estimates is found in US guidance (US EPA 2005a p. A7):

> As a default for oral exposure, a human equivalent dose for adults is estimated from data on another species by an adjustment of animal applied oral dose by a scaling factor based on body weight to the ¾ power. The same factor is used for children because it is slightly more protective than using children's body weight. This adjustment factor is used because it represents scaling of metabolic rate across animals of different size. Because the factor adjusts for a parameter that can be improved on and brought into more sophisticated toxicokinetic modeling when such data become available, they are usually preferable to the default option.
>
> For inhalation exposure, a human equivalent dose for adults is estimated by default methodologies that provide estimates of lung deposition and internal dose.

Where oral doses are expressed in parts per million (ppm) in the diet or drinking water, the dosage needs to be converted to mg/kg body weight using appropriate estimates of food or water consumption and body weights (see WHO 1987; Faustman & Omenn 1996).

Conversion factors commonly used for such conversions are cited in Table 5.

Note that, while these conversion factors may provide rough estimates of dose, it is always preferable to calculate actual doses from feed intake and chemical analytical data, where available.

Another conversion commonly needed in EHRA is changing the expression of concentration of substances in air from ppm to mg/m^3 (the preferred units). The conversion equation is:

$$mg/m^3 = ppm \times \frac{MW}{V}$$

where V (volume of 1 g mole) = RT/P

R is the universal gas constant (so that V = 24.5 at 25°C and 760 mm Hg) and MW is the molecular weight of the substance.

3.5.2 Route-to-route scaling

Often the toxicological data is not available for the most appropriate route of exposure for humans. For example, only oral carcinogenicity data may be available from which reference values have been calculated (e.g. cancer slope factor – CSF), whereas environmental exposure by oral, dermal and inhalational routes may be important. Thus, extrapolation from one route of exposure to another may be necessary; this needs to be assessed on a case-by-case basis depending on the available data.

One important consideration in route-to-route extrapolation is determining whether the adverse health effects are localised to the exposure site or whether they are a consequence of systemic distribution.

Table 5: Dose conversion from mg/kg (ppm) in diets to mg/kg/day in animal toxicity studies

Species	Weight (kg)	Food consumed per day (g)	1 ppm in food = mg/kg/day	1 mg/kg/day = ppm in food
Mouse	0.02	3	0.15	7
Rat – young – adult	0.1 0.4	10 20	0.1 0.05	10 20
Guinea pig	0.75	30	0.04	25
Rabbit	2	60	0.03	33
Dog – dry chow – moist chow	10	250 750	0.025 0.075	40 13
Monkey	5	750	0.05	20

If the effects are localised at the exposure site and not a consequence of the systemic distribution of the agent, then it is not appropriate to extrapolate the dose to a different route of exposure. If the effects are consequent to absorption and systemic distribution of the agent, then dose scaling between routes of exposure needs to account for the bioavailability of the agent by the different routes.

Therefore, bioavailability is an important consideration when extrapolating the applied dose to different routes of exposure. However, additional factors may need to be considered, such as physiological differences between species when extrapolating, for example, from inhalational exposure in animals to oral exposure in humans or vice versa. The assessor should include information about the bioavailability of the chemical agent in the experimental studies in the final report. There is further discussion of the importance of assessing bioavailability in Section 4.2.1.

In cases where bioavailability data is not available, important clues may be gained from the physical and chemical properties and physical state of the agent (e.g. liquid, solid, gas).

Unless specific dermal toxicity data is available, risk assessments involving the dermal route of exposure may necessitate the use of toxicity data derived from another exposure route (usually oral). The toxicity reference dose (e.g. TDI, RfD) can be adjusted as follows:

$$TDI_{dermal} = TDI_{oral} \times GAF$$

(GAF = gastrointestinal absorption factor)

A similar adjustment approach for non-threshold toxicity replaces TDI with CSF.

The basis for application of a GAF to an oral dose is outlined in the US EPA guidance on dermal risk assessment (US EPA 2004b). The adjustment is not necessary where oral bioavailability approaches 100 per cent because it is assumed that both the oral and dermal routes will deliver an equivalent target tissue dose. The US EPA guidance recommends the adjustment when:

- the toxicity data from the critical oral dosing study is based on an administered dose, and
- there is a scientifically defensible database that shows the gastrointestinal absorption from a medium similar to the one used in the critical study (water, feed or gavage vehicle) is significantly less than 100 per cent (i.e. <50 per cent). The 50 per cent cut-off is proposed since it is thought to reflect the intrinsic variability in data from absorption studies.

There are relatively few substances for which GAFs have been determined. US EPA guidance (US EPA 2004b) has tabulated GAFs for a few selected metals and metalloids, of which only five out of nine listed have GAFs below 50 per cent.

Route-to-route dose scaling is mainly done for adjusting doses used in animal studies, but it is an equally important consideration where human data is used.

3.5.3 Other factors in dose scaling

For inhalational exposure, doses expressed as mg/m^3 or ppm may need to be converted to mg/kg body weight in the test species by calculations based on the physical properties of the agent and minute volumes and respiration rates of the animal (Kennedy Jr & Valentine 1994; US EPA; 1994). However, more recent US EPA guidance (US EPA 2009a) proposes that such conversion is not needed for airborne substances where the toxicological information suggests a dose–response threshold (see Sections 4.6 and 4.12).

The process for converting doses derived from animal studies to an HED or concentration (HED/C) is detailed in US EPA guidance (US EPA 2005a).

A procedure for deriving an HED for inhaled particles and gases has been described by Di Marco and Buckett (1996).

3.6 DOSE–RESPONSE CURVE SHAPE, CONSTRUCT AND ANALYSIS

Elaboration of dose–response relationships is a fundamental element of contemporary risk assessment. The availability of human data on dose–response is often quite limited unless well-structured, ethically justified and controlled experiments have been conducted, Epidemiological studies may be useful for providing estimates of whether health risk in a population exposed to an environmental chemical, or group of chemicals, has been increased in comparison to an unexposed group or population. In some cases, the studied cohorts may be categorised in relation to exposures to one or more chemicals in the environment. However, such human data is rarely, if ever, as complete for the purposes of constructing dose–response relationships as that available from controlled toxicity studies using animals because they are based on estimates of exposure rather than actual measured exposure. In large part, the default database used in most health risk assessments relies on extrapolating human health risks based on dose–response relationships from animal studies.

There are different ways of characterising dose–response relationships. In risk assessment, dose–response relationships based on experimental studies in animals are often characterised as shown in Figure 7. The dose or exposure scale is generally logarithmic while the response scale represents the proportion of the tested animals that respond at the doses used. This scale may also be converted to a probit scale based on the properties of a Gaussian distribution curve. Such a transformation can linearise a large part of the dose–response curve.

Mathematical models may be used to fit the experimental data to smooth curve, but, as shown in Figure 7, the data points from which the curve is derived may be scattered around this relationship.

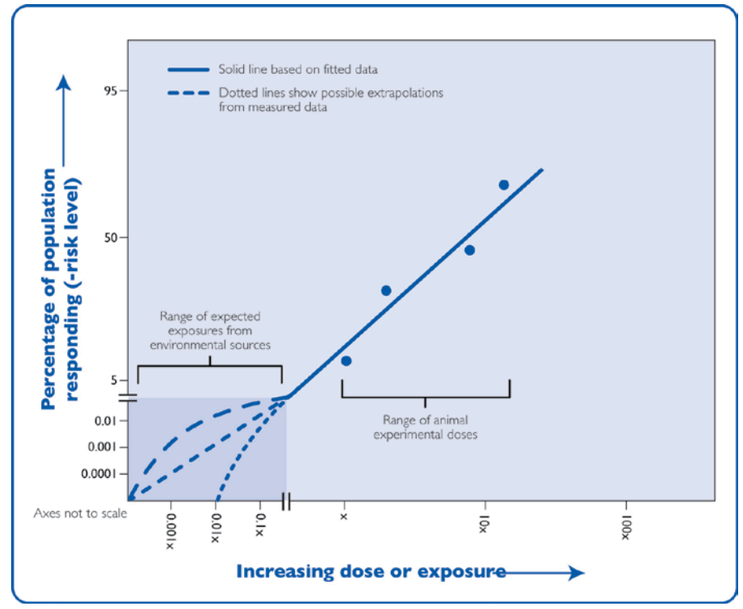
The shape of the dose–response curve below the experimental range can have multiple shapes depending on the model used. The choice of the model should, where possible, be based on mechanistic information.

It is obvious that different dose–response relationships can be elaborated where the chemical causes more than one form of toxicity (e.g. growth retardation, hepatorenal toxicity or neurotoxicity) (see Figure 8). The dose–response relationship furthest to the left (i.e. occurring at the lowest dose range) will usually drive the risk assessment process, unless it is considered that the effect (e.g. mild liver enlargement; changed enzyme activity with no obvious toxic consequence) is not considered to be 'adverse'. However, enzyme activity changes, such as the release of lactic dehydrogenase from tissues into blood, are commonly used as a surrogate for measuring organ toxicity, so would be relevant to risk assessment.

3.6.1 Individual versus population risk

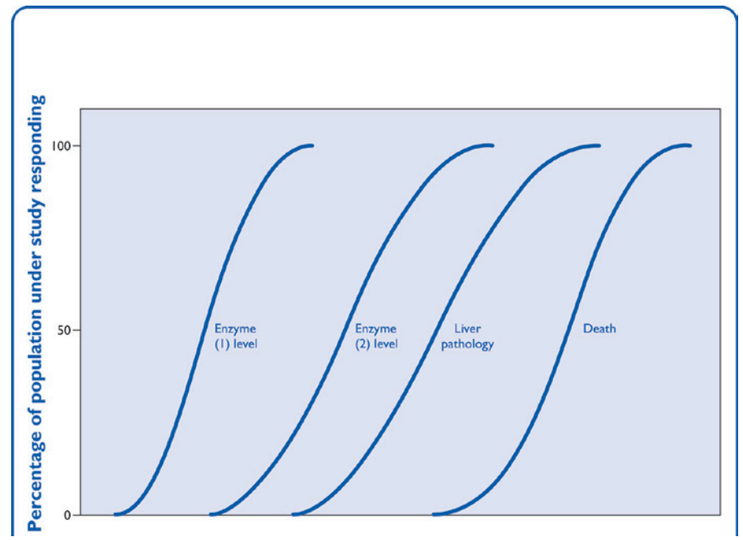
There is an extensive discussion of the conceptual models and implications for health risk assessment when extrapolating from dose–response curves that describe individual risk to those that characterise risk in a population (NRC 2008). Differences in individual sensitivity can markedly influence the shape of the population dose–response curve. For example, differences in sensitivity in a vulnerable group or bimodal differences in sensitivity can flatten out the lower part of the dose–response curve (Figure 9). Additionally, extrapolation of population risk can be influenced by variability in the background exposure or spontaneous

Figure 7: Hypothetical curve for an animal carcinogenicity study



Adapted from: Levy & Wegman 1988.

Figure 8: Different dose–response curves for different effects from a hypothetical substance

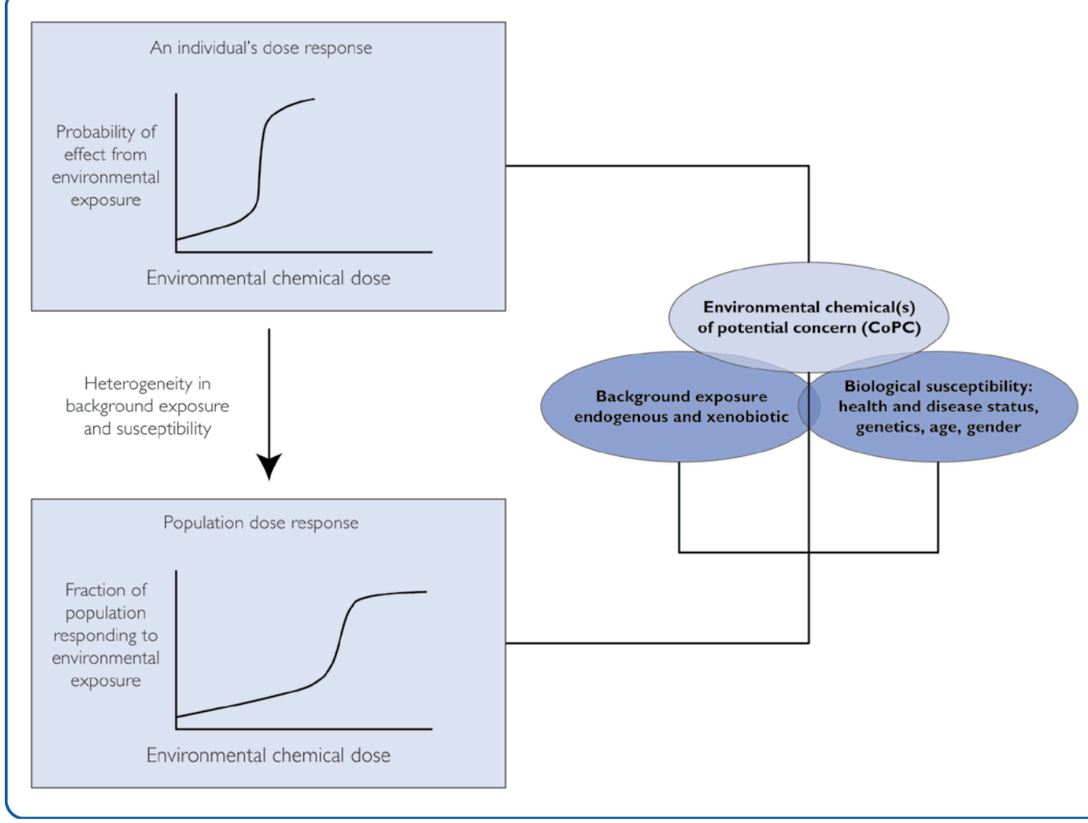


incidence of the disease that is characterised in the individual dose–response curves or by the assumptions made about whether low-dose effects are best expressed by linear or non-linear functions.

It is also possible to incorporate estimates of uncertainty associated with varying background exposure and heterogeneity of biological response and susceptibility, using upper-bound probability estimates of the population risk (see Figure 10).

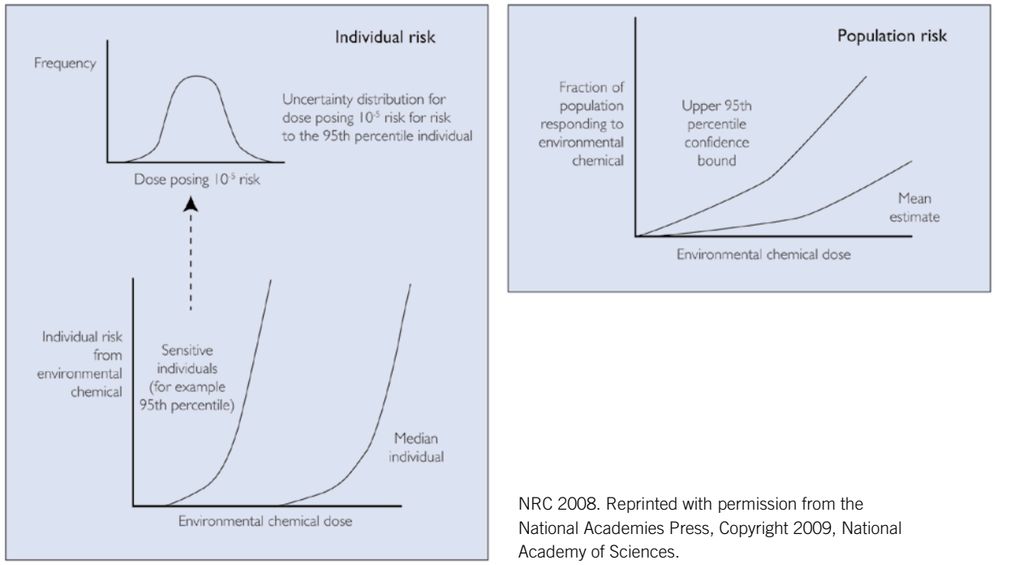
Individuals will vary in the lowest dose that can initiate the response, so that dose–response curves relating to more susceptible individuals will be furthest to the left, while more resistant individuals will have dose–response curves further to the right. This translates into a population dose–response relationship that may be curved at the bottom end. Where the population curve intersects the axes will be determined by whether the relationship is assumed to have a threshold above the background level of exposure (non-linear population extrapolation in Figure 11, Model 2) or a non-threshold relationship for both individuals and populations (Figure 11, Model 3).

Figure 9: A conceptual model that integrates individual risk dose–response, background disease incidence and variation in susceptibility into population dose–response



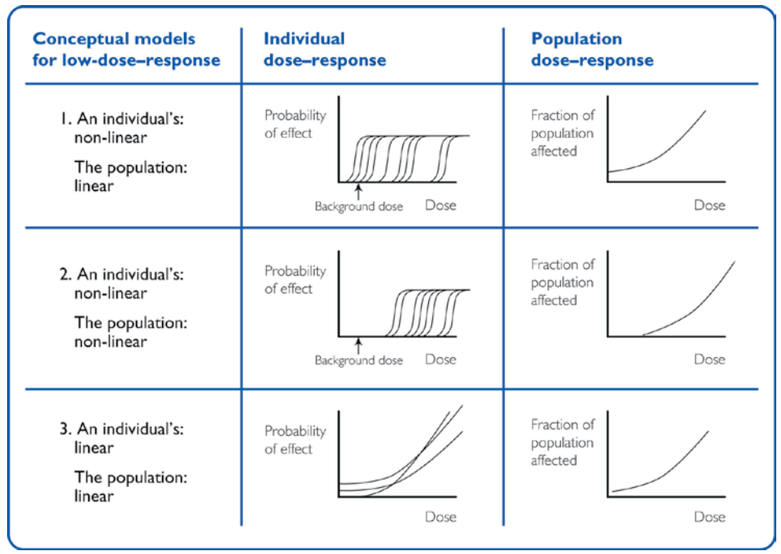
NRC 2008. Reprinted with permission from the National Academies Press, Copyright 2009, National Academy of Sciences.

Figure 10: Extrapolation from the individual risk to population risk incorporating uncertainty estimates



NRC 2008. Reprinted with permission from the National Academies Press, Copyright 2009, National Academy of Sciences.

Figure 11: Conceptual models describing low-dose linearity or non-linearity of individual and population dose–response curves



NRC 2008. Reprinted with permission from the National Academies Press, Copyright 2009, National Academy of Sciences.

3.7 HORMESIS

The basic tenet of dose–response assessment in EHRA is that the relationship between dose and toxic effect is monotonic (i.e. response always decreases as the dose or exposure is lowered).

The basic concept of hormesis is that dose–response curves in toxicology are non-monotonic. That is, they may resemble a J-shape or possibly U-shape at low dose and that, as the dose or exposure gets lower, the risk of harm may actually increase rather than decrease, at least over part of the dose–response range. In fact, it has been argued that non-monotonic dose–response relationships are quite common in toxicology, and that there are rational mechanistic and toxicokinetic explanations for the phenomenon (Connolly & Lutz 2004).

Hormesis theory presents important implications for risk assessment, since extrapolation of risk to very low levels of exposure may be confounded by such a fundamental change in shape of the dose–response curve. The issue becomes even more complex when it has been shown that, for the same chemicals, both hormetic and non-hormetic responses may be observed in different tissues (Borak & Sirianni 2005).

One of the more controversial aspects of hormesis theory is that part of the low dose–response relationship may describe a phenomenon where the risk of harm is actually less than in unexposed (controls); that is, low dose exposure may have a beneficial or protective effect.

The case for beneficial effects of toxic chemicals at low doses has been supported to some extent by analyses of cancer rates for dioxin). Tuomisto et al. (2005) reported that cancer rates for some sites in rats were actually lower at low doses than in controls, implying a protective effect. The study went on to note that cancer risk, as reflected in case-control studies in the Finnish population, appeared to decrease slightly in those groups showing higher levels of dioxin body burden, where dioxin exposure through food is the dominant exposure matrix.

While the concept of hormesis attracts vigorous debate in the scientific literature (Calabrese 2005; 2008; Thayer et al. 2006; Mushak 2007) and has been addressed to some extent in the most recent US guidance on risk assessment (NRC 2008), there is no consensus that the basic concepts of risk assessment (i.e. that it is possible to define low doses or exposures where risk is either negligible or non-existent) should be modified to accommodate the concept of hormesis.

The extent to which the debate over hormesis has polarised the scientific community is well set out in a number of commentaries (e.g. see Axelrod et al. 2004; Calabrese 2004; Kaiser 2003; Renner 2004).

A separate but related consideration is the concept of the acceptable ranges of oral intakes for essential trace elements, where there is a need to ensure tolerable intakes do not fall below the minimum requirements. WHO (1996) regards iron, zinc, copper, chromium, iodine, cobalt, manganese, molybdenum and selenium as unequivocally essential for human health. However, it is not clear that the toxic effects associated with vitamin deficiency and overdose is an appropriate example of hormesis. The adverse effects associated with vitamin overdose are different from those associated with vitamin deficiency. Therefore, it is inappropriate to depict these different adverse effects as a continuous dose–response curve or as an example of hormesis.

3.8 DOSE–RESPONSE MODELLING

There are various ways of managing dose–response data using mathematical equations to derive dose–response curves that fit the experimental data. They assume that the toxic effect results from the random occurrence of one or more biological events. These are known as stochastic events (Klaassen 1996).

Mechanistically derived models have been particularly used for cancer modelling (especially based on radiation exposures). The simplest form is a 'one-hit' linear model in which only one 'hit' or critical cellular interaction results in the alteration of a cell. This model would propose that a single molecule of a genotoxic carcinogen would have a 'minute but finite chance of causing a mutational event' (Klaassen 1996). From these models more complex models based on multi-hits or multi-stage events have been derived. Although conceptually based on biological mechanisms, most of these models do not rely on independently validated parameters describing the mechanisms, but rely on fitting curves to empirically observed data.

More recently these models have been adapted to take into account information based on knowledge of the relevant physiology and toxicokinetics (physiologically based toxicokinetics/ pharmacokinetics modelling – PBPK). These models take into account the effective dose at the target organ. A further development has been to make generalised mechanistic models take into account specific biological processes. An example of these biologically based dose–response models is the Moolgavkar–Venson–Knudson model that uses a two-stage model for carcinogenesis (Klaassen 1996).

3.8.1 Mathematical models used in risk extrapolation

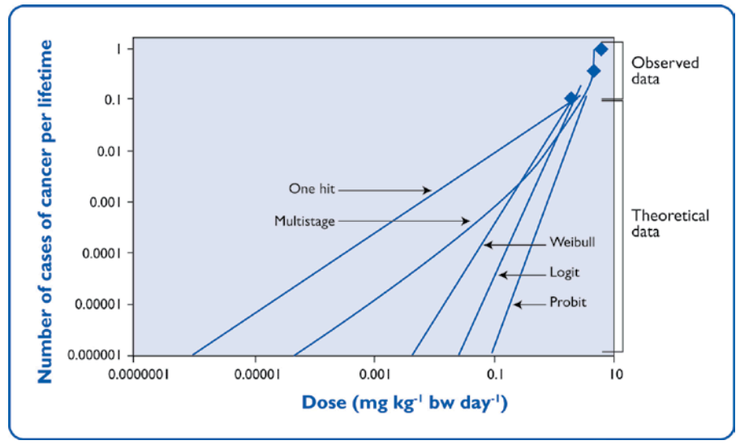
The BMD approach requires that all the experimental data is used to derive a 'best fit' in determining the shape and position of the dose–response curve. A variety of mathematical models have been developed to facilitate such curve fitting. They may be subdivided into:

- statistical or distribution models
  - log-probit
  - logit
  - Weibull
- 'mechanistic' models
  - one-hit
  - multi-hit
  - multi-stage
  - linearised multi-stage
  - stochastic two-stage model (Moolgavkar–Venson–Knudson)
- model enhancement
- time–tumour response
- physiologically based toxicokinetic models
- biologically based dose–response models.

A schematic diagram illustrating the critical effect on risk levels if an inappropriate extrapolation model is chosen is shown in Figure 12.

For a more extensive coverage of the mathematical concepts and models used in quantitative risk assessment, refer to monographs by David Vose (2008) or Dennis Paustenbach (2002), or more general monographs on the principles of environmental health risk assessment (Fjeld et al. 2007; Robson & Toscano 2007).

Figure 12: Examples of the potential variability when different models are used to extrapolate risk



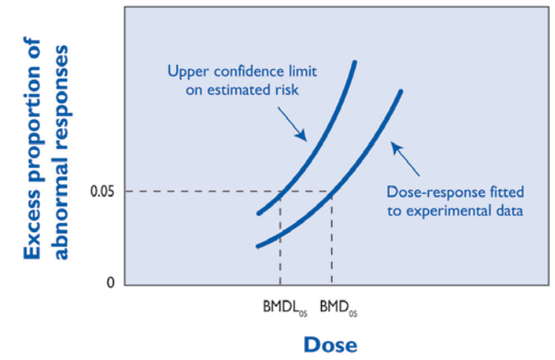
3.9 BENCHMARK DOSE APPROACH

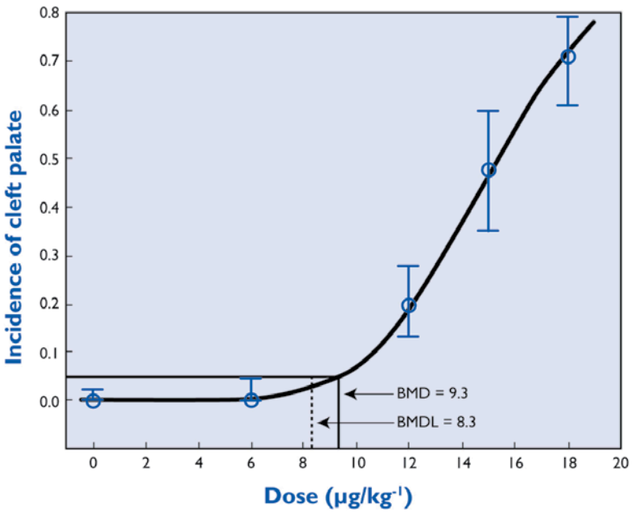
The BMD approach has been used in dealing with both cancer and non-cancer endpoints. It is described in EHC170 and a modified version for use with carcinogenic soil contaminants was described in NHMRC (1999a).

There has been more recent discussion of the utility of the BMD approach in the literature (for reviews, see Falk-Filipsson et al. 2007; Travis et al. 2005; Sand et al. 2008) and various guidelines have been issued by national bodies regarding it applications in risk assessment (e.g. see US EPA 1995b, 2000b; Appel et al. 2001; Health Council of the Netherlands 2003; IPCS 2009c). The influence of model- and dose-level selection has been evaluated, as well as assumptions about the underlying nature of the data (continuous or dichotomous) (Slob 2002; Sand et al. 2002; 2003; 2006). Imprecision in exposure estimates based on epidemiological studies can lead to a derived BMD estimate that is biased towards a higher and less protective level (Budtz-Jorgensen et al. 2004).

The BMD corresponds to a predetermined increase (between 1 and 10 per cent but commonly 5 per cent) of a defined effect in a test population. Mathematically it is the statistical lower confidence limit on the dose that corresponds to that predetermined increase, derived by extrapolation from the upper confidence level of the dose–response curve (Figure 13), although some agencies are using a best estimate rather than a lower confidence limit (IEH 1999b).

Figure 13: Graphical illustrations of the benchmark dose





(a) stylised to show the extrapolation of BMD_{05} and its lower confidence limit (based on upper confidence limit of incidence)

(b) illustration of BMD calculation from real data on malformations induced by TCDD Sand et al. (2008) using data from Birnbaum et al. (1989)

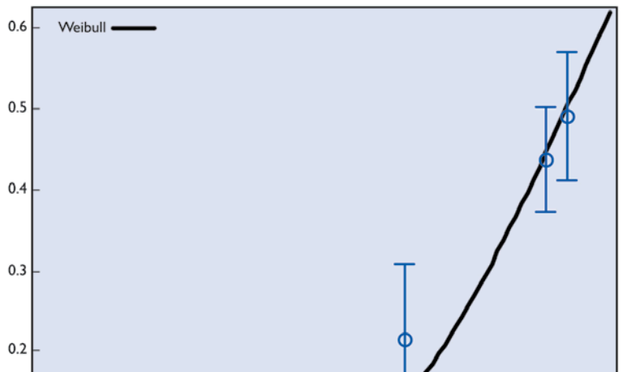
The stylised example (Figure 13a) shows the derivation of the BMD_{05}, the dose estimated to result in a 5 per cent increase in a defined effect. It also shows the derivation of the lower confidence limit (BMDL), which some agencies prefer as the estimate of BMD_{05}. The dose–response relationship (Figure 13b) shows the derivation of the BMD_{05} (9.3 μg/kg) and $BMDL_{05}$ (8.3 μg/kg) for some real toxicity data – malformations in mice after maternal dosing with TCDD (Sand et al. 2008, using data from Birnbaum et al.1989). It should be noted that, in this case, the figures derived from the BMD approach are between the NOAEL (6 μg/kg) and LOAEL (12 μg/kg) that are predicated by the dose selection in the study.

Fully maintained software for using the BMD approach sanctioned by the US EPA is available from the US EPA National Center for Environmental Assessment (NCEA) website at <http://cfpub.epa.gov/ncea/index.cfm>, along with tutorials on how to use the software.

As part of the NHMRC process for evaluating the mean BMD methodology (see Section 11.7), the CSIRO developed software that could calculate a mean BMD from most of the above mathematical models, and derive goodness-of-fit parameters that could aid selection of best-fitting curves. An illustration of the application of the CSIRO software modelling for a specific chemical of potential concern (COPC) is shown in Figure 14.

For developmental toxicity the BMD_{05} values have been similar to statistically derived NOAELs for a wide variety of developmental toxicity endpoints (Klaassen 1996). BMD approaches are also being developed and tested in regard to acute inhalation toxicity (Fowles et al. 1999), to the relationship between the BMD and the MTD (Gaylor & Gold 1998), and to addressing statistical procedures available for calculating BMDs and their confidence limits for non-cancer endpoints (Gaylor et al. 1999).

Figure 14: Analysis of data on formaldehyde toxicity using different mathematical curve-fitting models



Mathematical model	BMC_{05} (mg/m³)	GOF
Weibull	9.05	0.35
Gamma	9.01	0.58
Probit	8.91	0.90
Log-logistic	9.04	0.58
Multistage	5.83	0.01
Quantal linear	2.39	0.00
Quantal quadratic	5.52	0.0001
Average excluding those models with poor fit	9.00	

(a) The fitted curve using the Weibull model and (b) a table showing the goodness of fit (GOF) and derived BMD estimates for a number of different curve-fitting models.

Particular advantages of the BMD approach include that it:

- takes into account information from the entire dose–response curve rather than focusing on a single test dose such as is done with the NOAEL approach (i.e. uses all available relevant information)
- uses responses within or near the experimental range rather than relying on extrapolations to doses considerably below the experimental range
- uses a consistent benchmark-response level that crosses a range of studies and endpoints
- is less influenced than NOAEL approaches by the arbitrary selection of doses (Crump 1984)
- is able to be rigorously described
- allows potency comparisons between endpoints and between studies.

Its disadvantages are that it may not be possible to define the shape of the dose–response curve because of limited dose groups or the number of animals per group. It also requires greater statistical expertise than the NOAEL-type approach (IEH 1999b).

Use of a BMD_{05} provides a more data-sensitive and less model sensitive endpoint than using BMD_{01} (Klaassen 1996; NHMRC 1999).

When the benchmark response is within or near the experimental range of the data, the corresponding values of the benchmark doses are not greatly sensitive to the choice of the model used. The best scientific choice of a model would be a biologically based mechanistic model. Sand et al. (2002; 2003; 2006) confirmed that a BMD range within 5–10 per cent provides an estimate that is less dependent on the dose–response model used and the variation in the data. However, the method used to calculate confidence intervals can influence the precision of the BMD estimate (Moerbeek et al. 2004). Also, study designs with more dose levels improve precision, even when they result in fewer test animals per dose group (Slob et al. 2005).

An important proviso is that where the data fit is relatively poor (more than 18 per cent coefficient of variation), the ability to estimate either a BMD or a NOAEL is more compromised.

3.10 THRESHOLD VERSUS NON-THRESHOLD RESPONSES IN RISK ASSESSMENT

A longstanding convention in risk assessment has been the different treatment of dose–response relationships where either a threshold or a non-threshold relationship is assumed.

There may even be a dichotomy in understanding of the meaning of a 'threshold'. The toxicologist interprets a 'threshold' as a dose or level of exposure where the toxicity response or adverse effect measured is no greater than the background. That is, there is no measurable incremental risk. On the other hand, an epidemiologist may consider a threshold to be a point where the incidence of disease begins to exceed background, especially when the background incidence is not zero.

This concept of a threshold is not necessarily the same as a NOAEL. A NOAEL is defined on the basis of it being the highest dose used in a toxicity test where there is no appreciable increase over controls in the incidence of the adverse effect. The NOAEL is therefore influenced by dose selection and other parameters of the experimental system and a true threshold could be set at a different level if more appropriate doses had been selected in the defining study (see Section 5.6).

3.10.1 Threshold approaches

A threshold occurs when the dose or exposure has not reached a critical level sufficient to trigger a response. The concept is well established in relation to receptor-mediated mechanisms, when a certain degree of receptor occupancy is required to trigger a response. A threshold may also appear to occur when biological mechanisms act to reverse the toxic effects. These can include a saturable capacity to inactivate a toxic chemical by metabolism or excretion, or where tissue damage can be repaired up to the point of irreversibility or below a critical dose or level of damage.

A threshold may also be apparent where the level of cell damage or cell death does not reach a stage where tissue damage is apparent, or where the tissue function is compromised. It may be possible to 'kill' individual cells without consequence for the organism as a whole.

The approach with these models is to derive exposure limits such as an ADI, a provisional tolerable weekly intake (PTWI), tolerable daily intake (TDI) or RfD (see Section 5.6). This approach makes no attempt to calculate a level of risk at low exposures. Rather, it derives a dose that is apparently without effect in a human population or suitable animal model, and then applies a factor to derive an exposure that has a high likelihood that no effect will occur in the general human population.

3.10.2 Non-threshold approaches

Non-threshold models assume linearity between the lowest experimentally derived dose and the zero dose (the origin). This implies there is a calculable probability of an adverse effect (risk) no matter how small the dose. This does not mean there is no dose that could be considered safe unless safety must be equated with zero risk (Hrudey & Krewski 1995).

Numerical estimates of risk probabilities are generated by fitting one or more mathematical models to the data in the experimental dose range and extrapolating the upper 95 per cent confidence limit of the curve fitting to the low environmental exposure doses. For example, low-dose extrapolation using a linear model is a default approach for cancer risk assessment in the United States (US EPA1996a; 2005a) and is one approach which has been used (perhaps inconsistently) by WHO for genotoxic carcinogens in deriving drinking-water guidelines (WHO 1993a; 2006a).

The revised US EPA guidelines for cancer risk assessment state that:

> *'A linear extrapolation approach is used when the mode of action information is supportive of linearity or mode of action is not understood.'*
>
> *'When adequate data on mode of action provide sufficient evidence to support a nonlinear mode of action for the general population and/or any subpopulations of concern, a different approach – a reference dose/ reference concentration that assumed nonlinearity – is used.'*
>
> *'When the mode of action information indicates that the dose–response function may be adequately described by both linear and nonlinear approach, then the results of both the linear and nonlinear analyses are presented.'*
>
> *'Absent data to the contrary, the default assumption is that the cumulative dose received over a lifetime, expressed as a lifetime average daily dose or lifetime average daily exposure, is an appropriate measure of dose or exposure.'*
> (US EPA 2005a pp. A8–A10)

Studies where the significant endpoint includes a neoplastic change (carcinogenesis) are usually assumed to represent a non-threshold dose–response relationship, especially where the chemical of concern has been shown to have genotoxic potential. For non-threshold dose–response relationships, the excess incidence (i.e. incidence corrected for background) of induced cancer is assumed to be zero only at zero exposure.

Low-dose linearity assumes a positive slope of the dose–response curve upward from zero dose and implies that a single, irreversible genetic event at the initiation stage of carcinogenesis leading to transformation of a cell is sufficient by itself to lead to the development of cancer. The major difficulty in this debate is the impossibility of experimentally testing the shape of the dose–response curve at extremely low doses (Purchase & Auton 1995).

A transformed cell that has acquired the potential to develop into a tumour will probably realise that potential only rarely (US EPA 1996a), most likely because of the natural large-scale repair of DNA damage and other defence mechanisms of the body (DOH 1991; Abelson 1994). Furthermore, while it is generally accepted that mutagens and mutations play a role in the development of cancer, carcinogenesis is more than mutagenesis, with a number of non-mutagenic as well as mutagenic events taking place during the process (see Section 11.4). The shape of the dose–response curve at any one of these steps, not just the mutagenic events, can influence the shape of the dose–response curve for the carcinogenic response. Factors, such as genetic make-up, lifestyle and other environmental factors may also have a modifying influence on the processes of carcinogenesis.

This can introduce some complexity into the quantitative aspects of the risk assessment. The incidence of cancer in unexposed or control animals is rarely zero, and finite values may even approach 100 per cent. More of a problem is that the background or spontaneous incidence can be quite variable from study to study, and influenced by such things as the age and breeding (strain) of the animals used, the experimental conditions (e.g. caging and feeding conditions and other aspects of animal husbandry), the rigour of the pathological investigations, the magnitude of the doses used, especially when escalated towards a dose level defined as the maximum tolerated dose (MTD), and even the way the test chemical is administered (gavage or via feed).

It follows that the numerical data derived from any one study may not be reproducible in another repeat study of the same design. It is also important to acknowledge this fact, since there is a general belief that numbers generated in a toxicity study are inviolable or sacrosanct as inputs into a risk assessment.

Furthermore, the extrapolation methodology used to estimate the disease incidence (or risk) at dose levels well below the high dose levels actually used in the study is quite model-dependent. The requisite knowledge or understanding of the toxicological mechanisms that account for nuances in the shape of the dose–response relationship at low dose may be incomplete.

The above uncertainties about the quantitative data derived from an animal testing program need to be borne in mind when extrapolating risk estimates within the same experimental species, let alone when attempting to extrapolate from one species to another (e.g. to humans). The resultant numerical risk estimates may need to be extensively qualified by reasonably large statistical confidence limits. Risk assessment may conventionally focus on the extreme ends of these confidence limits in order to be protective of more vulnerable members of the population. This can inject quite a degree of unfounded conservatism into a risk assessment, especially when 'worst-case' estimates for several different parameters are incorporated into the risk assessment methodology.

3.10.3 Outputs from a non-threshold risk assessment

The outcomes of a non-threshold risk assessment are either:

(i) the dose describing at which the chemicals produce a predetermined risk level – note that this requires some judgement on what constitutes an acceptable level of risk (what may constitute an 'acceptable' or 'target' risk estimate is discussed in Chapter 5); or

(ii) the estimated risk level at any particular dose.

The parameter from which (ii) is derived is the cancer slope factor (CSF) or unit risk factor (URF), which is the probability (or risk) of the response (e.g. cancer) per unit of intake (usually expressed in mg/kg body weight per day) or exposure concentration, respectively, over a lifetime of exposure.

It should be noted that the CSF is derived conservatively and is based on a linear extrapolation to zero dose from the upper-bound estimate of a dose at which the incidence of cancer can be predicted (e.g. a BMD – see Section 3.9). The URF is similarly derived using concentration as the exposure measure. The implication is that the relationship that best describes the low-dose behaviour of the dose–response curve is linear in this region, although it may become non-linear at higher doses.

Earlier versions of US EPA guidance on non-threshold dose–response extrapolation referred to the output parameter from the linearised multi-stage models as q1*, defined as:

> An upper bound, approximating a 95 per cent confidence limit, on the increased cancer risk for a lifetime exposure to an agent. This estimate, usually expressed in units of proportion (of a population) affected per mg/kg/ day, is generally reserved for use in the

low-dose region of the dose–response relationship, that is, for exposures corresponding to risks less than 1 in 100.

Dose selection in non-threshold models has been discussed by Lovell and Thomas (1996), who suggest that the estimate of q1* is so dependent on the doses selected that it is almost independent of, or at least insensitive to, the actual tumour incidences in the dose groups. Specifically, the highest dose in an animal bioassay has overwhelming influence on the estimate of q1*, thus leading to the overestimation of risk at very low doses, with the extent of overestimation increasing as the environmental exposure becomes lower. Typically, the highest dose in a carcinogenicity bioassay is the maximum tolerated dose (MTD), a dose that causes no more than a 10 per cent decrease in body weight and no other overt toxicity. The MTD is very much greater than doses expected from non-occupational environmental exposures. Therefore, the dose that is the least relevant to environmental risk assessment has the greatest influence on low-dose risk estimates.

An extremely insightful finding about the meaning of the q1* estimated in this manner was revealed by Krewski et al. (1993). This analysis considered the relationship between the q1* calculated using the linearised multi-stage model or equivalent no threshold model and the MTD, the maximum tolerated dose. Because carcinogen bioassays are very expensive experiments (costing millions of dollars if the full protocols are followed), which are limited in the number of animals and dose levels that can be tested, there is a practical need to maximise the potential for detecting a carcinogenic response in the typical 2-year duration of the experiment. This need has been met by doing a range-finding experiment to determine the maximum dose of carcinogen that the experimental animals can tolerate as a daily dose so that they can survive for the 2-year duration of the experiment to have an opportunity to contract cancer. For most carcinogens tested, the MTD is a very high dose, which may not be far removed from an acutely toxic dose.

Krewski et al. (1993) showed in an analysis of bioassay results for 191 individual carcinogens plotting the upper-bound estimate for the q1* versus the MTD that these values were highly negatively correlated (r = –0.941) for values that spanned nine decades in MTD and ten decades in CSF, a result that could not conceivably be achieved from 191 truly independent experiments.

This finding shows that carcinogens with a very high MTD (low toxicity) had a very shallow q1* and carcinogens with a very low MTD (high toxicity) had a very steep q1*. A primary determinant of the q1* for any of these chemicals was its MTD, given the procedure used for determining the CSF. Krewski et al. (1993) described how this outcome was created by the relatively small range of possible outcomes from the cancer bioassay once the dose, which according to the range of MTDs varied over a much larger range, was determined.

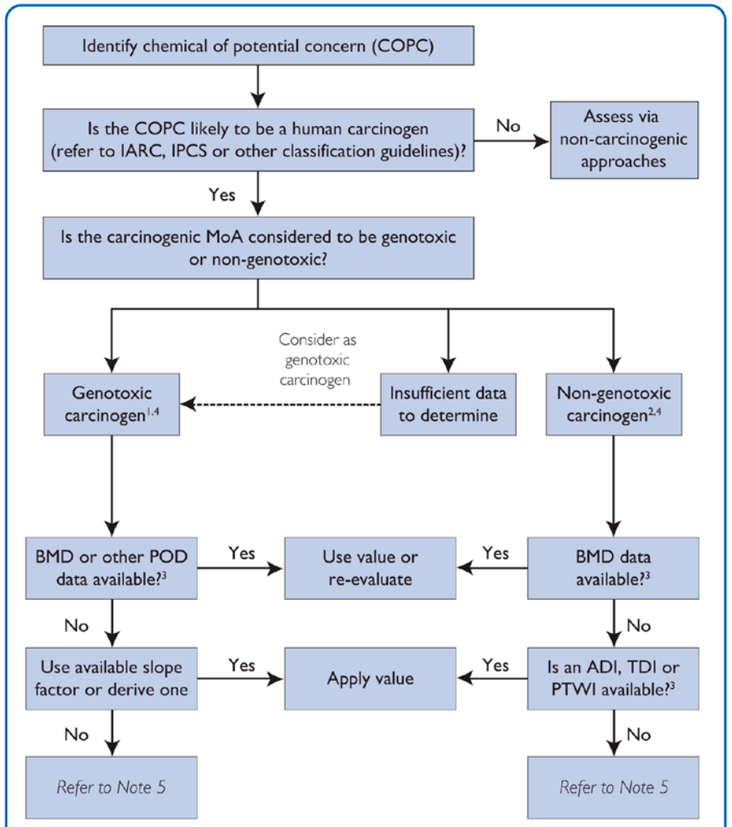
Non-threshold models currently in use are inflexible and generally do not take account of the complexities of the events between exposure to an agent and the induction of a neoplasm. Risks estimated at doses below the range of experimental data can vary considerably depending on the model used, even though the various mathematical models used generally fit the experimental data equally well (Crump 1985; Paustenbach 1995). The numerical expression of the estimated level of risk falsely gives the impression that it represents an exact measure of actual risk. This numerical expression provides little or no information on the uncertainties related to the estimated level of risk, nor does it allow comparison with values for non-cancer health effects.

It is notable that the latest US EPA guidelines on carcinogenic risk assessment (US EPA 2005a) make no reference to either CSF or q1*. The above definition is now simply linked to the term 'slope factor'.

Bodies such as WHO, US EPA and California EPA Office of Environmental Health Hazard Assessment have all developed CSF and URF for carcinogens. However, the values can vary by an order of magnitude depending on the studies used and the model chosen to derive the factor.

The step-wise process for deciding on the dose–response data to adopt for the EHRA of carcinogens (or potential carcinogens) is set out in Figure 15. This decision-making process recommends use of a BMD approach to selecting a POD for risk assessment once a decision has been made on classification of the COPC, as a carcinogen and a carcinogenic risk assessment approach is warranted. Where appropriate BMD data is not available, alternative dose–response data should be sourced, which may include the use of CSF (for genotoxic carcinogens) and ADI/TDI (for non-genotoxic carcinogens).

Figure 15: Decision-making process for choosing dose–response data in risk assessment of carcinogenic substances



Notes:

1. Should a non-threshold dose-response methodology be applied (i.e. derivation of a slope factor) an appropriate target risk level will need to be agreed. A target risk level of 10^{-5} is recommended in this enHealth guidance and in the contaminated sites NEPM
2. Where evidence suggests a threshold approach is appropriate, risk characterization may proceed using a BMD, or ADI/TDI approach. There will be no need to consider setting an acceptable or target level of cancer risk.
3. Refer to Chapter 5 for the source of preferred informations sources.
4. If a chemical has both genotoxic and non-genotoxic modes of action, results from both analyses should be presented.
5. Assessment not possible on the basis of presented data; consider whether acceptable data may be found through a more thorough search for relevant toxicological and/or epidemiological data and weight of evidence analysis.

3.11 THRESHOLD VERSUS NON-THRESHOLD APPROACHES – IMPLICATIONS FOR RISK ASSESSMENT

The important conceptual distinction between non-threshold methods and those that derive an acceptable exposure from the NOAEL using a safety or uncertainty factor is that this approach makes no attempt to determine a finite level of risk at low exposures, whereas non-threshold methods make an estimate of the risk at low exposures using (usually linear) extrapolation from a point higher in the dose–response relationship (see benchmark dose methodology in Section 3.9).

An NRC report (NRC 2008) acknowledges this important distinction, but also advances an argument that the two processes may be harmonised by redefining the RfD/RfC as risk-specific dose estimates describing the proportion of a population likely to be susceptible below the adjusted NOAEL.

The NOAEL is assumed to be the threshold dose for the effect. Both threshold and non-threshold approaches have advantages and disadvantages.

The advantages of the threshold approach are that the NOAEL is relatively easy to determine, and the process is simple to use, easy to understand and allows the use of expert judgement. In the few cases where epidemiological data has become available, the ADIs derived by this method have been validated (Lu & Sielken Jr 1991).

Additionally, the approach has been applied seemingly in a consistent fashion by WHO over three decades in deriving ADIs for pesticides. The safety-factor approach remained essentially unchanged until 1994 (WHO 1994a),

although a number of articles were published suggesting modifications or improvements (e.g. Calabrese & Baldwin 1994; Calabrese & Gilbert 1993; Crump 1984; Johannsen 1990; Lewis et al. 1990).

A potential disadvantage of relying on the NOAEL as the starting point for threshold-type risk assessment is the precision with which the real NOAEL can be estimated (see Section 5.6). Another possible limitation is the need to consider that a finite background level of disease may give an impression of non-linearity at low dose.

Additional limitations of the threshold approach include: the NOAEL is often perceived as a biological threshold, whereas it is a threshold limited by the experimental protocol; risk is expressed as a fraction of the guidance dose (e.g. ADI); it makes limited use of the dose–response slope; the choice of safety factors has been arbitrary to some extent; and the process does not generate a range of estimates of risk, but rather a single estimate of a dose below which no adverse effects are likely to be produced.

Because it provides numerical estimates of risk at all doses, the non-threshold approach, in principle, has the potential advantages (if the estimates are correct) of allowing computation of comparative risks in the sub-experimental range, which may be a useful tool in risk management and communication-potency comparisons between chemical agents at a particular risk level and estimates of the increased risks if a particular dose is exceeded. It has been argued (McMichael 1991) that risk estimates using this approach approximate those seen in humans in some cases and where there are disparities they are overestimates of the risks.

Both the threshold and non-threshold methods are likely to be unduly influenced by the selection of doses. The choice of the NOAEL is limited to one of the doses included in the experimental design. The biological no-effect dose may occur at this dose or possibly at a lower dose that is not included in the study. The closeness with which the selected NOAEL truly reflects the actual no-effect dose has an obvious impact on the degree of protectiveness in the derived ADI, PTWI or RfD. Furthermore, the precision with which the NOAEL can be assessed is influenced by the biological effects monitored, the number of animals in the test groups, the spontaneous incidence of the adverse effect, and the criteria used to determine when the incidence in a test group exceeds that in the controls (Renwick & Walker 1993).

3.12 SCIENCE POLICY AND THE SELECTION OF THRESHOLD AND NON-THRESHOLD MODELS

In deciding between a threshold or non-threshold approach to risk assessment, it is important to recognise that one is entering the realm of science policy. Such policy is commonly applied to genotoxic carcinogens by many regulatory agencies around the world, although it is being less rigorously applied in contemporary risk assessments to carcinogens for which there is reasonable evidence that they act through non-genotoxic mechanism(s) (NRC 2008).

The fact that a distinction may be made between a genotoxic and a non-genotoxic mechanism for a carcinogenic response will be based on the available evidence. However, it does not mean that a non-genotoxic carcinogen does not affect the genetic material of the cell under some circumstances or that a genotoxic effect is the only event required for the development of cancer by a genotoxic carcinogen.

With advances in biological knowledge, the classification and assessment of carcinogenic risk is now being guided by mechanistic, pharmacokinetic and other relevant data. The US EPA undertook a review of its science policy with regard to carcinogenic risk assessment (US EPA 2005a) and incorporated evaluation of MoA and its use in framing the risk assessment approach, as well as suggesting a more narrative approach to classify carcinogenicity. The use of MoA information and its impact on determining how to make low-dose extrapolation in EHRA was reviewed at a 'state-of-the-science workshop on issues and approaches in low-dose extrapolation' in April 2007 (White et al. (2009).

While the US EPA approach continues to rely almost exclusively on the non-threshold, low-dose extrapolation for cancer risk assessment as in the past, there is a growing acceptance that a threshold approach may be valid where the scientific data justifies an assumption of non-linearity at low dose (e.g. where cytotoxicity is a necessary precursor to the carcinogenic response).

It may be argued that the impetus for applying non-threshold methodology to carcinogenic risk assessment was the initial premise that all carcinogens are mutagens (Ames et al. 1973). One mutation or one DNA damage event was considered sufficient to initiate the process that leads to the development of cancer. A more contemporary hypothesis is that cancer formation requires a series of mutagenic events, perhaps in a defined sequence, not just a single event of DNA damage.

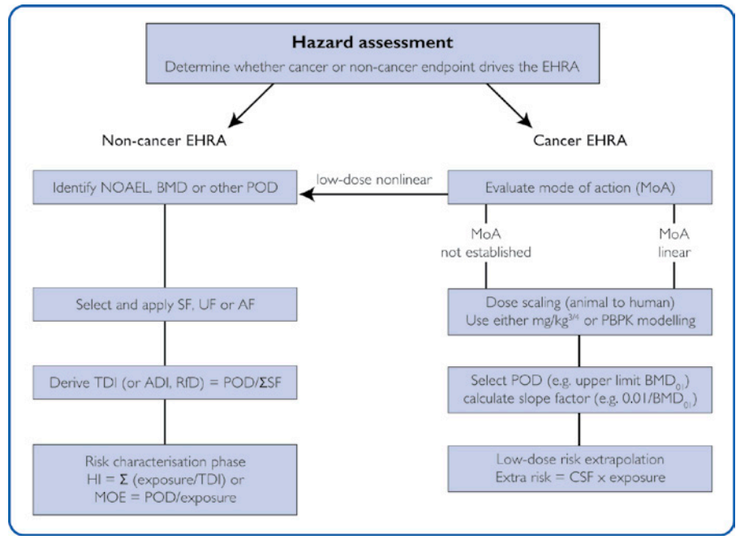
This premise that carcinogenicity and genotoxicity are inextricably linked has been questioned, even by Ames himself (Ames 1987), who has repeatedly pointed out that many animal carcinogens are not genotoxic, and that many naturally occurring processes (dietary elements and even oxygen itself) can produce a mutagenic yield in cellular DNA at theoretically much greater values than that associated with most environmental chemicals. It is also apparent that some chemicals can influence the quite complex multi-stage process of carcinogenesis by modifying events downstream from the initiating genetic defect (i.e. via epigenetic or promoter-type mechanisms), or even by simply amplifying the spontaneous development of cancers by increasing the rate of cell turnover, sometimes following a significant cell-damaging event.

The US National Research Council, working in conjunction with the US EPA, the National Institute of Environmental Health Sciences (NIEHS) and the National Academy of Sciences (NAS), has carefully enunciated the science policy decisions that underpin the use of threshold and non-threshold risk assessment approaches in the United States. The outcomes of this project were communicated through various consultations and reports, culminating in the release of the seminal report *Science and decisions: advancing risk assessment* in December 2008 (NRC 2008).

Figure 16 is a flow chart outlining the step-wise process whereby decisions can be taken about whether to use a threshold or non-threshold approach.

No Australian environmental health authorities have enunciated a specific policy on when a threshold or non-threshold approach should be used in EHRA. However, it is common practice, accepted by most state and territory jurisdictions, that a non-threshold approach consistent with that used by the US EPA should be used when assessing carcinogenic risk for genotoxic carcinogens. This implies that the endpoint of the EHRA is an estimate of risk that needs to be compared with an acceptable or 'target' risk level (see Section 5.10).

Figure 16: Decision tree for choosing a threshold or non-threshold model for risk assessment



NRC 2008. Reprinted with permission from the National Academies Press, Copyright 2009, National Academy of Sciences

When a non-threshold model is adopted the resulting prediction may be useful for demonstrating a plausible upper bound of cancer risk, but estimates by these methods are not intended to be used as estimates of "expected" cancer risk. In this regard, a quote from an award speech by Joe Rodricks, one of the creators of the U.S. policy approach for carcinogens is instructive (Rodricks 2007):

> The linearized multistage model was selected because it seemed to have some basis in the leading mechanistic hypotheses regarding the carcinogenic process, and also because it seemed highly likely that the model – because of its 'linearization' at low dose – would not underestimate low dose risk, that it would, in fact, place an upper bound on low-dose risk. Actual risk might be as large as the upper bound, but could be lower and could even be zero. It is not the case that risk assessors, at least those who truly understood the problem of low-dose extrapolation, have ever claimed that risks predicted in this fashion are known to be accurate, even ignoring the uncertainties introduced by the fact that most risk assessments are based on animal, not human data.

This enHealth document supports the application of sound scientific principles to assessing carcinogenic risk. These principles are articulated further in Chapter 5, and include consideration of carcinogenic MoA in determining whether a threshold or non-threshold approach is more suitable for the EHRA.

Chapter 4: Exposure assessment

4.1 INTRODUCTION

Exposure assessment requires a determination of the magnitude, frequency, extent, character and duration of exposures in the past, currently and in the future. There is also the identification of exposed populations and potential exposure pathways. Environmental monitoring and predictive models can be used to determine the levels of exposure at particular points on the exposure pathways. The contaminant intakes from the various pathways under a range of scenarios, including worst-case situations, can then be estimated (US EPA 1989).

Exposure assessment is one of the more critical and complex areas of risk assessment. Due to the complexity and scale of the EHRA process, a concise 'cookbook' on exposure assessment is not practicable. Similarly, the issues are often sufficiently complex and 'situation-specific' that a manageable and complete algorithm for decision making cannot be drafted. This chapter attempts to summarise useful guidance on exposure assessment to assist the decision-making process. Exposure assessment is identified as part of Phase II of the expanded framework for EHRA outlined in Figure 2.

Where possible, the information is prescriptive about certain aspects of exposure assessment. Having specific requirements for the content of investigations and having them presented in uniform, coherent and logically developed reports will enable more efficient, accurate, timely and transparent decision making and a greater consistency of environmental health decision making across Australia.

The aim is to provide:

- details on conducting appropriate exposure assessments
- a range of exposure factor data relevant to Australia
- where appropriate, default point estimates or, in some cases, probability distributions of exposure data for use in exposure assessments.

Basic elements to consider in planning an exposure assessment are:

- *Purpose:* the reason the study is being undertaken and how the results will be used.
- *Scope:* exactly what are the study areas, the population to be assessed, compounds and media to be measured.
- *Level of detail:* what level of accuracy is required in the estimate of the exposure for this to be meaningful, given the level of knowledge available about the toxicological links between exposure dose-effect and risk. What are the resource constraints and how can resources be most efficiently used?
- *Approach:* what methods will be used to determine exposure and do these accurately represent pathways of exposure that will affect risk. What is the nature of sample collection? (e.g. How many are needed? From where? How frequently?) How will the data be handled, analysed and interpreted (US EPA 1992).

This enHealth document, along with the companion *Australian exposure factor guidance* document aims to assist with this process by collating and tabulating data that could be useful to estimate exposures in an EHRA, and to fill any knowledge gaps where default assumptions may need to be made. Using data from the *Australian exposure factor guidance* document is discussed further in Sections 4.13 and 4.14. However, it is emphasised that inputting data based on valid measurements of parameters describing the exposure scenarios under consideration are always preferable to using the default assumptions that are a common feature of Tier 1 assessments.

The information in this chapter, and in the complementary *Australian exposure factor guidance* document, includes material published in some key US and IPCS guidance documents on exposure assessment, including:

- *Public health assessment guidance manual* (1992), United States Agency for Toxic Substances and Disease Registry (ATSDR)
- *Guidelines for exposure assessment* (1992), US EPA
- *EHC210 Principles for the assessment of risks to human health from exposure to chemicals* (WHO 1999b)
- *EHC235 Dermal absorption* (WHO 2006c)
- *EHC237 Principles for evaluating health risk in children associated with exposure to chemicals* (WHO 2006d)
- *Exposure factors handbooks* (2009b), US EPA.

It also includes exposure factor information presented in the proceedings of the five National Workshops for the Health Risk Assessment and Management of Contaminated Sites, data developed from research by the South Australian Department of Human Services and data sources from the international literature.

4.2 TERMINOLOGY USED IN EXPOSURE ASSESSMENT

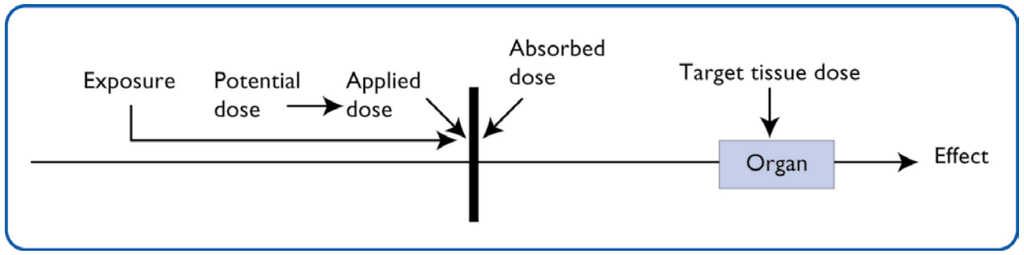
The terminology used to define exposure and the factors that can influence the extent of exposure by various pathways are explained in Table 6.

Table 6: Explanations of terms used in assessing dose and exposure

Term	Explanation
Bioavailability	A generic term defined as the fraction of a contaminant that is absorbed into the body following dermal contact, ingestion or inhalation. It is expressed as the ratio (or percentage) of the absorbed dose (systemic dose) to the administered dose.
Absolute bioavailability	The mass of a contaminant that is absorbed and reaches systemic circulation following dermal contact, ingestion or inhalation.
Relative bioavailability	The comparative bioavailability of different forms of a chemical or for different exposure media containing the chemical expressed as a fractional relative absorption factor. In the context of environmental risk assessment, relative bioavailability is the ratio of the absorbed fraction from the exposure medium in the risk assessment (e.g. soil) to the absorbed fraction from the dosing medium used in the critical toxicity study.
Bioaccessibility	Generically, it is the ability for a chemical to come into contact with the absorbing surfaces in an organism. It is related to solubility and dissolution, since absorption usually can only occur from a liquid or gaseous phase, and not from a solid phase. It is defined as the fraction of a contaminant in soil that is soluble in the relevant physiological milieu (usually the gastrointestinal tract) and available for absorption. This can be assessed by validated *in vitro* test systems. There are only a few such test systems and these have been found to be applicable to only a limited number of contaminants. In conjunction with bioavailability, it can be a significant factor determining the amount of a substance that might be absorbed from soil at a contaminated site.
Exposure	Concentration or amount of a particular chemical that reaches a target organism, or system or (sub)population in a specific frequency for a defined duration. Exposure is usually quantified as the concentration of the agent in the medium integrated over the time duration of contact.
Exposure concentration	The exposure mass divided by the contact volume or the exposure mass divided by the mass of contact volume, depending on the medium.
Exposure duration	The length of time over which continuous or intermittent contacts occur between a chemical and the exposed population.
Exposure event	The occurrence of continuous contact between chemical and exposed population.
Exposure frequency	The number of exposure events within an exposure duration.
Exposure route or pathway	The way a chemical enters an organism after contact (e.g. by ingestion, inhalation or dermal absorption). The pathway usually describes the course a chemical or physical agent takes from a source to an exposed organism. An exposure pathway describes a unique mechanism by which an individual or population is exposed to chemicals or physical agents at or originating from a site. Each exposure pathway includes a source or release from a source, an exposure point, and an exposure route. If the exposure point differs from the source, a transport/exposure medium (e.g. air) or media (in cases of inter-media transfer) is also indicated.
Exposed population	The people who may be exposed to the contaminant. Synonymous with 'receptor'.
Dose	The amount of a substance available for interaction with metabolic processes or biologically significant receptors after crossing the outer boundary of an organism.
Potential dose	Amount of a chemical contained in material ingested, air breathed or bulk material applied to skin.
Applied dose	Amount of chemical in contact with the primary absorption boundaries (e.g. skin, lungs, gastrointestinal tract) and available for absorption.
Internal (absorbed) dose	The amount of a chemical penetrating across an absorption barrier or exchange boundary via either physical or biological processes.
Target tissue (biologically effective) dose	Amount of chemical available for interaction with any particular organ or cell.

The terminology relating to dose in Table 6 is represented graphically in Figure 17.

Figure 17: Representation of dose and exposure



Adapted from: US EPA 1992.

It is normal practice in Australia to estimate the exposure to a chemical or other agent that people are likely to receive for a specific situation, process or facility. This is then compared with the dose likely to not cause an effect (estimated in the effect or toxicity assessment phase) in the risk characterisation step to determine if the situation or facility poses an unacceptable risk to people. This is likely to be a conservative way to address the assessment of risk given that it is likely to overestimate the target tissue dose for the actual organ or cell where effects will occur.

4.2.1 Significance of bioavailability and bioaccessibility

Many substances are able to tightly bind to environmental matrices such as soil or sediment. In most situations, there will be little or no information about the bioavailability of a contaminant in the situation under investigation. It is, therefore, normal practice to assume 100 per cent bioavailability. If reliable information is available it can be used to justify the use of a value other than 100 per cent. The companion *Australian exposure factors guidance* document includes more extensive guidance on where specific bioavailability data may be available.

The bioavailability of the substance from the media consists of two major processes, bioavailability and bioaccessibility. Together these terms represent the amount of a substance that may reach the systemic circulation of a human 'receptor' following exposure to an environmental contaminant. Bioaccessibility refers to how much of the chemical dissolves into bodily fluids, such as inside the stomach, enabling the chemical to reach and cross biological membranes and enter the circulatory system. Bioavailability then refers to how much of the dissolved chemical can cross the absorption barriers.

Essentially: bioavailability = bioaccessibility × absorption

The significance of understanding bioavailability and bioaccessibility in assessing exposure is discussed in more detail in the *Australian exposure factors guidance* handbook, with chemical-specific data summarised where it is available. Much of the following text is extracted from Sections 3.4 and 4.1 of that companion document, and definitions of the various terms describing bioavailability and bioaccessibility are defined in Table 6.

Oral exposure is commonly the main route for entry of chemicals into the body. Given its importance, in the absence of specific bioavailability data, it has been common practice to conservatively assume the oral bioavailability of a chemical from environmental media will be at least the same as the bioavailability of the chemical in toxicity experiments underlying the derivation of the guideline value (GV) (i.e. a relative bioavailability of 100 per cent) (UK EA 2009, US EPA 2007a).

Oral bioavailability relates to the fraction of an orally administered dose of chemical that reaches the systemic circulation (RIVM 2009).

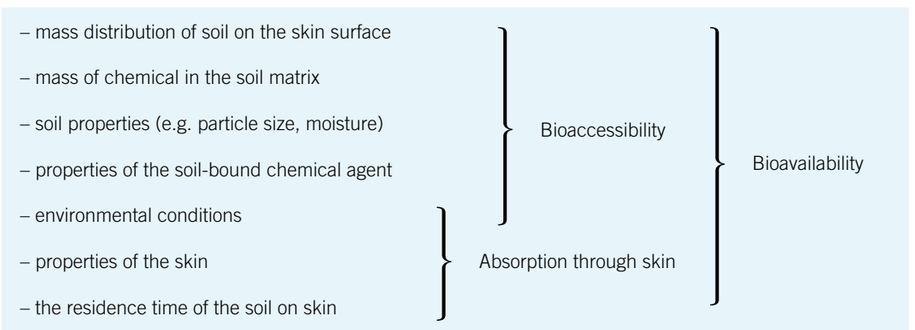
The term 'relative bioavailability' refers to a comparison of absolute bioavailabilities. Therefore, it is the ratio of the bioavailability of a substance in one exposure context (i.e. physical chemical matrix or physical chemical form of the substances) to that in another exposure context (commonly an administered dose in an experimental animal study).

Estimates of bioaccessibility and overall bioavailability (i.e. bioaccessibility plus absorption) can be determined from experimental studies – *in vitro* systems mimicking biological conditions for bioaccessibility estimates and *in vivo* (whole animal) models for bioavailability.

The processes of bioaccessibility and absorption affect the bioavailability of all chemicals from environmental media, but are of special importance for metals. This is because metals can exist in a variety of chemical and physical forms, and different forms of a given metal can be absorbed to a different extent. For example, a metal in contaminated soil may be absorbed to a greater or lesser extent (but generally somewhat lesser) than when ingested in drinking water (US EPA 2007a).

The overall dermal bioavailability from soil should be chosen on a chemical-specific basis following a review of the scientific literature. The US EPA has published default dermal bioavailability factors for 23 chemicals (US EPA 2007a). Additional dermal bioavailability and bioaccessibility data is detailed in Section 3.4 of the *Australian exposure factors guidance* document. Toxicity studies generally use applied dose rather than actual measures of the absorbed dose and a default bioavailability of 100 per cent was assumed for oral studies of highly water-soluble substances. The NRC (2003) prefers to use the term 'bioavailability processes' to encapsulate the mechanisms involved in the dissolution, transport and absorption of environmental contaminants by a receptor organism.

Some factors that influence bioavailability are listed below.



4.3 PLANNING AN EXPOSURE ASSESSMENT

4.3.1 identifying release of a chemical or agent to the environment

A chemical or other agent will be released to the environment from a facility, situation or process in a variety of ways. The first step in planning an exposure assessment is to work out how the chemical or agent gets into the environment.

As discussed in Section 1.2 examples of when EHRAs may be undertaken include:

- new additives to food or potable or recreational waters
- new and existing chemicals assessment
- contaminated sites assessment
- assessment of major planning developments
- assessment of hazardous developments
- pollution impact assessment at existing facilities
- changes to climate, landform, geography or demography that may impact on disease vectors and parasites
- situations where environmental standards or guidelines are unavailable
- environmental changes that will increase traffic flow and may increase the risk of injury or air pollution, such as new traffic corridors
- changes that may impact on the microbiological or chemical safety of food chains and food supplies
- situations where there is a high level of public interest in or concern about environmental health issues
- situations where vulnerable populations may be affected by environmental health issues such as the placement of schools
- legislative or policy changes
- designating housing setbacks from industry and transport corridors
- where health impact assessments are undertaken.

In all these situations, chemicals may be released directly into food, air, water, soil or waste. They may also indirectly or accidentally be released to these media.

In each of these situations, consideration of how the chemical or other agent is used and how it may be released to the environment is required. How often, how much, what happens in unusual circumstances like plant malfunctions, and how variable the release rate may be, all must be considered.

4.3.2 Identifying fate and transport of a chemical or agent

Once a chemical or other agent reaches the environment, its fate and transport needs to be considered.

Transport away from the release point is fairly obvious. If released to air, then the air containing that chemical or other agent may be blown away and diluted. If released to water in a river, then the chemical may flow away and be diluted quite a distance while release to a lake or tiny creek may stay close to the release point for an extended period. If released to groundwater then it may not flow far. If released to soil or waste, then it may be trucked away from the original source and reach a whole new location.

Fate of a chemical or other agent describes the reactions it may undergo once released. A chemical can react with other chemicals like humic acids, be broken down by chemical processes like hydrolysis or photolysis, be broken down by micro-organisms in the environment, or persist for an extended period. What actually occurs in a specific situation will depend on the characteristics of the environment into which it is being released.

4.3.3 Identifying exposure pathways

Exposure pathways are those processes that take a chemical or other agent from its point of release to the environment through to a situation where a person can be exposed. So it may be direct exposure because a chemical may be added to food which is then consumed. It may be indirect exposure where a chemical is released to the environment and people are exposed at some temporal or geographic distance from the initial release point via an exposure pathway consisting of more than one step. Often the exposure pathways will be fairly obvious, but there may be a range of situations such as the movement of contaminated groundwater or volatile chemicals from contaminated groundwater that are less obvious. These less obvious pathways also need to be evaluated.

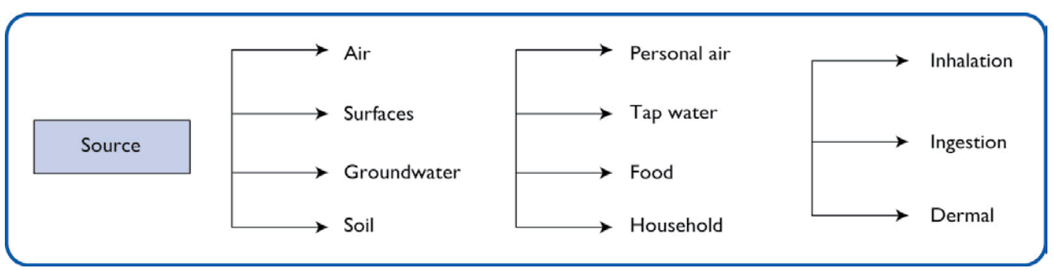
A summary of exposure pathways or routes and their interrelationships is shown in Figure 18. Interventions to reduce exposure may be made at a single point or multiple points.

A fundamental concept of risk assessment is that there must be plausible evidence of an exposure pathway linking the source of contamination and the exposed population. Where this linkage exists, an assessment of the nature and significance of the exposure pathway is required to determine the level of risk and the extent to which the proposed pathway is 'complete'.

A pathway may be considered 'complete' where there is documented evidence of a source (i.e. presence of one or more COPC at the site under consideration for the EHRA); there is evidence that the COPCs are actually released from their sources; that there are transport pathways and mechanisms that could convey the COPCs from source to the various sites where 'receptors' are located; and that the potential for human contact has been established for contaminated environmental media (air, water, soil food, surfaces) at each point in the transport chain.

The development of a conceptual site model (CSM) may be quite useful in identifying and quantifying the exposure pathways.

Figure 18: Exposure pathways



Adapted from: McKone 1993.

4.3.4 Identifying potentially exposed populations

To potentially pose a risk it is necessary to have an exposed population for the agents of concern. Part of the information to be collected in developing a conceptual site model is whether people may live or interact with the agent of concern or with media contaminated with the agent of concern. Often this pathway is quite obvious but the possibility of less obvious exposure pathways should not be overlooked.

There are situations, however, where it can be quite difficult to establish whether people might be exposed. When the behaviour is rare or limited in extent in the general population establishing whether there are actually likely to be people who meet the criteria can be difficult. For example, determining whether a particular location is actually a spot where people like to go fishing and so could be exposed via consumption of the fish they catch.

4.4 DEVELOPMENT OF A CONCEPTUAL SITE MODEL

Once all the information discussed above is collected, it should be collated together to enable the development of a conceptual site model.

4.4.1 Conceptual site models

Developing a conceptual site model (CSM) can materially assist the process of understanding how human 'receptors' may be exposed to chemicals from relevant environmental sources. The CSM describes the sources of contamination, the pathways by which contaminants may migrate through the various environmental media and the populations (human or ecological) that may potentially be exposed (NEPC 2010).

CSMs are site- and scenario-specific and describe the pathways by which chemicals transfer from environmental sources (e.g. soil, groundwater, airborne emissions from an industrial facility). A CSM may be based on diagrams or flow charts, but it must be supplemented with detailed information on COPC concentrations in various media, transfer characteristics, and receptor characterisation.

A CSM is generally a written description of the site that is accompanied by a schematic, graphical interpretation that depicts what is known or has been inferred about the site. CSMs are an important tool for visualising the pathways by which human exposure to chemicals from a variety of environmental sources may occur. The CSM usually includes diagrammatic representations of elements of the pathways between source and receptors, but it must always include more detailed textual descriptions of the characteristics of each source, pathway and receptor.

It is important to be clear whether an exposure pathway is 'completed' (i.e. there is reasonable evidence that there are 'receptors' who would actually be exposed in the given scenario) or whether the exposure pathway is 'potential' (i.e. it may need to be considered in a holistic EHRA), but where the contribution to overall exposure may be so slight or limited that it would have little impact on the risk estimate.

A detailed conceptual site model should include information on the following:

- the contaminants – concentration, distribution and media in which they occur (soil, water, sediment or air)
- physical characteristics of the environment for contaminated sites – including soil type, porosity, potential preferential pathways, vadose zone thickness, groundwater gradient and velocity, and hydraulic conductivity of the saturated zone
- characteristics of the exposed populations – exposed populations may be humans residing or working at the site or adjacent areas, future occupiers of the site after redevelopment, or environmental populations such as ecosystems in receiving environments such as natural surface waters.

CSMs are particularly important in EHRA of contaminated sites but they are also useful for elaborating exposure pathways associated with airborne pollutants from industrial sites.

4.4.2 Examples of CSM diagrammatic representations

There are many ways of depicting a CSM using diagrams and flow charts. Two examples are shown in Figures 19 and 20.

Figure 19: Representation of a CSM flow chart for potential airborne exposures from an industrial site

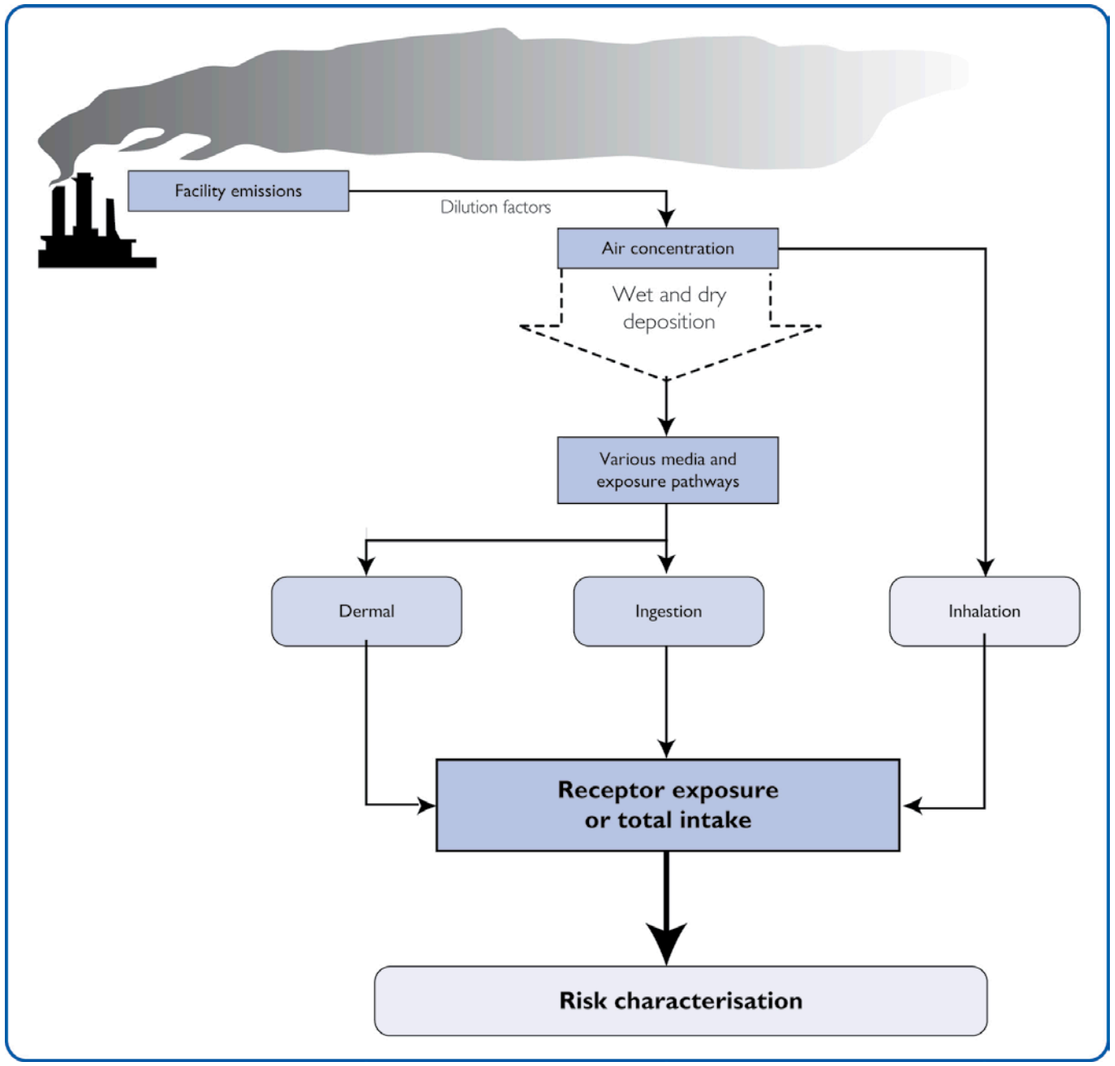
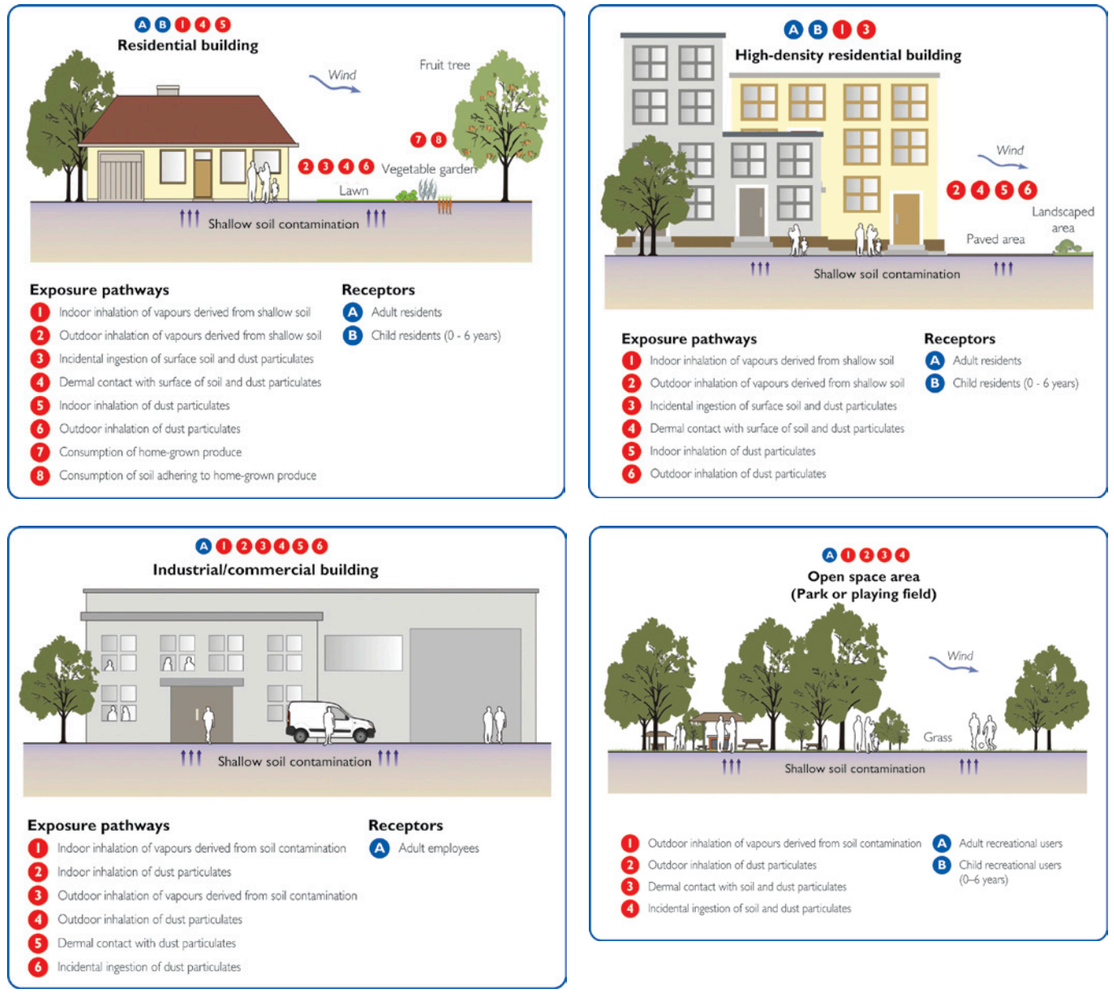


Figure 20: CSM site diagrams representing potential exposure pathways from a contaminated site



Reproduced with permission from the contaminated sites NEPM Schedule B(7) (NEPC 2010).

4.5 APPROACHES TO QUANTIFYING EXPOSURES

An initial requirement for exposure assessment is an understanding of the presence (or absence) of an agent and its concentrations and distribution, including any fluctuations over time. Guidance on sampling and analysis of environmental media is summarised in Chapter 8.

Accurate and useful exposure assessment requires a detailed understanding of both the strengths and weaknesses of the exposure assessment techniques, and the specific exposure factors used in the assessment. Considerable effort needs to be made to accurately characterise the population or individuals for whom the exposure assessment is relevant.

> *Direct measurement of the exposures of the (potentially) affected population provides the best exposure data but this is not always available or practicable and default exposure factor data are often required.*
>
> (Langley 1993a p. 90)

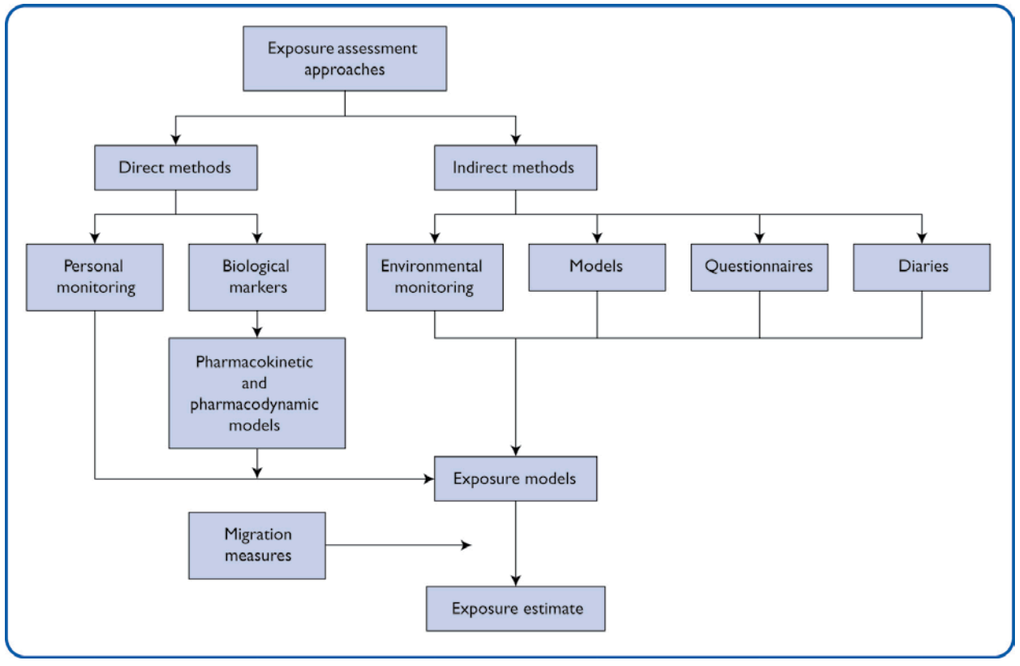
Figure 21 outlines the integration of direct and indirect measurements of exposure. Most EHRA processes rely on indirect estimation of exposure, using environmental monitoring data and models to quantify chemical transport through the identified exposure pathways. Chapter 13 outlines further information on the use of modelling in exposure assessment.

4.5.1 Measurement of exposure

Accurate and useful exposure assessment requires a detailed understanding both of the strengths and weaknesses of the exposure assessment techniques, and the specific exposure factors used in the assessment.

Direct measurement of the exposures of the (potentially) affected population provides the best exposure data, but this is not always available or practicable (except perhaps at the Tier 3 level) and default exposure factor data is often required.

Figure 21: Components of exposure assessment



Adapted from: National Academy of Sciences (NAS) 1991.

4.5.2 Determinants of exposures

The principal determinants of the level of exposure are:

- concentration of the agent in the relevant medium
- exposure duration
- exposure frequency
- exposure fluctuations (continuous or intermittent)
- whether exposure pathways are completed or potential.

Monitoring environmental media can provide useful indications of exposure levels, providing there is a strong correlation between environmental media levels of exposure and personal doses.

A powerful method of direct exposure assessment is using biomarkers as part of a biological monitoring process, but this approach has its limitations. The role of biomonitoring is discussed in Chapter 14.

4.5.3 Quantification of exposure

The quantification of exposure can be done in three ways:

- The exposure can be measured at the point of contact (the outer boundary of the body) while it is taking place, measuring both exposure concentration and time of contact and integrating them (point of contact measurement).
- The exposure can be estimated by separately evaluating the exposure concentration and the time of contact, then combining this information (scenario evaluation).
- The exposure can be estimated by dose, which in turn can be reconstructed from internal indicators (biomarkers, body burden, excretion levels) after the exposure has taken place.

Each of these methods is a separate entity using different information, and so each one can be useful in verifying or validating the results of other methods (ATSDR 1992).

Commonly, chemical levels will be measured at the point of release to the environment, as this is likely to be the point where concentrations are highest and so it will be the easiest to measure. Such data may also be available from monitoring required by regulation. More information is provided in Chapter 8.

It is important to have a good understanding of the strengths and weaknesses of the sampling design and the analytical methods used in any measurement.

An understanding of transport and fate models for the agent(s) in question is also important. Transport and fate will be affected by (Fiksel & Scow 1983):

- environmental exposure medium (e.g. air, surface water, soil, groundwater or biota)
- geographic scale (e.g. global, national, regional or local)
- pollutant source characteristics (e.g. continuous, intermittent or instantaneous releases from industrial, residential and commercial point or area sources)
- the nature of the risk agent (e.g. whether it is a specific agent or group of agents)
- the receptor population (e.g. humans, animals, plants, micro-organisms and habitats, as well as specific sub-populations exposed to high levels of the agent or who are particularly sensitive to exposure)
- exposure routes (e.g. ingestion, dermal contact or inhalation)
- environmental conditions (e.g. pH, presence of organic matter, clay content, temperature and meteorological)
- the time frame (e.g. retrospective, current or prospective).

Modelling may be used to estimate the concentration that people may be exposed to when measurement is not practical or possible in the time frame required. See Chapter 13

The initial release of a chemical may be modelled for facilities that are yet to be built. These will be based on the engineering of the facility and the way chemicals are to be used.

Measured data for the release of a chemical from a facility may be available but models may commonly be used to describe the transport of the chemical away from the point of release, such as air dispersion models, or to describe the fate of the chemical in the environment through consideration of half-lives, effect of organic carbon or other characteristics.

In developing sampling plans for chemical agents and assessing exposure, an understanding of the movement of chemical agents within and between environmental compartments and the effects of environmental partitioning will be necessary (see Section 4.9.1).

4.6 EXPOSURE ASSESSMENT CALCULATIONS

A generic formula for calculating chemical intake from various media has been formulated and described in US RAGS-A guidance on risk assessment (US EPA 1989). The generic formula is:

$$I = \frac{C \times CR \times EFD}{BW} \times \frac{1 \times CF}{AT}$$

I = intake of chemical (usually expressed as mg/kg bw/day)

C = average chemical concentration in media over the exposure period (e.g. mg/L, mg/kg or mg/m_3)

CR = contact rate the amount of contaminated media contacted per unit time or event (e.g. L/day)

EFD = exposure frequency and duration (how long and how often exposure occurs)

EFD may be based on the product of two parameters

EF exposure frequency (e.g. days/year) and ED exposure duration (e.g. years)

BW = body weight, usually averaged over the exposure period (e.g. kg)

AT = averaging time period over which the exposure is averaged (e.g. hours, days, months, years)

CF = conversion factor, if units in above parameters don't match

The first term in the above equation calculates an intake based on frequency and duration of exposure, adjusted for body weight, while the second term adjusts the intake on the basis of the selected averaging time, and incorporates an adjustment where units in the individual terms do not match.

Where specific routes of exposure are considered, this basic equation may be modified to incorporate route-specific information. Examples are:

For oral ingestion:

$$I\ (mg/kg/d) = \frac{C\ (mg/kg) \times AoF \times IGR\ (mg/d) \times EF\ (d/yr) \times ED\ (yr) \times CF\ (10^{-6})}{365\ (days\ in\ a\ year) \times AT\ (yr) \times BW\ (kg)}$$

In this case, AoF is an oral absorption factor or bioavailability estimate (unitless) and ingestion rate (IGR) of the medium (e.g. food, soil, water) replaces CR. For water intake, the units for C and IGR are expressed in mg/L and L/d respectively.

For inhalation of volatiles:

$$I = \frac{C\ (mg/m^3) \times IR\ (m^3/h) \times LR \times ET\ (h/d) \times EF\ (d/yr)}{365 \times AT\ (yr) \times BW\ (kg)}$$

IR = inhalation rate; LR is a lung retention factor (unitless); ET = exposure time

Note that this equation is based on the US EPA RAGS-A approach (US EPA 1989). The most recent US EPA guidance (RAGS-F) proposes that there is no need to calculate an intake rate based on concentration, inhalation rate and lung retention. RAGS-F recommends calculating a modified exposure concentration (EC), which is then compared directly with the RfC derived from toxicological studies.

Using the RAGS-F approach (which is the recommendation of this updated enHealth document), the equation converts to:

$$\text{Exposure concentration (EC)} = \frac{C\ (mg/m^3) \times ET\ (h/d) \times EF\ (d/yr)}{AT\ (yr)}$$

For inhalation of dusts $$I = \frac{C\ (mg/kg) \times IR\ (m^3/h) \times ET\ (hr/d) \times EF\ (d/yr)}{365 \times AT\ (yr) \times BW\ (kg) \times PEF\ (m^3/kg)}$$

PEF = particle emission factor

For dermal contact with soils

$$I = \frac{C\ (mg/kg) \times AH\ (mg/cm^2/d) \times SA\ (cm^2) \times AF \times EF\ (d/yr) \times ED\ (yr) \times CF\ (10^{-6})}{365 \times AT\ (yr) \times BW\ (kg)}$$

AH = soil adherence; SA = surface areas of skin exposed; AF = skin absorption factor

For dermal contact with water

$$I = \frac{DA_{event}\ (mg/cm^2/event) \times SA\ (cm^2) \times EV\ (events/d)\ EF\ (d/yr) \times ED\ (yr)}{365 \times AT\ (yr) \times BW\ (kg)}$$

DA_{event} = dose absorbed per event; EV = event frequency (events/d); EF = exposure frequency (d/yr)

The above equation is the first of a series provided in RAGS-E guidance from the US EPA (US EPA 2004b; Eq 3.1) for chemical intake via water exposure. Equations are also provided for DAevent, the choice of equation depends on the relationship of the exposure time (ET, in the DAevent equations) with the time to reach steady-state water concentrations. In estimating intake of organic chemicals from contact with water, the risk assessor may also wish to incorporate consideration of bioavailability (skin absorption) (Eq 3.8, US EPA 2004b). Reference values for many of the equations parameters are available for a large range of compounds in Appendix B, US EPA (2004b); it should be noted however that many are calculated rather than being empirically derived, the latter are usually preferable providing the experimental conditions are appropriate.

Volatilisation of organic compounds from water may need to be taken into account. Blando and Cohn (2004) describe an approach using equations defined by the US EPA (1994).

The averaging time (AT) will depend on the nature of the adverse effect (e.g. acute or chronic) being assessed and the exposure scenario (e.g. short periods or long term) predicted from the issues identification phase of the risk assessment or the relevant conceptual site model.

- For non-threshold adverse effects (e.g. genotoxic carcinogenesis) and threshold effects where exposure may be assumed to be over a lifetime (e.g. via food or drinking water) the AT value most commonly used is 70 years. It should be noted that where the exposure is less than lifetime, this calculation may underestimate risk if the effect does not depend upon lifetime exposure.
- For adverse effects considered to exhibit a threshold, the AT depends on the nature of the toxicity data and the assumed period of exposure. Where a shorter period better reflects the expected exposure paradigm, a shorter averaging time may be used. For example, where exposure at a domestic dwelling may be from contaminated soil, vapour from contaminated groundwater or emissions from a local industry, an AT value of 30 years (the average period of residency) may be considered suitable (see Section 7.1 of the *Australian exposure factor guidance* handbook). Similarly, ATs for occupational exposures may be restricted to the anticipated time and frequency of exposure, or assumed working lifetime in a particular industry. For risk assessments in which exposure may be for a particular life stage, the appropriate AT is the time span of the life stage; for example, for young children from birth to the sixth birthday the AT is 6 years (US EPA 1989).

4.7 DATA ANALYSIS AND EVALUATION

4.7.1 Nature of the exposure assessment

Risk models for carcinogens, in particular that of the US EPA, use lifetime time-weighted average doses to determine the dose–response relationships. This is based on using data from lifetime animal studies to determine dose–responses and hence slope factors. Consideration must be given to the nature of the exposures that occur, as in many instances exposures may be episodic or quite variable. In some instances, an exposure at a critical period may be of more concern than average exposure over a long period.

If exposure ends, this must be taken into account in the exposure assessment, as it may mean risk ceases (e.g. the risk of trauma when hazardous machinery is fenced) or decreases (e.g. the decreasing risk of lung cancer related to the period since smoking ceased). In these situations, account must be taken of the influence of these episodic exposures rather than using a lifetime average exposure.

4.7.2 Assessing past exposures

There are often considerable difficulties in assessing historical exposures. It may be possible if there have not been disturbances of the environmental media and the substance is relatively inert (e.g. the amount of lead in soil is unlikely to change). However, in a situation where there has been disturbance of the environmental media (e.g. soil movements) or changes to physico-chemical characteristics of the substance it may be difficult or impossible to accurately assess past exposures (e.g. to volatile hydrocarbons in the surface stratum of soil). It may only be possible to make some crude classification of exposure as 'high', 'medium' or 'low'. A 'crude dichotomy' of exposure may be all that is possible (e.g. 'exposed' versus 'not exposed').

Errors in estimating historical exposures may have significant consequences for the accuracy of the risk characterisations. If the exposure estimates are being used in epidemiological studies it may mean that it is impossible to determine whether there is a dose–response relationship or an accurate level of association between the level of exposure and a health effect (US EPA 1992).

US EPA (1992) provides several references detailing approaches for determining and estimating past exposures.

4.7.3 Dealing with data gaps

Constraints in time resources and money always lead to data gaps. Ways of addressing these data gaps include (US EPA 1992):

- collecting new data – this is the preferred option, although it must be recognised that the time taken to collect new data may compromise getting a resolution of the problem if the need for a risk assessment is urgent
- using models to estimate exposure values
- inserting conservative assumptions – however, the assessor should be aware of the flow-on consequences of utilising conservative assumptions, particularly a series of such assumptions
- using professional judgement – although this should depend on extensive experience rather than

anecdotal information and the assessor must account for such judgements in the uncertainty analysis

- narrowing the scope of assessment if the data gaps appear to be in one pathway or exposure route – this is likely to be the least satisfactory option if such a route or pathway is important, or if there is insufficient knowledge to determine the significance of the pathway.

4.7.4 Dealing with censored data

Heyworth (1991) provides a summary of the three essential methods for dealing with censored data:

1. *Simple substitution methods:* Simple substitution methods refer to those methods that substitute a single value, such as one-half the detection limit for each censored value. While these methods are commonly used they have no theoretical basis. The choice of the substitution value is essentially arbitrary and the estimates of summary statistics will be biased by these fabricated results.

2. *Distribution methods:* The distribution method uses the characteristics of the assumed distribution of the data. For environmental monitoring the log-normal distribution is usually assumed and values of data above and below the reporting limit are assumed to follow this distribution.

   Estimates of the mean and standard deviation are computed using the best match from the observed data and percentage that fall below the limit. Estimation methods include maximum likelihood estimation and probability plotting procedures. These methods will produce unbiased estimates only when the observed data fits the distribution exactly and the sample size is large. This, of course, is a rare case. However, they provide better estimates than those obtained by simple substitution.

3. *Robust methods:* The robust method combines the observed data above the detection limit with extrapolated below-limit values to compute summary statistics. In contrast to the distribution method, the actual data above the reporting limit is used to fit a distribution rather than assuming a distribution.

   This method has the advantage that estimates of extrapolated values can be directly retransformed and summary statistics computed in the original units, thereby avoiding transformation bias. Also, this method is not as sensitive to the fit of the distribution for the largest observations because actual observed data is used to fit the distribution.

   The probability plotting method used to fit the distribution in robust methods can be computed quite readily by most commercially available statistical packages. The US EPA website includes ProUCL 4.0 software for managing data that contains data requiring censoring along with technical support documents (Singh et al. 2007). See <http://www.epa.gov/esd/tsc/software.htm>.

Censoring data to manage data gaps or to replace analytical data that is below the limit of reporting is discussed in Section 8.7.

4.8 INDIVIDUAL EXPOSURE, DOSE AND RISK ESTIMATES

Several aspects need to be addressed in ascertaining estimates appropriate to an individual in a population:

- To what extent does the dose–response relationship take into account the normal variability in a population? If a highly susceptible sub-group can be identified, is this incorporated into the dose–response relationship?
- Has the appropriate dose for use in the dose–risk relationship been identified? It must be ascertained whether reference data applies to absorbed/internal/potential/applied/effective doses and the relevant data is used appropriately in the model. This may require allowances for variations in bioavailability. Bioavailability data from animal studies should not necessarily be considered to be similar for humans.

Given that these two factors have been adequately resolved, estimates of exposure, dose and risk can then be undertaken for individuals or narrow sub-populations. Often the high-end risk will need to be estimated because this will often be a driving force for any risk management measures.

The US EPA points out 'the high end segments of the exposure, dose and risk populations may represent different individuals' (US EPA 1992 p. 22921). This is due to variations in bioavailability absorption, intake rates, susceptibility and other variables between the segments of the population; that is, 'a high exposure does not necessarily result in a high dose or risk, although logically one would expect a moderately to highly positive correlation among exposure, dose and risk' (US EPA 1992 p. 22921). The treatment of dose–response relationships that reflect differences between individuals and populations is discussed in Section 3.6.1.

4.8.1 Population exposure assessments

Risk estimates may need to be made for populations with high exposures, average exposures and unusual exposure circumstances (e.g. very high, short-duration exposures).

Populations at higher risk because of high exposures or increased susceptibility are of particular importance in risk management, as often mitigating actions will be framed around these groups. Increased susceptibility may be due to physiological or metabolic factors. It may not be possible to identify such people and for this reason the conservatism built into EHRA should provide protection for more susceptible sub-groups. High-exposure people are usually defined as being 'above the 90th percentile of the population distribution, but not higher than the individual in the population who has the highest exposure' (US EPA 1992 p. 22901).

There is a need for caution when applying the concept of 'worst case exposures'. These exposures are often based on the accumulation of a range of unlikely but individually plausible scenarios. Such worst exposure cases are often worse than any remotely plausible case because they can represent a 'hypothetical individual and an extreme set of conditions [that] will usually not be observed in actual populations' (US EPA 1992 p. 22901). The term 'worst case exposure' is to be contrasted with the more practical term 'maximum exposed individual', which describes 'an individual that does, or is thought to, exist in the population'.

4.8.2 Use of bounding estimates

The US EPA (1992) has used the concept of an upper bounding estimate. This estimate is essentially marginally higher than the highest exposure, dose or risk incurred by the person in the population with the highest exposure, dose or risk. It is therefore a useful point against which to judge estimates of exposure, dose or risk to see whether they are plausible.

A lower bounding estimate may be set to define the level at which exposures, doses or risks become insignificant or trivial.

Given an upper and a lower bound estimate, values outside this range may be discounted from further consideration and this can be a useful method for rationalising exposure scenarios providing there is an awareness of conservative assumptions used to set the bounding estimates.

4.9 ISSUES IN EXPOSURE ASSESSMENTS

All exposure pathways must be considered for health risk assessment, although one exposure pathway may be dominant. As the total amount of a chemical absorbed by a person's body influences the risk to health, exposure assessment must take into account all sources of exposure irrespective of whether these are from food, water, the workplace, outdoor air, or a combination of these and other sources (Langley 1991a; IEH 1999a).

In large-scale contamination (i.e. regional), more exposure pathways will be involved than in small-scale (very localised) contamination.

Children usually receive a higher exposure to environmental agents per unit body weight than adults because of behavioural and physiological factors (e.g. hand-to-mouth activities for soils, higher respiration rates per unit body weight, increased gastrointestinal absorption of some substances).

For soil contaminants, ingestion is usually by far the most important exposure route for small children.

Bioaccumulation may be a significant concern for some substances with long biological half-lives (e.g. cadmium, organochlorine pesticides), and this factor should be considered.

4.9.1 Environmental distribution

In developing sampling plans for chemical agents and assessing exposure, an understanding of the movement of chemical agents between environmental compartments and the effects of environmental partitioning will be necessary.

Partitioning will reflect the fact that substances tend to move to the environmental compartment for which they have the most affinity (Calamari 1993; 1999). Transformation may occur in any environmental compartment.

Fugacity modelling (assessment of the escaping tendency of a chemical from one environmental phase to another) enables an estimation of which compartment will contain most of the agent and where the highest concentrations in the 'unit of world' are (Mackay 1991). Mackay's 'unit of world' is a hypothetical box 1 km square and 6 km deep that includes air, terrestrial and aquatic biomass, soil, water and sediment. Because environmental data is typically distributed on a log scale, the approximate estimates provided by fugacity models are very useful for distinguishing where a contaminant is likely to reside, even if the absolute distribution estimates are not completely accurate. Simpler fugacity models assume equilibrium, which in many cases does not apply, but the equilibrium estimates still demonstrate partitioning tendency.

Especially where monitoring data is inadequate, fate models are useful for estimating chemical concentrations. These models can span a wide range of complexity in terms of spatial dimensions and temporal assumptions (i.e. steady state versus non-steady state).

Types of fate models include (from WHO 1999b, p. 42):

- *simple dilution models*, where either: a measured concentration in an effluent is divided by a dilution factor; the chemical release rate is divided by a dilution factor; or the chemical release rate is divided by the bulk flow rate of the medium
- *equilibrium models*, which predict the distribution of a chemical in the environment based on partitioning ratios or fugacity

- *dispersion models*, which predict reductions in concentrations from point sources based on assumed mathematical functions or dispersion properties of the chemical
- *transport models*, which predict concentration changes over distance and can represent dispersion, biochemical degradation and absorption.

4.9.2 Environmental persistence

The terrestrial, aquatic and atmospheric fate of agents needs to be considered as part of the exposure assessment. The agents may be relatively inert (e.g. asbestos) or subject to biodegradation and abiotic degradation. Persistent substances, or those with long half-lives in the environment or biota, may provide opportunities for exposure to increase over time, such as if there are movements in the population, as can occur with residential redevelopments of an old industrial area.

The source of the data and the relevant environmental comparisons with Australian conditions should be taken into account. For example, Australian soils and climatic factors may result in different environmental persistence for some pesticides in Australia compared with North American or northern European conditions.

4.10 EXPOSURE ASSESSMENT OF VOLATILE AGENTS

Exposure assessment for volatile agents may need to be undertaken under a variety of circumstances and media (e.g. soil/landfill, water, spills, consumer goods) and will often depend on modelling. It is beyond the scope of this enHealth document to discuss modelling of volatile agents in detail.

In relation to modelling issues of volatiles in groundwater, there are more extensive discussions of such modelling in:

- *Review of schedule B(4) of the contaminated sites NEPM* (NEPC 2010)
- CRC-CARE technical reports on the development of health screening levels for petroleum hydrocarbons in contaminated soil and groundwater (CRC-CARE 2009)
- *Evaluating vapor intrusion pathways at hazardous waste sites* (ATDSR 2006)
- *OSWER draft guidance for evaluating the vapor intrusion to indoor air pathway from groundwater and soils (subsurface vapor intrusion guidance)* (US EPA 2002b). This document is expected to be updated by late 2012.

For groundwater the commonly used models (e.g. the Johnson and Ettinger (1991) model) for assessing vapour intrusion into indoor spaces from soil or groundwater combines convective and diffusive parameters into a one-dimensional model describing these transfer functions. Some of the limitations of these models have been discussed by Turczynowicz and Robinson (2007). They can overestimate or underestimate indoor air concentrations depending on whether attenuating factors such as biodegradation, or finite versus infinite sources, are taken into account. Under-estimation may occur when factors such as preferential pathways and non-diffusive vapour ingress (e.g. advection) are addressed. The Johnson and Ettinger model has been used extensively in Australia. For example, with appropriate modification, it was used for the assessment of petroleum hydrocarbon vapour intrusion in the CRC-CARE development of health screening levels (HSLs) (CRC-CARE 2009) and in the revision of the HIL for total petroleum hydrocarbons (TPH) in the contaminated sites NEPM (NEPC 2010).

Currently, field monitoring data is the most appropriate to use in assessing exposures to volatile substances. Environmental fate and modelling characteristics present problems for using short-term field monitoring data. This is particularly marked for the decay of exposures to finite sources of volatile substances. The validity of assumptions of biodegradation and finite source are difficult to assess without an empirical database showing paired measurements of indoor air versus sub-slab and soil concentrations that test the applied assumptions and exposures in building developments on potentially contaminated land.

The failure rate of vapour intrusion models (as routinely applied by consultants in Australia) is currently unknown. To avoid potential public health and financial impacts, vapour modelling should be appropriate and reasonably conservative. Demonstration of conservatism may be achieved by analysis of model uncertainty in which sensitive parameters are identified and the influences of changes in model input values are compared with the output from the assumptions of an exposure scenario considered realistic. The risk manager may then choose a modelling scenario that provides an appropriate level of precaution. Arguably, the ultimate test of modelling appropriateness is collection of indoor air and measurement of the contaminants of concern. This of course requires attention to sampling and statistical issues to ensure confidence in the resulting data.

4.11 EXPOSURE DURATION

Exposure is rarely constant over time. It can vary substantially over periods ranging from a few minutes, to hours, days or weeks. Variation is common where the source emits the COPCs intermittently, e.g. from the stack of an industrial facility.

Figure 22: Illustration of potential variability in exposure patterns

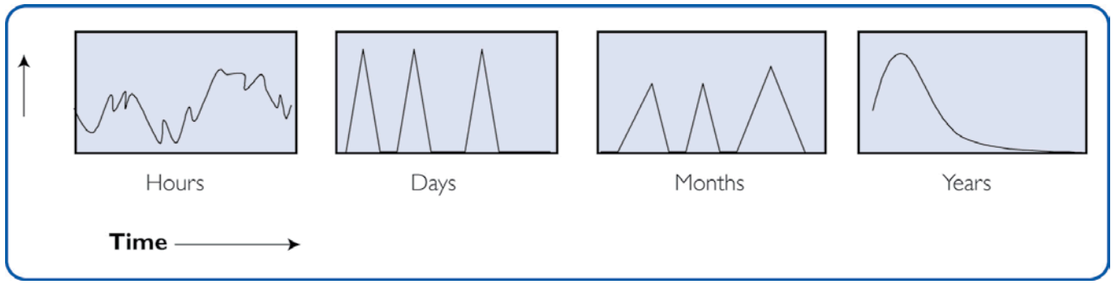


Figure 22 shows some of the potential exposure scenarios that may occur over a period of time. It demonstrates the difficulties in describing a variety of exposure scenarios with a single number.

If fluctuations in exposure are relatively large, the risk assessor may need to consider whether averaging the exposure over a relevant exposure period could underestimate risk, especially where acute effects associated with peaks in exposure could result in more critical effects on health.

4.11.1 Short-term versus long-term exposure data

Short-term data applies to data collected over minutes, hours, days or months. Long-term data refers to years or 'lifetimes'. Long-term exposure estimates based on a 'snapshot' of data at a particular point in time may not accurately represent the fluctuations that occur during exposures lasting years or a lifetime. Populations are not randomly mobile or static, and using short-term data to estimate long-term exposures tends 'to underestimate the number of people exposed, but to overestimate the exposure levels to the upper end of the distribution even though the mean will remain the same' (US EPA 1992 p. 22917). Conversely, the fact that long-term data often tends to even out exposures can mean that significant variations due to short-term conditions or activities can be missed.

The issue of whether estimates of long-term or short-term exposure should be used in the risk assessment may also be informed by the nature of the predicted health effects. For example, short-term exposure estimates (particularly peak exposures) may be more relevant where an acute effect (e.g. sensory or mucosal irritation) is the effect driving the risk assessment.

4.11.2 Adjusting exposure duration

It is possible that toxicological data against which inhalational exposures is to be benchmarked has been developed from studies where the duration of exposure is different to the exposures modelled or expected in an environmental health risk assessment. The conventional approach to adjusting the toxicological benchmarks is the application of Haber's Law, or a variation of that approach. This assumes that concentration and time of exposure are equally important in producing the effect.

Haber's Law states the product of the concentration (C) and the time of exposure (t) is equal to a constant level or severity of response (K) for a specific toxicological effect. Therefore:

$$C \times t = K$$

This equation is equivalent to the area under the exposure curve. The area under the dose–response curve (AUC) that measures the total delivered dose to a target tissue is an analogous concept. However, not all substances follow this simple relationship, and a more general exponential relationship of:

$$C^n t = K$$

n is a chemical-specific parameter that is also specific for a specified health endpoint.

This is required to describe the effects of concentration and exposure time for some toxicological endpoints (Ten Berge et al. 1986).

Guidance on using such data for setting air quality standards (NHMRC 2006) recommends that, as a general rule, averaging times should be consistent with or, where practical, shorter than the elicitation time for the effect of concern. This is an appropriate and conservative approach for public health protection. Therefore:

- for compounds whose effects require chronic exposure, the averaging time should be a year
- for pollutants whose effects are rapid in onset and/or readily reversed, the averaging time should be of the order of days, hours or less.

Ideally, the averaging time will be determined from the experimental data used to identify and measure the key adverse effect. However, issues such as the practicality of measurement, the desire for uniformity or the possibility of short-term spikes within the averaging period may influence the final regulatory time frame assigned to an air standard.

For acute systemic adverse effects, toxicokinetic information on the active compound (either the parent molecule or metabolite) may inform selection of an appropriate averaging time. For example, where an adverse effect that is not dependent upon accumulation of toxicity is identified in a clinical or animal study, but is observed some time after steady-state blood or body-burden concentrations are achieved, the averaging time for a standard could be determined by adjusting the experimental observational period (often this is the same as the exposure period) according to the half-life of the active molecule. It takes approximately five half-lives to achieve steady-state conditions for blood or body burden concentrations of the pollutant. This means that an air standard averaging time less than the equivalent of five half-lives will confer additional conservatism if the numerical value of the standard has been established on a NOAEL identified from the longer experimental observational period.

Further detail on the approach to adjust inhalational exposure times, including adjustment to the application of Haber's Law, are detailed in NHMRC (2006).

4.12 ADJUSTMENTS FOR SENSITIVE SUB-POPULATIONS

Identifying sensitive groups that may be exposed to an agent of concern is often a critical component in an EHRA. Where such groups are likely to have different exposure characteristics it is necessary to apply appropriate adjustment factors to the exposure assessment. Where the inhalational exposure route involves children, note the previous advice on EHRA-proposed adjustments, taking into consideration the different breathing rates and volumes applicable to children, as well as their activity patterns. Guidance on the selection of suitable respiratory rates, volumes and activity patterns for children is summarised in Sections 5 and 6 of the *Australian exposure factor guidance* document.

It should be noted that recent US guidance on inhalational exposure for children (US EPA 2009a) recommends a different approach that does not necessarily adjust exposures based on different breathing and activity patterns. The basis for this recommendation is that US guidance on chemical-specific data for toxicity assessment recommends that, where appropriate (e.g. for mutagenic carcinogens; see Section 5.8), early-life exposures be factored into the relevant reference dose assessments. Where this is done, the US EPA considers that it is no longer necessary to adjust the exposure estimates for age-related ventilation rates or body weights (see Section 4.6 for the RAGS-F exposure equations).

While it is recognised that some Australian air quality GVs-based inhalation risk assessments may have been carried out using the previous approach, updated enHealth guidance in this document recommends the RAGS-F approach be used from now on in Australia.

4.13 DEFAULT VALUES FOR EXPOSURE ASSESSMENTS

As outlined earlier in this chapter, exposure assessment is the process of measuring or estimating the intensity, frequency and duration of human or other population exposures to risk agents (Covello & Merkhofer 1993). Where specific data is unavailable (or inadequate) to quantify exposures by all the relevant pathways, the risk assessor usually relies on modelling or other means of calculating the exposure amounts and fluxes. Such exposure models and calculations may, in turn, need to make extensive use of default values to estimate the various model inputs where direct measurements are absent.

Tabulated data on human anatomical and physiological parameters and human activity patterns relevant to Australia has been compiled in the *Australian exposure factor guidance* document (AEF).

Historically, many of these types of data have been sourced from various international sources, although the amount of relevant Australian data sources has been increasing. Australian data that may provide a useful source of default estimates for air, water, soil and food-based risk assessment have been juxtaposed with overseas data to allow an appreciation that not all overseas-sourced data truly reflect the current Australian population. In some cases, it may be necessary to use defaults that are more population- and site-specific (e.g. air quality data specific to a region). Stakeholder consultation may be useful in establishing such site-specific data.

Another use of default parameter estimates is in initial screening assessments or 'back of the envelope' appraisals to establish whether there is a need to move to site-specific appraisals.

Tables 7 to 9 summarise the recommended default parameters for adults, children and non-age-dependent exposure factors from the *Australian exposure factor* (AEF) document. Internal references in these tables refer to relevant sections in the AEF.

Table 7: Summary of suggested exposure factors for adults[a]

Parameter	Suggested value Average (95th percentile)[a]	Units	Comment
Anatomical and physiological parameters (Sections 2.1.4 & 2.2.4)[b]			
Body weight	78 (107)	kg	M & F combined*
	85 (114)	kg	M
	70 (100)	kg	F
	70	kg	Lifetime average M & F combined
Body height	169 (181)	cm	M & F combined
	176 (188)		M
	162 (174)		F
* Note these heights and weights may differ for uniform populations of a specific race (Sections 2.2.1 & 2.2.2)[b]			
Dermal exposure parameters (Section 3.2.4)[b]			
Total skin surface area	20,000 (24,000)	cm^2	M & F combined. See Table 3.2.3[b] for surface area of specific body parts.
	21,000 (25,000)		M
	19,000 (23,000)		F
Exposed skin surface area	6,300 (7,900)	cm^2	M & F combined*
	6,700 (8,100)		M*
	5,900 (7,500)		F*
* These defaults approximate the sum of forearms, hands, lower legs and feet. The actual exposed body parts should be used as in indicated by the exposure scenario. Section 3.2.4[b]			
Oral exposure parameters			
Drinking water intake M & F combined (gender-specific data not available)	2 1.2 (2.8)	L/d	Lifetime tap water (i.e. community supply) intake. Includes water used in food preparation. Excludes commercially purchased bottled water and water intrinsic to purchased food and beverages (i.e. milk). Less than lifetime tap water intakes. (Water intake may be much larger with high activity ± tropical or arid areas). Can be used for pregnancy but 50% increase during lactation. (Section 4.1.3)[b]
Food intake	1,400	g/day	M & F combined*
	1,550		M*
	1,200		F*
* Average food intakes not including beverages (e.g. juices, tap water, coffee) but including milk for ≥19 yrs; upper intakes from recent Australian food surveys are not readily available. For intakes of individual food groups see Section 4.3.4[b] and Tables 4.3.1a,b,c.[b]			
Soil ingestion	50 (60)	mg/day	Section 4.5.3[b]
Incidental water ingestion while swimming	25 (125)	mL/hr	Average (upper estimate) (Section 4.6.3)[b]

Parameter	Suggested value Average (95th percentile)[a]	Units	Comment
Inhalation exposure parameters			
Inhalation rate (long term-exposures)	15 (20)	m^3/day	M & F combined (Section 5.1.3)[b]. For specific inhalation rates by activity or for short-term exposures, see Section 5.1 (Table 5.1.2)[b]
Activity patterns			
Total time indoors	20 (24)	hr/day	Average (upper estimate) (Section 6.2.3.2)[b]
Time indoors (at home)	20 (24)	hr/day	Average (upper estimate) (Section 6.2.3.2)[b]
Time spent outdoors	3	hr/day	Approximate average value for Australian adults (Section 6.2.3.2)[b]. Upper estimate not available.
Time spent swimming	0.5	hr/day	For general population*
	1.5		For people who swim regularly*
* Estimates for Australians assuming all outdoor sports activity is swimming (Section 6.2.4.3)[b].			

M = Male (adult)

F = Female (adult)

a The summary tables provide suggestions for possible values for use in screening risk assessments. An average (i.e. central) and reasonable maximum value is provided. The latter is in parenthesis and when data permits is the 95th percentile, otherwise it will be an 'upper' estimate as indicated. In general, an 'upper estimate' is a reasonable maximum value. The specific sections in the AEH should be consulted for additional explanations. It is the ultimate decision of the risk assessor to choose the most appropriate value to use on a case-by-case basis. Wherever possible, data which is specific for the risk assessment scenario, chemicals and receptors of concern should be used ahead of the values in this table.

When separate values for males or females are not provided in the summary tables the recent data used for generating the tables did not contain this information. Older agency publications (e.g. from Australia, US EPA, Canada, the Netherlands) may have such data and the risk assessor should seek and justify the use of this information as needed.

b The references to sections and tables refer to those in the *Australian exposure factors guidance* document.

Table 8a: Summary of suggested exposure factors for 2–3-year-old child (male and female combined)

Note: Values are average or 95th percentile (in parenthesis)[a]

Parameter	Suggested default	Units	Comment and internal reference
Anatomical and physiological parameters			
Body weight	15 (17)	kg	Section 2.2.4[b]
Body height	96 (106)	cm	Section 2.1.4[b]
Dermal exposure parameters			
Total skin surface area	6,100 (7,000)	cm^2	Section 3.2.4. See Table 3.2.5 for specific body part data[b].
Exposed skin surface area	2,300 (2,700)	cm^2	Section 3.2.3, 3.2.4[b]
Oral exposure parameters			
Drinking-water intake	0.4 (1)	L	Tap water intake. Includes water used in food preparation.(Section 4.2.3)[b].
Food intake (excludes beverages except for milk)	1,100	g/day	Upper percentile not available. For intakes of individual food groups, see Section 4.4.4 and Table 4.4.2[b].
Soil ingestion	50, (100), [100]	mg/day	Central tendency for outside soil. (Reasonable maximum, outside soil) [Central tendency outside soil + indoor dust]. (Section 4.5.3)[b]
Incidental water ingestion while swimming	50 (~ average) 150 (~ upper estimate)	mL/hr	Section 4.6.3[b]
Inhalation exposure parameters			
Inhalation rate	9.5, (15.9)	m^3/day	Section 5.1.3[b]. For specific inhalation rates by activity or for short term exposures, see Section 5.1; Table 5.1.2[b]
Activity patterns			
Frequency of hand to mouth	13 (37) (indoors)	contacts/ hr	Section 6.1.1.3[b]
	5 (20) (outdoors)		
Mouthing duration	Varies	hrs/d	Mean and maximum values differ by object mouthed (Section 6.1.1.3)[b]
Time spent indoors	21.9 (total)	hrs/d	Upper estimate not available (Section 6.1.2.3)[b]
	16 (21.6) (at home)	hrs/d	Section 6.1.2.3[b]
Time spent outdoors	2	hrs/d	Average. Upper percentiles not available (Section 6.1.2.3)[b]
Playing on sand/gravel	0.9	hrs/d	
Playing on grass	1	hrs/d	
Playing on dirt	0.8	hrs/d	
Time spent swimming	23	hr/year	Average value. Upper estimate not available (Section 6.2.4.3)[b]

Table 8b: Summary of suggested exposure factors for 1–2-year-old child (male and female combined)

Note: Values are average or 95th percentile (in parenthesis)[a]

Parameter	Suggested default	Units	Comment and internal reference
Anatomical and physiological parameters			
Body weight	11 (13)	kg	Section 2.2.4
Body height	81 (86)	cm	Section 2.1.4
Dermal exposure parameters			
Total skin surface area	5,300 (6,100)	cm^2	Section 3.2.4. See Table 3.2.5 for specific body part data.
Exposed skin surface area	1,600 (1,900)	cm^2	Section 3.2.3, 3.2.4
Oral exposure parameters			
Drinking-water intake	0.3 (0.9)	L	Tap water intake. Includes water used in food preparation (Section 4.2.3).
Food intake (excludes beverages except for milk)	720 (1,700)	g/day	For intakes of individual food groups, see Section 4.4.4.
Soil ingestion	50, (100), [100]	mg/day	Central tendency for outside soil. (Reasonable maximum, outside soil) [Central tendency outside soil + indoor dust]. (Section 4.5.3)
Incidental water ingestion while swimming	50 (~ average) 150 (~ upper estimate)	mL/hr	Section 4.6.3
Inhalation exposure parameters			
Inhalation rate	8.0 (12.8)	m^3/day	Section 5.1.3. For specific inhalation rates by activity or for short term exposures, see Section 5.1;Table 5.1.2
Activity patterns			
Frequency of hand to mouth	20 (63) (indoors)	contacts/ hr	Section 6.1.1.3
	14 (42) (outdoors)		
Mouthing duration	Varies	hr/d	Mean and maximum values differ by object mouthed (Section 6.1.1.3)
Time spent indoors	22.6 (total)	hr/d	Upper estimate not available (Section 6.1.2.3)
	17.8 (24) (at home)	hr/d	Section 6.1.2.3
Time spent outdoors	1.4	hr/d	Average. Upper estimates not available (Section 6.1.2.3)
Playing on sand/gravel	0.7	hr/d	
Playing on grass	1.1	hr/d	
Playing on dirt	0.9	hr/d	
Time spent swimming	21	hr/year	Average value. Upper estimate not available (Section 6.2.4.3)

a See Footnote (a) to Table 7. For screening risk assessments and establishing guidelines the most sensitive receptor is assumed to be a 2–3-year-old (Table 8a) or a 1–2-year-old (Table 8b).

b The references to sections and tables refer to those in the *Australian exposure factors guidance* document.

Table 9: Summary of non-age-dependent exposure factors

Note: Values are average or 95th percentile (in parenthesis)[a]

Parameter	Suggested value	Units	Comment
Anatomical and physiological parameters			
Life expectancy	82 [*]	yrs	Male and female combined (Section 2.4.1)[b]
	79		Male (Section 2.4)[b]
	84		Female (Section 2.4)[b]
*Upper estimate not available. Many national and international agencies use 70 yrs as the assumed lifetime exposure to environmental agents.			
Dermal exposure parameters			
Soil adherence	0.5 (1.7)	mg soil/cm^2 skin	Applicable for outdoor and indoor residential child and adult exposures (Section 3.3.1)[b]. For specific activities and body parts see Tables 3.3.3, 3.3.4, and 3.3.5[b].
Dermal bioavailability	Organics: 1 Inorganics: 0.0001	Unitless	Chemical-specific. These table values are to be used only when other reasonable information is not available. Bioaccessibility of inorganics from soil or other media can be approximated with experimental tests (Sections 3.4 and 4.0)[b].
Dermal bioaccessibility	Organics: 1 Inorganics: No default		
Shower and bath frequency	1 (2)	#/day	Central estimate (upper estimate) for adults and children (Section 3.5.5)[b]
Shower duration	8 (16)	mins	Section 3.5.5[b]
Shower volume	72	L	Volume and flow rate of non-water saving shower (Section 3.5.5)[b]. Upper estimates not available.
Shower flow rate	9	L/min	
Bath duration	21	mins	Bath duration for adults and children combined (Section 3.5.5)[b]. Upper percentile not available.
Bath volume	Insufficient data	L	Insufficient data (Section 3.5.4)[b].
Oral exposure parameters			
Oral bioavailability	Organics: 1 Inorganics: No default	Unitless	Chemical-specific. These table values are to be used only when other reasonable information is not available. Bioaccessibility of inorganics from soil or other media can be approximated with experimental tests (Sections 3.4 and 4.0)[b].
Oral bioaccessibility	Organics: 1 Inorganics: No default		
Inhalation exposure parameters			
Building air exchange rate	*Residential:* 0.6 *Commercial:* no recommendation	#/hr	The residential value is midpoint of range for 'closed' Australian dwellings. Air changes will be higher with open doors/windows, ceiling fans and air conditioning. A single value is not suggested for commercial buildings. Ur estimates not available (Section 5.2.4)[b].
Indoor particle deposition rate	No recommendation	#/hr	Markedly differs from house to house. No suggested value (Section 5.3.1)[b].
Floor area of residential dwelling	*Houses:* 190 *Other:* 120 *All:* 180	m2	Average values. There is a wide range of values for houses and other types of dwellings (Section 5.4.1)[b]. Upper estimates not available.

Parameter	Suggested value	Units	Comment
Air volume of residential dwellings	*Houses:* 460 *Other:* 280 *All:* 420	m^3	Average values. Assumed ceiling height is 2.4 m. There is a wide range of values for houses and other types of dwellings (Section 5.4.1)[b]. Upper estimates not available.
Uptake (product of retention and absorption) of inhaled contaminants	Default 100%	unitless	Chemical-specific value should be used if available (Section 5.5)[b].
Background particulate levels for urban ambient air	PM^{10}; 17 (39) $PM_{2.5}$; 7 (16) Ratio; [0.4]	µg/m^3 [unitless]	National average of all urban monitoring sites – 50th percentile, (95th percentile). Proportion of PM_{10} that is $PM_{2.5}$. (Section 5.6.2)[b].
Fraction of indoor dust from outside soil	50 (100)	%	Average (maximum) (Section 5.7)[b]
Activity patterns			
Time spent in transit	1	hr/day	Total time by all travel modes (Section 6.2.3.2)[b]. Upper estimate not available.
Frequency of swimming	52 (150)	d/yr	Approximate median (upper estimate). Actual frequency will depend on the Australian locality (Section 6.2.4.3)[b].
Residence and population mobility parameters			
Duration of residence	10 (35)	yr	Section 7.1.3[b].

a See Footnote (a) to Table 7. For screening risk assessments and establishing guidelines the most sensitive receptor is assumed to be a 2–3-year-old.

b The references to sections and tables refer to those in the *Australian exposure factors guidance* document.

It should be noted that body weight is a parameter used to adjust dose and exposure in many aspects of risk assessment. Because of demographic differences, various authorities around the world have used slightly different default values for age-related body weight in their risk assessment equations. For example, the default generic value used by Health Canada (1994) is 70 kg, a value commonly found in many health risk assessment documents, including the *Australian drinking water guidance* (NHMRC, NRMMC 2004). The WHO figure is 60 kg and the US EPA (1997a) *Exposure factors handbook* lists 71.8 kg (increased to 80 kg in the 2009 update).

Both the Australian and US handbooks provide extensive tables of age-related body weights and growth curves. The default adult (combined sexes) body weight value recommended in the *Australian exposure factor guidance document* is 78 kg, with 85 kg and 70 kg recommended for only males or females respectively, and 70 kg recommended for whole-of-life body weight average (see Table 7).

In Australia, it is generally assumed that the most sensitive individual is the 2–3-year-old child (enHealth 2003, 2004; NEPC 1999). Data is provided throughout the *Australian exposure factor guidance* document for a 2–3-year-old child. However, in Table 8b, information is also provided for a 1–2-year-old child. The risk assessor should determine which age bracket most closely resembles the most sensitive individual for their exposure scenario.

It is important to recognise that many toxicological reference values and environmental guidance values will have been derived using exposure factors that are now outdated.

4.14 SOURCES OF EXPOSURE ASSESSMENT DATA

It is recommended that the *Australian exposure factor guidance* (AEF) document be used as the primary source of exposure data. While the information contained in this document is not intended to be a comprehensive compendium of exposure parameters, it has been compiled with a view to providing guidance to Australian risk assessors. While a number of 'recommendations' have been made regarding parameter values in the text (summarised in Tables E1 to E7), risk assessors and others using the information should check to ensure the suggestions presented are suitable for the scenarios they are evaluating. Australian data should be used where it is available.

It is important that the risk assessor consult the text and, if necessary, the primary information source, prior to using any of the summary information contained in Tables E1 to E7.

Australian exposure factor information has been sought and juxtaposed with overseas data to allow an appreciation of the fact that not all overseas data reflects sectors of the current Australian population. If Australian information is not available, overseas data may be used, but will require justification in the risk assessment as to why they are applicable in Australia.

It should also be appreciated that the information may not be current at the time the risks assessor consults this document. Some values may be more than a decade old and Australian demographics and behaviour may have changed from the time the information was first gathered. For example, there are currently many more people of Asian, Indian and African descent residing in Australia. People are more mobile and, due to water restrictions in most states and territories, shower durations and garden irrigation are different from 10–20 years ago. These examples highlight the necessity to make sure exposure parameter values are contemporary and 'fit for purpose'. It is the risk assessor's responsibility to ensure this is so.

Other useful sources of information and data (of a somewhat older vintage) include:

- *Exposure scenarios and exposure settings* (Taylor & Langley 1998)
- *Exposure factors in risk assessment* (Langley & Sabordo 1996)
- *1996 Australian exposure factors* (Langley, Taylor & Dal Grande 1998).

The Australian Bureau of Statistics can provide a range of Australian data.

The US EPA has developed an exposure factors program to further advance the science of exposure assessment in risk assessment. The US EPA *Exposure factors handbook* (US EPA 1997a) is a comprehensive collation of exposure information used in US risk assessment practices. This is continually updated, and the most recent update as released for consultation in October 2009 and published early in 2012.

The American Industrial Health Council's *Exposure factors sourcebook* (1994) provides examples of probability distributions for a range of exposure factors. These largely relate to the US population. These, and similar US-based data, should only be used if they can be demonstrated to be relevant to the Australian population.

Chapter 5: Risk characterisation

5.1 INTRODUCTION

Risk characterisation is the final step in the risk assessment process that:

- integrates the information from hazard identification, dose–response assessment and exposure assessment
- discusses chemicals of potential concern (COPC) and quantifies risks associated with these specified chemicals
- identifies the contributions to risk from all the relevant exposure pathways, and aggregates these risk estimates
- considers the possibility that multiple COPCs may have cumulative effects, and considers options for best integrating the effects of combined exposures (see Chapter 12)
- describes the risks to individuals and populations in terms of nature, extent and severity of potential adverse health effects
- provides an evaluation of the overall quality of the assessment and the degree of confidence the risk assessors have in the estimates of risk and conclusions drawn; this should be based on appropriate uncertainty and sensitivity analyses
- communicates results of the risk assessment to the risk manager
- provides key information for risk communication.

Risk characterisation is identified as part of Phase II of the expanded framework for EHRA outlined in Figure 2.

The overall objective of the risk characterisation stage is to determine that exposures to COPCs from the environmental source under consideration do not exceed a level considered to be protective of human health. In practice, this means that the estimated total exposure (including background where relevant) does not exceed a toxicological reference value or a health-based guideline value, usually one that has been set using the same principles of health risk assessment set out in these enHealth guidelines (see Section 5.5).

The final risk characterisation is limited by the available data, and this should be discussed in the uncertainty assessment. The process requires considerable expertise. If data is collected and analysed according to the principles and guidelines in this enHealth document, the process will become more transparent and consistent. Some parts of the risk assessment process such as 'data collection' and 'exposure assessment' will be, at least in part, quantitative and possibly based on modelling or extrapolations from measured data. These guidelines are intended to assist the qualitative process of determining whether environmental health intervention is required or not required.

Risk characterisation may involve comparing environmental data, exposure data, intakes and biological monitoring results with established criteria, including guideline values (GVs) established or published by authoritative sources.

Due to the complexities of the matter, the risk characterisation process cannot be reduced to a 'cookbook'. In this context, the guidance in this document consistently recommends that the choice of default parameters, GVs or risk assessment methodology must include an assessment of their suitability for use in the EHRA at hand. In other words, care must be taken to ensure that published or derived health-based GVs are 'fit for purpose'.

5.2 KEY PRINCIPLES IN ENVIRONMENTAL HEALTH RISK CHARACTERISATION

There are a number of key principles for health risk characterisation:

1. Protection of human health is the primary objective. Human health risk assessment is generally undertaken with an appreciation that the health risk assessment is part of a larger assessment that encompasses ecological risk assessment. However, actions based on the risk characterisation taken should always adequately protect public health and the environment, putting these responsibilities before all other considerations.
2. Risk assessments should be transparent (Schreider et al. 2010). The nature and use of default values and methods, assumptions and policy judgements in the risk assessment should be clearly identified and documented. Conclusions drawn from the evidence should be distinguished from policy judgements, and the influence of 'scientific judgement' made clear.
3. Risk characterisations should include a summary of the key issues and conclusions of each of the other components of the risk assessment, as well as describing the nature and likelihood of adverse health effects. The summary should include a description of the overall strengths and limitations of the assessment and conclusions.
4. To protect public health and the environment an appropriate degree of conservatism must be adopted to guard against uncertainties. There should be a detailed description of the areas of uncertainty and an analysis of the effects of these on any derived values.
5. Risk characterisations (and risk assessments) should be undertaken using methodologies outlined in this enHealth document, noting that methodologies may be revised as needed to maintain consistency with best scientific practice. Reports should follow a consistent general format (see Chapter 7), bearing in mind the need to recognise the unique characteristics of each specific situation.
6. Risk assessors should review the most up-to-date scientific literature relevant to the risk assessment under consideration and to the toxicological profile of the identified COPC. Information in appropriately peer-reviewed articles should be accorded greater weight than information in articles that are not peer-reviewed.
7. Variations in risk assessments as a result of particular statutory requirements, resource limitations, and other specific factors should be explained as part of the risk characterisation. For example, a reason will be required to explain why certain elements are incomplete.

5.3 QUANTITATIVE AND QUALITATIVE RISK CHARACTERISATION

The level of risk estimated in any risk assessment can be described either qualitatively (i.e. by putting risks into categories such as 'high', 'medium' or 'low') or quantitatively (with a numerical estimate). Current risk assessment methods described in this enHealth document provide quantitative estimates of risk but the precision of any such estimate will be limited by the data available to use in the assessment.

Differentiation of the approaches used in qualitative and quantitative risk assessments are informed by definitions that have been developed for each of these two processes.

Qualitative assessment: An inquiry process that generates non-numerical data, providing an 'understanding of a social or human problem, based on building a complex and holistic picture formed with words, reporting detailed views of informants and conducted in a natural setting' (Creswell 1994); or 'a classification process, where objects or materials are assigned to some class on the basis of tests made against established or implied criteria' (Ellison et al. 1998)

Quantitative assessment: The application of a set of scientifically measurable, reproducible and mathematically sound data values to estimate value, probability and associated risk of loss.

In quantitative risk assessment, reporting of a measurement is an approximation or an estimate of the value of the subject being measured. Such a result should only be considered complete after it has been evaluated and the uncertainties in the measurement explained. There are different ways of measuring uncertainty. Statistical analysis allows for the evaluation of random events and from those arising from a systematic effect.

Accounting for uncertainty in a qualitative assessment is the acknowledgement that the original classification has been made on the basis of the available evidence and that misidentifications may have occurred. There may be a lack of evidence in an observation that has caused it to be placed in a particular class, and this may result in either a 'false positive' or 'false negative' classification. Methods used in describing the identification of any classification should reflect the uncertainties associated with the evidence, permit updating on the basis of further evidence, and consider the probability of both types of error. Other desirable features include a lack of ambiguity, ease of calculation, clarity (especially for the presentation of the results) and broad acceptability of the reasoning for the determined classification.

Numerical estimates of risk may be an outcome of a quantitative assessment, but with the qualification that these are mathematical constructs incorporating various degrees of uncertainty. Numbers may give a misleading implication of accuracy, especially when based on poor or uncertain information (Langley 2003). The numbers generated should never be portrayed as being highly precise or accurate (IEH 1999b). The risk level should never be expressed in a way that suggests a greater degree of precision than is warranted by the data, for example, a risk level of 4.73×10^{-6} (i.e. using three significant figures rather than 5×10^{-6}) is probably meaningless in the context of an EHRA outcome (Langley 2003).

Variability associated with the identified hazards, the nature of the exposed populations or groups of people ('receptors'), and limitations in the toxicological and exposure data will all contribute to these uncertainties.

The most conservative mathematical models used in quantitative EHRA can be virtually insensitive to the actual experimental data and should be viewed only as a risk management solution, not a risk assessment technique (IEH 1999b). The extent to which manipulation of the input data can influence the resultant risk estimates should be determined using 'sensitivity analysis' techniques (see Section 5.15).

Estimates do not have to depend on the use of numbers to be useful; ordinary language may be used to indicate the level of risk. A finely divided ranking system can give a relatively accurate indication of quantity without using numbers (ACDP 1996).

Clearly defined qualitative categories can enable reliable and effective risk management decisions.

Some attempts have been made to integrate the concepts of qualitative and qualitative risk estimates by assigning explanatory statements to the various qualitative categories of risk and, in some instances, a numerical probability range associated with these descriptors. Examples of this approach may be found in an assessment by Biosecurity Australia (2006) of risk estimates associated with the importation of apples from New Zealand, and in WHO/FAO guidelines on risk characterisation of microbiological hazards in foods (WHO 2009).

While the qualitative descriptors indicating the likelihood of the occurrence of event may amplify the understanding of simple descriptors such as low, medium and high risk, the linking of these descriptors to quantitative probability estimates is a different matter. The quantitative probability estimates in the above examples were developed for quarantine and food quality risk assessment scenarios. They are not necessarily appropriate for linking qualitative descriptors with probability estimates of human health risks associated with environmental chemical exposures, nor have such linkages been endorsed by enHealth or any other public health authority for this purpose.

Qualitative risk conclusions can also be used to avoid the false sense that the extent of the risk is known precisely. Using terms such as 'high', 'medium' or 'low' may have different interpretations for different groups so they should be clearly defined. Such definitions, or risk categorisation matrices, can be found in some Australian guidance documents (e.g. AS/NZS ISO 31000:2009; *OGTR risk assessment framework*, see Section 1.3.2).

Qualitative risk characterisation may need to be put into context or compared with other risks relevant to the community. However, if comparisons do not directly relate to alternative options, they may be counterproductive and caution should be exercised in making such comparisons, especially if like is not compared with like, or if comparisons are being used to imply acceptability. Flippant comparisons are certainly likely to be counterproductive (DOH 1998). Comparisons should be used only where the evidentiary base and the method for risk estimation are similar and where the uncertainties in all the comparative estimates are shown (Thomas & Hrudey 1997).

It is important to consider contingent risks. This requires not looking at risks in isolation so that, for example, the risks of immunisation or chlorination are considered in the context of the risks of not having immunisation or chlorination.

5.4 RISK ESTIMATION

Risk estimation combines the estimated intakes calculated in the exposure assessment with the toxicological reference values (TRVs) (threshold and non-threshold where relevant) from the toxicity assessment to produce numerical indices of likely health effect. The risk estimation methodology differs for threshold and non-threshold compounds due to the different modes of chemical effect.

5.4.1 Threshold risk estimation

For threshold compounds, the intake for each exposure pathway is divided by the appropriate threshold, TRV (allowing for intakes from other sources where relevant) to produce a simple ratio, termed a 'hazard quotient' (HQ) (commonly used/historical term) or 'risk quotient' (RQ) (WHO recommended term).[2] The HQs for all exposure pathways for each contaminant can be summed to produce a total hazard index (HI) or risk index (RI) (see Section 12.4).

$$\text{Hazard quotient (HQ)} = \frac{\text{Intake (mg/kg/day)}}{\text{Threshold TRV (mg/kg/day)}}$$

$$\text{Hazard index (HI)} = \Sigma \text{ Hazard quotients}$$

The HQ/RQs for all exposure pathways for all contaminants should be summed to produce a total HI/RI, unless evidence is available to show this is not appropriate. When summing these HQs, the following should be taken into consideration:

- HI/RIs should be calculated for each age group (category) separately.
- HI/RIs should be calculated separately for chronic, sub-chronic and shorter-duration exposures.
- Ideally, HI/RIs should be categorised into groups of chemicals that induce the same type of effects or that act by the same mechanism of action. However, this process is not simple and requires a thorough understanding of the toxicology of the chemicals concerned, and must only be undertaken by an appropriately qualified toxicologist. If this segregation is not performed carefully, an overestimate (or, less commonly an underestimate) of the true risk could result. When toxicological information is lacking or unclear, it can be assumed that the chemicals act by the same mechanism of action with summation of the HQ/RIs. It should be noted that this will result in an overestimate of the true risk. See Section 12.4 for a wider discussion of the use and limitations of the HI/RI approach.
- HI/RIs should represent the exposure pathways that have the potential to expose the same individual or sub-population, making sure to consider areas of highest exposure for each pathway for both current and future land uses. All exposure pathways should be summed unless information is available that indicates the same individual or sub-population cannot be exposed by a particular pathway.

2 Note that the WHO preferred terms are 'risk quotient' and 'risk index'. The terminology HQ/HI is retained in here because of its widespread popular use in the USA and Australia, including in the revised NEPM on contaminated sites assessment.

5.4.2 Non-threshold risk estimation

Where non-threshold TRVs are adopted (that is, assuming a linear low-dose relationship), risks are estimated as the additional probability of an individual developing cancer over a lifetime as a result of exposure to the carcinogen. The estimated intake for each exposure pathway and non-threshold TRV are multiplied to produce pathway-specific estimates of increased lifetime cancer risks (ILCR).

However, for those carcinogens where appropriate benchmark dose data is available and suitable for use, the risk estimation method outlined above for threshold compounds applies.

$$\text{ILCR} = \text{intake (mg/kg/day)} \times \text{TRV(mg/kg/day)}^{-1}$$

$$\text{ILCR} = \text{exposure concentration (mg/m}^3\text{)} \times \text{TRV(mg/m}^3\text{)}^{-1}$$

ILCR estimates from all pathways and all contaminants should be summed to produce a total additional increased lifetime cancer risk, unless evidence is available that suggests that this is not appropriate. When combining ILCR estimates, the US EPA (1989) identifies several limitations that may be considered. These include:

- As each non-threshold TRV is an upper 95th percentile estimate of potency, and because upper 95th percentiles of probability distributions are not strictly additive, the total cancer risk estimate might become artificially more conservative as risks from a number of different carcinogens are involved.
- It will often be the case that substances with different weights of evidence for human carcinogenicity are included. The cancer risk equation for multiple substances sums all carcinogens equally, giving as much weight to Class 2 as to Class 1 carcinogens. In addition, non-threshold TRVs derived from animal data will be given the same weight as non-threshold TRVs derived from human data.
- The action of two different carcinogens may, or may not, be independent.

In practice, it will often be the case that there is insufficient information to make a well-informed decision as to whether it is reasonable or not to sum ILCRs across either pathways or contaminants (see also Chapter 12) Where more information is available, a decision to assess contaminants or pathways as independent and non-additive should be supported with reference to the toxicology of the contaminants concerned. This assessment should be undertaken by an appropriately qualified toxicologist. It is recommended that the following approach should be followed under most circumstances.

- ICLR estimates should normally be summed across pathways unless specific evidence is provided that the same person/group of people cannot be exposed by the different pathways.
- ICLR estimates should normally be summed across contaminants unless specific evidence is provided by a qualified toxicologist that this is not the appropriate approach.
- It is recognised that synergistic (that is, more than additive) effects are possible; however, the practical difficulties of quantifying the synergy in a risk assessment are significant. Unless evidence for synergistic effects is available, the potential for synergistic effects may be omitted from the assessment. Additive effects are much more common and are covered in risk assessment by summing risks across chemicals and pathways.

5.5 TOXICOLOGICAL REFERENCE VALUES (TRVS) IN COMMON USE

As outlined in Chapter 3, both threshold and non-threshold methodologies for utilising dose–response data rely on establishing a point of departure (POD) for further analysis. For threshold responses (all toxicological endpoints other than carcinogenesis), the POD is usually the NOAEL, LOAEL or a BMD.

This POD is then divided by a series of safety factors (SFs – sometimes termed uncertainty or modifying factors) to determine an acceptable or tolerable daily intake (ADI or TDI). The general methodology for assessing the ADI or TDI is outlined in Section 5.5.1.

Table 10 describes the toxicological reference values possibly generated by the threshold approach.

The term ADI is generally used for pesticides that are deliberately used on food or crops and may appear as residues in the food chain. The term TDI is generally used when a chemical is a food or environmental contaminant. Both the ADI and TDI are conceptually similar to the RfD. The difference in terminology arises because the terms have been coined by different agencies.

Table 10: Environmental health criteria derived from a threshold approach

Toxicity reference value	Units	Description or definition
Acceptable daily intake (ADI)	mg/kg/day	The daily intake of a chemical that, during a lifetime, appears to be without appreciable risk, on the basis of all the facts known at the time (WHO 1994a). The term ADI is generally used for chemicals such as pesticides, which may be present in foods within their maximum residue level (MRL) because of their permitted uses in agriculture. The ADI and RfD are conceptually the same; the terminology differs because of development by different authorities.
Tolerable daily intake (TDI)	mg/kg/day	An estimate of the intake of a substance that can occur over a lifetime without appreciable health risk (WHO 1994a). This is conceptually the same as the ADI and RfD but used when the substance is an unintended contaminant in food or an environmental medium such as air, water or soil. This terminology avoids the implication that the contaminant is 'accepted'.
Reference dose (RfD) (US terminology)	mg/kg/day	An estimate (with uncertainty spanning perhaps an order of magnitude) of a daily oral exposure to the human population (including sensitive sub-groups) that is likely to be without an appreciable risk of deleterious effects during a lifetime.
Reference concentration (RfC) (US terminology)	mg/m^3	An estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive sub-groups) that is likely to be without an appreciable risk of deleterious effects during a lifetime.

Note that the US EPA envisages that, while the RfD or RfC are generally only used for non-cancer endpoints, it may be necessary to derive both an RfD (or RfC) and a non-threshold cancer risk slope factor where both cancer and non-cancer endpoints need to be considered in the risk assessment. The only exception is where a carcinogenic response drives the risk assessment, but it is considered that a non-threshold approach is warranted because of the proposed mode of action.

The time base may be altered from daily intake to weekly or monthly intakes where the exposure pathways or toxicokinetic behaviour of the chemical warrant a longer period of averaging. For example, the term 'tolerable monthly intake' (TMI) is applied to dioxins because of the very long half-life for elimination and the use of body-burden estimates in humans and animals to adjust intakes (Joint FAO/WHO Expert Committee of Food Additives – JECFA 2002; OCS 2004).

Sometimes the word 'provisional' is attached to the term (e.g. the provisional tolerable weekly intake – PTWI). This is generally done when further data is required to establish an acceptable or tolerable intake but a temporary GV is required by the risk managers. A provisional ADI or TDI may incorporate an additional SF in the calculation because of the inherent uncertainty.

There are other sets of health-based guideline values derived from occupational health and safety (OHS). These include threshold limit values (TLV), short-term exposure limits (STEL) permissible exposure limits (PEL), to be used in an environmental health risk assessment. While they are often derived using comparable methodology, extrapolating from animal toxicity studies, human exposure studies and epidemiological studies, they are based on the protection of workers (who are on average healthier than the whole community) during the course of a normal working shift and a normal working lifetime. They may use different levels of protection and safety factors than used for the general community, such as tolerating relatively minor adverse effects. It would be unusual for OHS-based guideline doses to be used in an environmental health risk assessment, although the risk assessor may need to be aware of potentially conflicting situations where it may not be clear whether the risk estimates of an EHRA should be applicable to both the general community and/or to workers within an exposure scenario.

5.6 DETERMINATION OF NO(A)ELs, ADIs (RFD) AND TDIs FOR HUMANS

The determination of an acceptable or tolerable daily intake (ADI or TDI) involves establishing an overall NOAEL for a chemical that is generally the lowest NOAEL in the most sensitive species.

This approach of using the lowest NOAEL is justified unless there is evidence of one or more of the following:

- from pharmacokinetic/metabolic studies that the most sensitive species shows a different toxicokinetic behaviour than humans and is therefore less relevant as a predictor of human toxicity than another toxicity test species
- that the toxic effect that has the lowest NOAEL is not relevant for humans, or
- that the lowest NOAEL is derived from an inadequate or invalid study.

Thus it is emphasised that the full database must be used and all relevant findings correlated when determining the most appropriate health endpoint.

It is important to note also that in public or occupational health risk assessments, establishing a NOAEL is likely to be influenced by a consideration of the relevant route(s) of exposure and experimental design.

The selection of the NOAEL can be significantly influenced by:

- the selection of doses used in the study – the 'real' NOAEL is likely to lie somewhere between the apparent NOAEL and LOAEL doses (if there is a relatively wide margin between doses used in the study, a higher NOAEL might have been obtained if doses had been more appropriately spaced)
- the number of test subjects in the dose levels – a smaller number of test animals per dose in a study compromises the statistical power of being able to discriminate between dose levels that produce an 'effect', compared with those where the incidence of disease or toxicity is comparable to the 'controls' or untreated animals
- the extent to which disease or toxicity associated with administration of the test agent can be discriminated from disease processes that occur naturally during ageing. This is particularly true of neoplastic responses, where it may be difficult to 'score' the number of neoplasms at different stages of the life span of the test animals, where the progression through a series of pathological changes is not well delineated. Where the progression of toxicity is time related and possibly reversible if exposure ceases, the duration of treatment becomes a more critical factor in the experimental design.

An ADI or TDI is derived from the NOAEL (or LOAEL) as follows:

$$\text{ADI or TDI} = \frac{\text{NOAEL}}{\text{SF}}$$

These exposure limits are derived by first determining the NOAEL or, if the NOAEL cannot be determined, taking the LOAEL and dividing the value by factors to account for:

- inter-species differences (extrapolating from animals to humans)
- intra-species differences (differing sensitivities between individuals)
- the severity of the adverse effect
- the quantity and quality of the scientific data.

The general approach to calculating an ADI/TDI follows the principles initially outlined in the IPCS *Environmental health criteria monograph* No. 104 (WHO 1990). The uncertainty inherent in extrapolation between and within species has generally been dealt with by using safety (uncertainty) factors.

Historically, the most common overall factor used by a number of regulatory bodies is 100, comprised of 10 to account for uncertainties in inter-species extrapolation, and 10 to account for intra-species variability. An additional factor of 10 is sometimes used if the NOAEL was not established in the study and the LOAEL used instead, if the study used to determine the ADI/TDI must be based on a relatively short-term study (e.g. 28–90 days) or if a large toxicological database has not been assessed. Application of additional factors needs careful consideration for new industrial chemicals, where the available toxicological databases may be less comprehensive than those for new agricultural chemicals, proposed food additives or medicines (human and animal). The overall factor can range from 10 to 10,000, depending on the source and quality of data, the biological relevance of the endpoint, and the hazard assessment (carried out on a case-by-case basis). In general terms only, a safety factor of 10 would apply when appropriate human data were available.

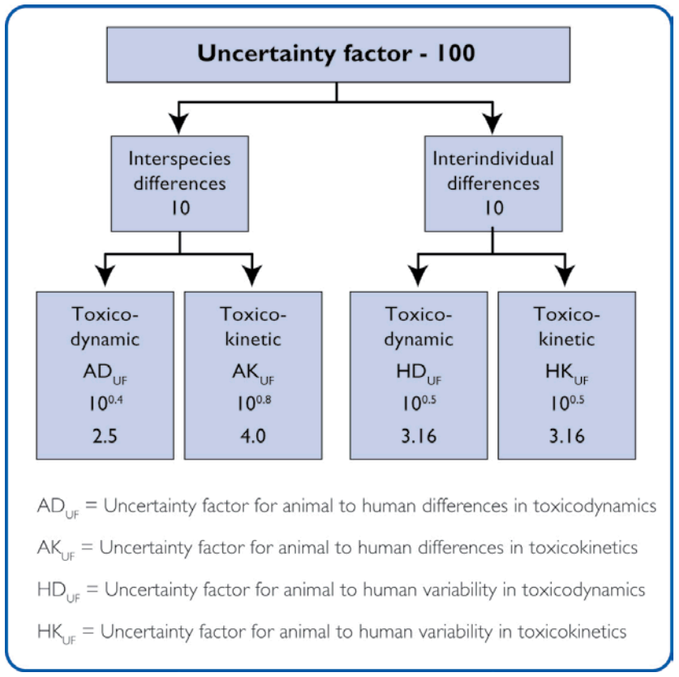
From the data available on humans and experimental animals, it appears that inter-species and intra-species differences are, in general, less than 10, hence the often-used overall safety factor of 100 for these two factors is conservative and adequately protective of public health (Johannsen 1990; Renwick & Walker 1993).

One of the outcomes of an IPCS program (IPCS 2005) to develop chemical-specific adjustment factors (CSAFs) was to further refine the breakdown of the conventionally used 100x safety factor by incorporating figures based on inter-species and intra-species toxicokinetic and toxicodynamic variation. The IPCS proposal for the default CSAFs is set out in Figure 23.

The IPCS program recommends using chemical-specific data to replace default values where adequate data is available, and outlines the nature of data on toxicokinetic and toxicodynamic variation that could be used. The program recognises that the combined uncertainty factor (CUF) based on CSAFs could be less than, or more than, the common default value of 100, but notes that this should be made transparent to the risk manager. It still recognises the need to add additional uncertainty factors when the quality of the studies is deficient or where significant data gaps occur.

The decision on the magnitude of factors to use is predominantly based on expert or informed judgement. While this approach to selecting the number and magnitude of the safety factors can appear to be somewhat arbitrary, improved knowledge of the biological processes that cause inter- and intra-species variation (e.g. metabolic and other pharmacokinetic rate differences) have generally supported the choice of the default safety factors.

Figure 23: Proposed subdivision of default uncertainty factors to be used in risk assessment



Reproduced from IPCS 2009b with permission from WHO.

It is generally accepted that if the magnitude of the overall safety factor approaches or exceeds 5,000, this is effectively an admission that there is insufficient knowledge of the environmental hazard under consideration and that the underlying data may be unsuitable to support a risk assessment. US EPA practice has been to not recommend an RfD/RfC if the combined uncertainty factor exceeds 3,000 (NRC 2008). Where a precautionary approach requires the application of such large uncertainty factors in setting a health-based guideline value, it is inevitable that when better information becomes available, the consequent change in the numerical value (often an increase) can reduce community confidence in its health-protectiveness.

However, Gaylor et al. (1999), when commenting on the possible use of a benchmark dose approach (using BMD_{10}) as a point of departure, recommended the use of a default uncertainty factor of 10,000 for irreversible adverse effects (e.g. cancer) with a smaller default uncertainty factor of 1,000 for reversible adverse effects. The argument was based on the fact that BMD_{10} approximates the LOAEL dose in conventional threshold-type risk assessments.

This update of enHealth guidance on EHRA commends the IPCS approach to selecting and justifying CSAFs and recommends that it be adopted in EHRA practice in Australia when relevant data is available.

5.7 TOXICOLOGICAL REFERENCE VALUES DERIVED USING A NON-THRESHOLD APPROACH

The two toxicological reference values that may be developed using a non-threshold approach are:

- *cancer slope factor (CSF)*: This is the plausible upper-bound estimate of the probability of a carcinogenic response per unit of intake over a lifetime; it is expressed in units $(mg/kg/d)^{-1}$
- *unit risk factor (URF):* This is an expression of carcinogenic potency in concentration terms, expressed as the probability of cancer per unit of an exposure medium (e.g. per μg/L of water, per $\mu g/m^3$ of air *or ppm).*

The CSF is used in EHRA to estimate the upper-bound probability that cancer will develop over a lifetime of exposure to a chemical at a specific level of intake. It is the slope of a linear extrapolation from the upper-bound estimate of a POD dose to zero.

The URF can be derived directly from inhalation or drinking-water studies depending on which media is being assessed. Where such data is not directly available, these unit risks can be derived by converting an oral CSF with units of $mg/kg/d^{-1}$ to a concentration of the substance in air, water or other media. These extrapolations often assume default intake rates for the specified media (e.g. an inhalation rate of $20m^3$ per day of air, or ingestion of 2 L/day of water and a body weight of 70 kg).

The conversion equation most commonly used is:

Inhalation URF $(\mu g/m^3)^{-1}$ =

$$\frac{\text{CSF } (mg/kg/d)^{-1} \times 20\ m^3/d}{70\ kg\ BW \times 1000\ \mu/mg}$$

(BW = body weight)

Where different default values are used (e.g. those recommended in the *Australian exposure factor guidance* document – see Chapter 4), it may be necessary to adjust CSFs derived by the US EPA or in Integrated Risk Information System (IRIS) databases.

Recent US EPA (2009a) guidance (RAGS-F) indicates that inhalation risks should be assessed by calculating an exposure-adjusted air concentration, which is then used in EHRA risk characterisation equations. This means that exposure estimates are no longer adjusted based on inhalation rate or body weight, and the only difference in risk estimates between a child and an adult is the exposure time.

The application of these risk factors is to calculate the probability of a finite increase in cancer risk over a lifetime, according to the equation:

Increased lifetime cancer risk (ILCR) = chronic daily intake (mg/kg/d) × CSF $(mg/kg/d)^{-1}$

or

Increased lifetime cancer risk (ILCR) = exposure concentration × URF

The outcomes of cancer risk estimates based on CSF or URF calculations are the prediction of an increased lifetime risk of developing cancer. The intake estimate (or exposure concentration) must be averaged over the lifetime of expected exposure (default 70 years). The ILCR must be clearly presented so that the cancer estimate over 70 years cannot be misrepresented as an estimate of annual cancer risk.

To convert a lifetime to an annual risk estimate is approximated by a simple division by the standardised lifetime duration (70 years in most jurisdictions). In reality for cancer, incidence will be greater in the later part of the 70-year window.

The step-wise process for deciding on the dose–response data to adopt for the EHRA of carcinogens (or potential carcinogens) is set out in Figure 15 (Section 3.10.3). This decision-making process recommends use of a BMD approach to selecting a POD for risk assessment, once a decision has been made on classification of the COPC as a carcinogen and a carcinogenic risk assessment approach is warranted. Where appropriate BMD data is not available, alternative dose–response data should be sourced, which may include the use of CSF (for genotoxic carcinogens) and ADI/TDI (for non-genotoxic carcinogens).

5.8 AGE-SPECIFIC ADJUSTMENT FACTORS

The US EPA has directed particular attention to the possibility that early-life exposure to a carcinogen may exacerbate risk to the extent that the default approach based on a whole-of-life CSF or URF may not be sufficiently protective (US EPA 2005a). The guidance is consistent with reviews of animal carcinogenicity bioassays relevant to the assessment of early-life susceptibility to carcinogens (Hattis et al. 2004; 2005). US EPA guidance on early-life exposure to carcinogens for which a mutagenic mode of action (MoA) has been reasonably established has been summarised in the 2005 supplemental guidance (US EPA 2005c) and incorporated into the most recent RAGS-F guidance (US EPA 2009a p. 23).

The guidance indicates that an additional safety factor should to be applied to mutagenic carcinogens as follows:

- tenfold adjustment for exposures during the first 2 years of life
- threefold adjustment for exposures from ages 2 to <16 years of age
- no adjustment for exposures after turning 16 years of age.

Carcinogens identified by the US EPA as having a mutagenic mode of action (as of 2005) are discussed in US EPA 2005a. This list includes benzo(a)pyrene, and the additional safety factors recommended by the US EPA have accordingly been incorporated into the derivation of HILs for benzo(a)pyrene in the revision to the contaminated sites NEPM (NEPC 2010).

5.9 COMBINING RISK ESTIMATES

Where there are several exposure pathways, the incremental lifetime cancer risk (ILCR) estimate is simply summed for each of the relevant pathways to get a combined risk estimate (see Section 5.4). However, some caution should be exercised in adopting this simple summation approach (US EPA 1989). Since the CSF is an upper 95th percentile estimate of cancer potency, simple addition of 95th percentiles is strictly not correct. Such an approach can add unnecessary conservatism to the aggregate risk estimate. The CSF are not weighted according to the strength of evidence that underpins their categorisation. All classes of carcinogenic categorisation are given equal weight, including those where either human or animal data (or both) drive the categorisation.

There may be different CSF estimates for a single chemical where the cancer data relates to different tumour sites. The EHRA usually uses the CSF that predicts the highest risk. If cancer potency and/or the type of tumour produced differs according to the route of exposure, the aggregate risk may need to reflect this difference.

For example, in the case of benzo(a) pyrene (BaP), the representative carcinogenic polycyclic aromatic hydrocarbon (PAH), the CSF used for carcinogenic risk assessment in the

US IRIS database is 7.3 mg/kg/d^{-1} (range: 4.5–11.7 depending on the database used). The CSF used for BaP carcinogenicity assessment of B(a)P in the contaminated sites NEPM review (NEPC 2010) is 0.5 mg/kg/d^{-1} for the oral route. An estimated CSF by the dermal route (25 mg/kg/d^{-1}) has been derived in data reported by Knafla et al. (2006), although these estimates have yet to be endorsed by any regulatory agency.

5.10 TARGET RISK LEVELS

When the risk assessment uses a threshold approach (e.g. derivation of an ADI or TDI based on application of safety factors to a NOAEL or LOAEL), there are no implicit target risk levels associated with any derived environmental standards or guidelines. The use of safety factors to further reduce an exposure that is assumed to be without appreciable risk over the specified period of time (usually 70 years, representing an entire life span) is presumed to deliver an acceptable level of 'safety'.

For an individual COPC, the aim of the risk assessment is to determine whether the exposure exceeds an appropriate risk-based guideline value. This is generally expressed as a hazard quotient (HQ), defined as the ration of the exposure divided by the GV (see Section 5.4.1).

When the risk assessment involves combining exposures to multiple agents the outcome of the risk assessment may be a hazard index (HI). In such a process, the 'target HI' is generally assumed to be 1. Chapter 12 includes discussion of the possible implications when the HI is above or below 1.

When the risk assessment uses a non-threshold approach (as is the case for most quantitative carcinogenic risk assessments) it is implicit that any derived environmental standards will attempt to protect the community by minimising exposures to a point where a specified level of risk will not be exceeded.

While acknowledging that setting a level of 'acceptable risk' is often necessary for decision-making purposes, setting the numerical value for such a risk level is a socio-political matter, requiring extensive consultation with stakeholders, including the community likely to be affected by the environmental hazard and those responsible for managing or ameliorating the risks. A socioeconomic or cost–benefit analysis of the risk management options should also be part of the process.

It is therefore important that all parties appreciate the real meaning of a 'target' risk level – often expressed as something like 1×10^{-6} (one in a million). It should not be taken to imply certainty that one person will get the disease if there are at least one million people exposed. It is simply a way of expressing risk, as a numerical expression of the likelihood of an event occurring under the defined conditions of exposure, based on extrapolation of dose–response data. The precision of the risk estimate is subject to many assumptions about the validity of the extrapolation methods used, the degree of conservatism built into the modelling (for example, if upper-bound limits are used rather than means or central estimates). It is therefore important that risk communication strategies address the need to explain the meaning and derivation of such numerical estimates of risk, including its likely level of conservatism and the background to the development of the target risk levels.

Reinforcing the view that setting an 'acceptable' risk target for carcinogens is a process based on policy, rather than science, WHO (1994a) stated that:

> ... crude expression of risk in terms of excess incidence or numbers of cancer per unit of the population at doses or concentrations much less than those on which the estimates are based may be inappropriate, owing to the uncertainties of quantitative extrapolation over several orders of magnitude. Estimated risks are believed to represent only the plausible upper bounds and vary depending upon the assumptions on which they are based.

What sort of target risk levels have been used in various risk assessment paradigms? For carcinogens, a target risk level of 1×10^{-6} is the one most commonly used. The origin of the 10^{-6} level has been attributed to US regulators designating this level as a negligible or essentially non-existent risk, from the legal point of view that *de minimus non curat lex* (the law does not deal with trifles) and 10^{-6} is a convenient quantitative expression of the *de minimus* concept (Langley 2003). The current review of schedule B(4) of the contaminated sites NEPM cites a commentary by Kelly (1991) sourcing the origin to the arbitrary 10^{-6} risk level used by the US Food and Drug Administration (FDA) in the 1970s to regulate carcinogenic residues in foods. Irrespective of the source, the general understanding is that a 10^{-6} risk level is either essentially zero, or it can be considered negligible in a regulatory sense.

However, it must also be acknowledged that the target risk level has been varied upwards to between 10^{-5} and 10^{-3} in different types of risk management situations. The higher risk levels are more commonly found in occupational exposure settings, or associated with the evaluation of contaminated sites. For example, the Dutch 'intervention levels' for management of contaminated soils are based on a carcinogenic risk target of 10^{-4} (de Bruijn et al. 2001).

The 'target' risk level to which some Australian environmental regulatory authorities aim is 1×10^{-6}, although this may depend on whether the risk is associated contamination of air, water or food, or whether the exposure is associated with a single carcinogen or is the outcome of multiple chemical exposures. In the latter case, a combined risk of 10^{-5} may be considered acceptable. The revision of the contaminated site NEPM (NEPC 2010) proposes a carcinogenic risk 'target' of 10^{-5}, irrespective of whether a single or multiple chemical exposures contribute to the combined risk.

5.11 PRINCIPLES FOR SETTING RISK-BASED ENVIRONMENTAL HEALTH GUIDELINE VALUES

Elements of the risk assessment methodology described in this enHealth document provide a framework for setting risk-based environmental health GVs or choosing appropriate GVs from published national and international sources (see Sections 5.11, 5.12).

Toxicity reference values are generally based on the most sensitive toxicological effects, when known. When selecting toxicological reference values for comparison with exposure estimates, or when using established health-based environmental GVs, it is preferable to use values that have been endorsed or formally approved by the relevant Commonwealth, state or territory environmental health agencies. If there are no approved and relevant toxicological reference values or health-based environmental guideline values, advice should be sought from an administrative authority at the relevant Commonwealth, state or territory level.

Health-based GVs should be derived using toxicological data or exposure criteria from agencies or organisations relevant to the state or territory (e.g. local or Commonwealth health agencies such as NHMRC or enHealth, or to which Australia is party (e.g. WHO). Section 5.12 of this enHealth document provides some additional guidance on a hierarchy of preferred alternative sources.

Where health-based GVs are available from multiple sources, the assumptions, defaults and science policy that underlie them should be taken into consideration and should be compatible or generally similar to relevant Australian practice.

In setting GVs, apportioning exposures between media should take account of published relevant data on background exposures and consider policies and practices detailed in Section 5.11.1. This apportionment should be documented for each substance for which a GV is derived.

Generic GVs developed using this methodology require endorsement by appropriate national health bodies. Situation-specific environmental health criteria developed using this methodology require endorsement by the appropriate health agency before being applied to a particular situation. The use of 'sensitivity analysis' approaches (see Section 5.15) will provide some insight on the validity of the data inputs and the level of uncertainty in the derived criteria.

A wide range of issues are considered by those establishing risk-based environmental health GVs. A good understanding of how a particular set of GVs has been calculated and what policy drivers are included in the considerations is needed by those using GVs to ensure that the values chosen for a particular risk assessment are appropriate to the situation being investigated. Some of the key issues that may have been considered when establishing GVs are:

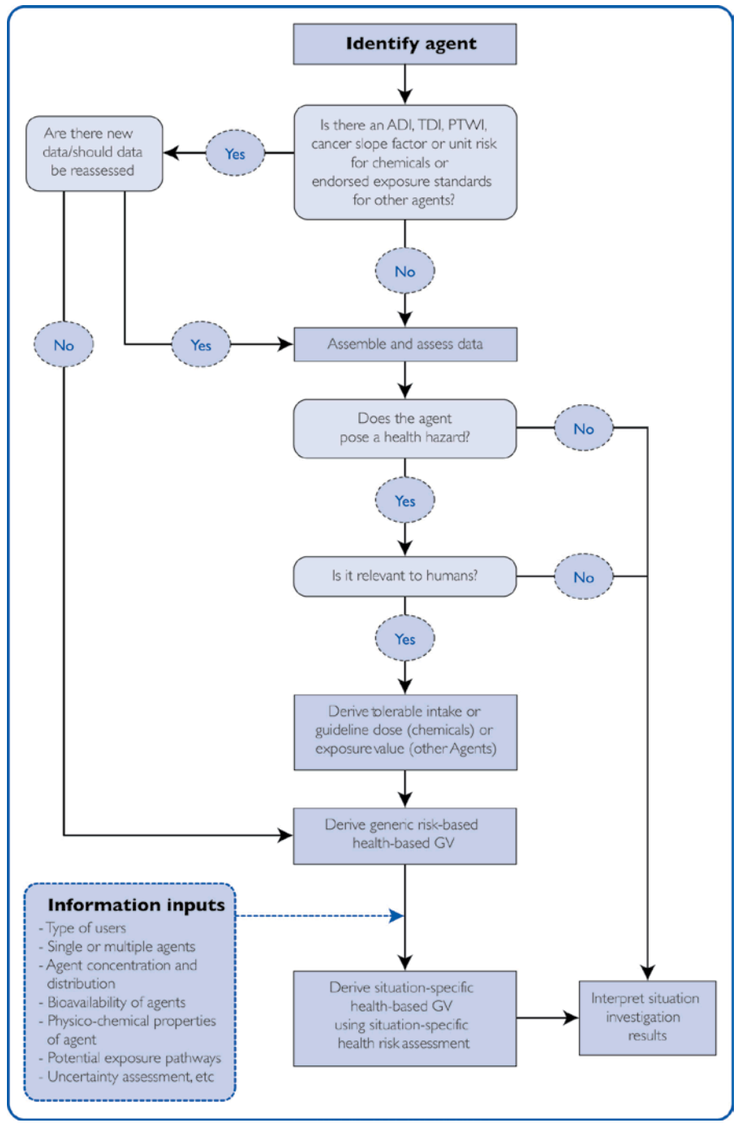
- Why is a GV being proposed and is it necessary?
- Are there alternative means of achieving the desired outcome? The large improvements achieved by the *Clean Air Act (1956)* in the UK occurred without any air quality standards.
- How will the GV be used? Is it simply a guideline or should it be considered to be a standard? Standards often have greater legal or regulatory standing than guidelines.
- What is the critical health effect? What is its nature, severity and reversibility?
- Are interactions with other agents relevant?
- Is there sufficient data to establish a GV?
- What population(s) will be affected?
- Are there any sensitive or particularly susceptible sub-populations who are exposed, and is the derived GV intended to be protective of these groups, or merely protective of the average person?
- Over what period of time will the population be exposed to the agent for which the GV is being set?
- Is the GV to be generic (applying to many situations) or situation-specific?
- What patterns of exposure are likely to occur? Are there likely to be short- or long-term fluctuations? Has the proposed guideline value appropriately matched the expected exposure patterns with the health-based data used to derive the benchmark?
- Are there difficulties in getting relevant and accurate background exposure data?
- What is the relationship between the derived tolerable intake and any background exposure? If the tolerable intake estimate is actually less than background, does this have implications for health consequences in the population, or is it because of the conservative way in which the tolerable intake has been derived? In such circumstances, should any actual experience or health information on the exposed community take precedence over a conservatively derived tolerable intake?
- What should be done if the tolerable intake is so conservative that it derives a guideline value so low that it cannot be measured or achieved by reasonable risk management practices? Should environmental health criteria ever be set to levels that cannot be measured by contemporary analytical techniques

and, if so, how can such criteria be managed in a regulatory context?

- How can tolerable intakes, which are usually based on ingestion, be applied to other exposure routes?
- Has the tolerable intake been set using gavage or bolus administrations rather than administration as part of the diet? Has the interpretation of such tests been compromised by the effects of a vehicle (e.g. an oil carrier) or the dietary matrix, such that the bioavailability of the test substance has been significantly altered?
- How is exposure occurring in the environmental setting, and to what extent do exposures vary over time and place?
- How do you deal with multiple pathways of exposure such as the relatively high exposures in tobacco smoke compared with dietary sources?
- Is the bioavailability of the substance known? While the amount of validated data on bioavailability from oral and dermal routes is steadily increasing, a conservative default assumption of 100 per cent bioavailability is usually assigned where such data is not available.
- Is there a background level of exposure to the agent from media other than the one for which GVs are being set? How do you apportion exposures associated with such background? How are these background or extraneous exposures managed in the standard-setting process? (See Section 5.11.1.)
- Who will be involved in setting the GV?

A decision tree detailing the use of EHRA to develop risk-based environmental health criteria is provided in Figure 24. This model uses a GV (e.g. ADI) that is apportioned between background exposures and exposures relevant to a particular exposure pathway (e.g. food, water, air or soil). This approach is based on chronic or subacute exposures. It may not be applicable for acute exposures such as when dealing with a respiratory irritant.

Figure 24: Decision tree for the development of risk-based environmental health GVs



A slightly different approach will be required where acute exposures may cause the hazard of concern to become manifest. Examples are: setting microbiological standards for water and food where acute exposures precipitate disease the setting of criteria for sulfur dioxide in air where acute exposures may cause exacerbation of asthma; and setting soil values for nickel or chromium (VI), both of which may cause allergic reactions from acute exposures in sensitised individuals. It is particularly relevant where the susceptible sub-population comprises a significant proportion of the total population, for example, approximately 20 per cent of Australian children have asthma and 10 per cent of women are allergic to nickel. In these situations, GVs (which, by definition, are based around long-term exposures) will be irrelevant and unsuitable for use. However, the remainder of the risk assessment process will still be relevant for establishing criteria.

For most criteria there is a significant margin of safety between the GV and typical exposures. Safety is enhanced by a further margin of safety arising from the process by which GVs are set. However, for some agents health effects may arise from background exposures such as exacerbation of asthma from urban ozone exposures or cardiovascular diseases associated with exposure to ultrafine particulates (PM_{10} and $PM_{2.5}$). In these instances, a risk management decision will need to be taken about the incidence of adverse health effects in the community. Further discussion about managing these types of environmental health problems with no apparent threshold is contained in the NHMRC guidance on setting ambient air quality standards (NHMRC 2006).

A useful source of equations used generally in toxicology is in Deralenko and Hollinger (1995), while equations used in EHRA processes specific for assessing contaminated sites is in Schedule B(7) of the contaminated sites NEPM (NEPC 2010).

5.11.1 Allocating background exposures in setting environmental health guidance values

Chemicals designated as of potential concern (COPC) in an EHRA may also occur naturally in the environment, including chemicals like benzene, chlorinated dioxins or petroleum hydrocarbons. The particular situation being assessed for risk may be due to the activities of people, but it will mean that any exposed population may well have exposure to a chemical outside that being considered in the risk scenario under consideration. It is important to consider these background exposures in a risk assessment to ensure that the situation being assessed does not result in exceedence of a TDI, for example.

Background exposure is also important to consider for chemicals that may not occur naturally but are now so ubiquitous that exposures from sources other than the scenario under consideration are inevitable.

For many agents, there may be several exposure pathways. Copper will be found in water and food (and, of lesser importance, in consumer products) so that setting a GV for copper in another medium (e.g. soil) or a particular foodstuff (e.g. shellfish) will need to take into account the range of other potential exposure pathways. The exposure route for dioxins likely to dominate in a risk assessment is the intake of fatty foods. While dioxins in soil, air and water may make only a small direct contribution to intake, these sources need to be considered in an EHRA in order to account for possible transfers into the food chain.

Intakes may need to be apportioned between the different exposure pathways. The apportioning of intakes raises other questions:

- What is the nature of the background exposures? Are they fixed or changing over time? Are they able to be altered? Are they voluntary (e.g. smoking) or involuntary (e.g. ambient air pollution)? To what degree should voluntary background exposures be taken into account?
- What is the background level of exposure to the agent from media other than the one for which GVs are being set? How are these background or extraneous exposures managed in the standard-setting process?
- How do you apportion exposures associated with such background? What percentage of the total tolerable intake should be used for establishing a set GV? Interagency cooperation will be required to enable appropriate apportionment.

The toxicity reference value (e.g. TDI) gives a measure of the total daily exposure to a substance that can be tolerated for a lifetime. In the derivation of health-based GVs, a proportion of the assumed total exposure to a substance is attributed to the exposure medium under consideration (e.g. soil, water, food, air) and the remainder to 'background'. Such background levels of exposure can be associated with the chemical concentrations present in the environment as a result of everyday activities (e.g. emissions from motor vehicles) or natural sources (e.g. dissolution of mineral deposits). Chemicals may be present in food, air, water and consumer products, and these all contribute to the quantity of the chemical that a person might be exposed to on a daily basis.

Apportionment of the toxicity reference value to a specific environmental medium can range from 100 per cent (representing negligible background exposure from any other source) to 20 per cent (i.e. 80 per cent to background) where there is a significant background intake from sources other than contaminated soil. The fraction attributable to the source medium

under consideration is then used to calculate the GV.

Various default values have been used in attributing proportions of the tolerable intake to specific environmental media. For example, in Canada, 20 per cent of total exposure is allowed from each of food, air, water, soil and consumer products in setting source-specific environmental criteria. US EPA guidance on aggregate and cumulative risk assessment (US EPA 2003a) does not include specific recommendations for attributing pathway-specific exposures, but it does suggest that where a pathway contributes less than 1–10 per cent of the population-adjusted dose (PAD), it may be discounted in an aggregate risk assessment. Other US guidance refers to a relative source contribution (RSC) factor, which defines the relationship between source-specific and background exposures. In the case of setting drinking-water guidance (US EPA 2000c), the RSC factor specified to account for non-water sources (except for carcinogens) was recommended to range from 20–80 per cent, with 20 per cent recommended as the default value in the absence of specific data. In some specific cases, where human exposure data has been available, the RSC has been set at 100 per cent.

There are some Australian precedents for assuming default values for attributable background exposures. For example, in setting water quality standards, the *Australian drinking water guidelines* (ADWG) (NHMRC-NRMMC 2004) allocates 10–80 per cent to drinking water as the source for setting ADWG GVs. The *Australian guidelines for water recycling: augmentation of drinking water supplies* (NRMMC, EPHC, NHMRC 2008) allocates 10 per cent to water as a default when the chemical has recognised industrial or other uses in Australia but between 20 and 100 per cent where there is no evidence of such use, allowing a greater proportion of intake to be attributable to water.

No specific attribution factor for airborne exposure has been determined in Australian guidance on setting air quality standards (NEPC 2009), but air quality environmental health criteria are often set on epidemiological data where background exposures would be accounted for in the measured health effects data. In guidance under development for management of contaminated sites (NEPC 2010), health investigation levels (HILs) for specific substances use soil attribution factors varying from zero (representing negligible background exposure form any other source) to 80 per cent, where there is a significant background intake from sources other than contaminated soil.

No specific recommendations for route-specific background exposure attribution factors (AF) were made in the original 2002 enHealth document, although such a factor was included in some of the equations recommended for calculating exposure such as the equation recommended in 2002 for the derivation of drinking-water GVs.

$$GL = \frac{RL \times BM \times AF \times ED}{IR \times UF \times AT}$$

GL = the guideline concentration of contaminant in water

RL = toxicity reference value (often a NOAEL)

BM = body mass, often 70 kg for an adult

AF = the proportion of total exposure attributable to drinking water

ED = exposure duration (if exposure is less than continuous)

IR = ingestion rate, often taken as 2L per day

UF = uncertainty factor applied to reduce the RL

AT = an averaging time of exposure (will be equal to ED if exposure is continuous)

Since the setting of default values for background is a somewhat arbitrary process, no specific recommendations for AFs are made in this update of enHealth guidance on EHRA. A case-by-case analysis of the available data should be made to determine what proportion of the tolerable intake estimate could be allocated to the exposure routes under consideration. If there is insufficient data available to make such a judgement, default values of 20 per cent for any specific pathway might be reasonable (as done in Canada), but it is again emphasised that reasoned analysis is preferable to using arbitrary default values.

This background attribution process is applied only to substances that demonstrate threshold dose–response characteristics. For substances exhibiting non-threshold characteristics (mainly carcinogens) background exposures are not considered in the EHRA, since the risk estimate output is extra risk (i.e. risk above background).

The adjustment in the calculation is:
$TDI_{(adjusted)} = (1 - background) \times TDI$

Where data suggests that background exposure is essentially negligible (contributing less than 5 per cent of the threshold TDI), the allocation to the specific exposure pathway of interest can be 100 per cent of the threshold TDI.

A complication arises when the background exposure exceeds the TDI because, in theory, a health-based GV that is sufficiently protective of human health cannot be calculated. A few approaches are available to address this problem. In setting soil-based GVs (equivalent to Australian HILs), the UK (UK EA 2009) arbitrarily allocates 50 per cent of the TDI to soil when the background actually exceeds 50 per cent. New Zealand guidance on contaminated site assessment (MfE 2010) allocates background on a case-by-case basis.

5.11.2 Using weight of evidence in setting environmental health guidance values

The NHMRC guidance document on setting ambient air quality standards (NHMRC 2006) also includes a useful discussion of the application of weight-of-evidence (WoE) criteria in standard-setting. It is equally applicable across the board in setting other health-based environmental criteria. The guidance notes that, where a range of evidence is available to support a standard-setting process, assessing the weight to be applied to any piece of evidence is a critical part of the analysis. In many cases, multiple sources of possibly conflicting data will be available, and choosing the most appropriate source for standard-setting becomes a crucial step. If the conflicting data results in different reference doses or concentrations derived by different regulatory authorities, it may be difficult to resolve the apparent conflicts unless there is sufficient clarity about the sources of data used, and how they were used to derive the values, and some judgement can be made about the relative weights to be given to the different data sources.

In reviewing the concepts that underscore WoE analysis, Weed (2005) noted there is no clear or universally accepted definition of what constitutes a WoE approach. Weed pointed out that WoE can be taken to include processes where the approach is:

- *metaphorical*, where it is implied that merely a collection of data or studies has been taken into consideration, but there is no clear indication of what type of weighting has been applied, or whether there has been any objective analysis
- *methodological*, where systematic interpretative analyses have been applied to a complete, or nearly complete dataset, such as systematic narrative reviews, meta-analysis or some other careful review of the quality criteria of the studies included
- *theoretical*, where the term is simply applied as a conceptual framework.

However, Weed (2005) noted that, in a review of 276 papers published between 1994 and 2004 (and which purported to apply a WoE analysis), approximately 50 per cent of them could be classified as simply applying a metaphorical meaning to the WoE concept, with no elaboration of how the concept was applied. This means that any study that professes to include WoE to underscore its outcomes needs to be carefully examined for the real meaning of the claim.

5.12 GUIDANCE ON SELECTING SOURCES OF TOXICOLOGICAL DATA AND ENVIRONMENTAL HEALTH CRITERIA

The original 2002 version of this enHealth document included a section that categorised sources of toxicological information and established environmental health criteria according to a hierarchy. This hierarchy, modified to include some additional sources, is reproduced in this section (5.12.1)

The sources are grouped into 'levels' in which the hierarchy generally signifies an order of preference. For example, it may be assumed that Australian guidance values accorded Level 1 status should take precedence over other sources, provided they are reasonably current. Other Level 1 sources may be more useful where it can be established that they are based on more recent data or set using more contemporary risk assessment methodologies. In this context, documents or GVs based on an internationally peer-reviewed risk assessment process (e.g. WHO, IPCS) may be accorded a higher status than Australian sources under some circumstances.

It may be that some international data sources may be preferable in the Level 1 hierarchy on the basis that they address specific aspects of toxicity in a more detailed way. For example, International Agency for Research on Cancer (IARC) monographs (when current) may present a deeper and more structured appreciation of the carcinogenic potential for a specific COPC than other documents.

While this section primarily refers only to sources of toxicological and TDI data, rather than derived GVs, there may be a need to encompass such GVs where they have been based on a transparent risk assessment process. In line with this philosophy, published Australian ADIs should be used in risk assessments when available, but other data may be used with appropriate justification. Different agencies are likely to have used differing risk assessment and standards-setting methodologies and these differences should be appraised by the risk assessor. All documents, particularly those in the second and third levels, require rigorous appraisal for relevance, validity and accuracy. An important consideration is the frequency with which documents and GVs are updated via a periodic review process.

5.12.1 Level 1 sources

- National Health and Medical Research Council (NHMRC) documents and documents from other joint Commonwealth, state and territory health organisations. These may be a source of Australian guidance values
- ADI list from the relevant office of the Australian Government Health portfolio (currently the Office of Chemical Safety – OCS) – regularly updated at <http://www.health.gov.au/internet/main/publishing.nsf/Content/ocs-adi-list.htm>
- World Health Organization (WHO) documents. Australia is a party to the WHO process and has incorporated its material in a variety of environmental health criteria. A range of documents include those from the WHO/ILO/

UNEP International Programme on Chemical Safety (IPCS), which produces environmental health criteria monographs and concise international chemical assessment documents (CICADs). Documents detailing international acceptable daily intakes (ADIs), tolerable daily intakes (TDI) or tolerable weekly intakes (TWI) may be found in evaluations by the WHO/FAO Joint Meeting on Pesticide Residues (JMPR) and by the Joint FAO/WHO Expert Committee on Food Additives (JECFA). The WHO has also published guideline documents on ambient and indoor air quality (WHO 2000b, 2010)

- enHealth Council documents
- National Environmental Health Forum documents distributed by the Commonwealth Department of Health and Ageing
- International Agency for Research on Cancer (IARC) monographs
- WHO/FAO Joint Meeting on Pesticide Residues (JMPR) monographs
- NICNAS Priority Existing Chemical (PEC) reports
- US Agency for Toxic Substances and Disease Registry (ATSDR) documents for general toxicological reviews and reference doses
- National Toxicology Program (NTP) carcinogenicity appraisals, which report in detail the results of carcinogenicity tests on a wide range of chemicals
- OECD Standard Information Data Sets (SIDS) and SIDS initial assessment reports (SIAR)
- EPA reference doses, cancer slope factors and unit risk factors
- TOXNET – sets of toxicology databases maintained by the US National Library of Medicine (includes TOXLINE, CCRIS, GENETOX, IRIS)
- The International Toxicity Estimates for Risk Assessment (ITER) database – a compilation of risk data from multiple sources (including some respected national risk assessment databases), first published by Toxicology Excellence for Risk Assessment (TERA), a non-profit US organisation. ITER is now also available via the TOXNET network
- A special issue of *Toxicology Applied Toxicology* (Vol 233, Nov 2008) included several papers on databases useful in risk assessment (see Waters & Jackson 2008; Wexler 2008; Woodall 2008; Woodall & Goldberg 2008; Wullenweber et al. 2008)
- Dutch National Institute for Public Health & the Environment (RIVM) reports and databases.

5.12.2 Level 2 sources

Peer-reviewed journals: These may provide opinions that do not meet general scientific agreement. and may be a suitable source where no GVs have been established by competent authorities.

Industry or commercially based publications: With justification, and acceptance by the local jurisdiction, the following may be suitable for use:

- European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) monographs, JACC reports and technical reports
- Chemical Industry Institute of Toxicology (CIIT) reports (now the CIIT Centers for Health Research, part of the Hamner Institutes of Health Sciences at Research Triangle Park, North Carolina
- International Life Sciences Institute (ILSI) and ILSI – HESI (Health & Environmental Sciences Institute) workshop and technical/project committee reports
- Toxicology Excellence for Risk Assessment (TERA) – a non-profit organisation that publishes reports on methodologies and selected issues/ chemicals and supports training in risk assessment
- Alliance for Risk Assessment (ARA) – a consortium of organisations formed to promote collaboration on risk assessment issues – maintains the Risk Information Exchange (RiskIE) database, which catalogues reports on risk assessment documents from various governmental and non-governmental sources
- Scientific societies and organisations that specialise in toxicology and risk assessment issues – the US Society of Toxicology (SOT) has long been a leader in developing and evaluating toxicological and risk assessment methodologies, and maintains specialist sections that undertake projects to advance risk assessment methodology. The Society for Risk Analysis (SRA) is another society that has played a leading role in advancing risk assessment and risk communication methodologies
- Unpublished industry reports submitted for regulatory purposes – these may have restricted availability but information may be available in evaluation reports from regulatory agencies that have reviewed individual reports
- Occupational health and safety sources – these may be a useful source for toxicological data and reviews, but occupational exposure criteria must not be used in a general public health context without appropriate adjustment for the different durations of exposure, the inclusion of susceptible sub-population in the general community (e.g. children) and the methodological differences in the setting of criteria.

5.12.3 Level 3 sources

These are sources not covered in Levels 1 and 2. Using this information requires justification that no other sources are available and an appraisal of the methodology detailing the level of conservatism and range of uncertainties inherent in the approach.

With justification, and acceptance by the local jurisdiction, they may be suitable to use if no GVs are available.

5.12.4 Selecting data sources

The reliability of the data used in a risk assessment should not be based solely on the position of its source in the above hierarchies. The most important consideration is that the data is supported by sound science and contemporary risk assessment methodologies, and that a WoE approach has been used to assess their worth. This means that risk assessors should not simply rely on 'looking up numbers' but should gain an appreciation of the currency of the source and the underlying science from which the numbers have been derived.

5.13 CHEMICALS WITH NO AVAILABLE TOXICITY DATA

When no suitable toxicological data is available for a COPC, it will be difficult to establish an appropriate toxicological reference value (ADI or TDI), and this will impact on the ability to derive an appropriate health-based GV. In the absence of a suitable GV, there may be grounds for assuming a level of exposure to a COPC that would be of no toxicological consequence, and hence the COPC could be discarded or discounted in an EHRA. This concept has been variously described as the threshold of toxicological concern (TTC) approach or the concentration of no toxicological concern (CoNTC) approach (Drew & Frangos 2007).

The thresholds determined using the TTC concept are intakes of chemicals below which a given compound of known structure is not expected to present a toxicological concern. On the basis of classical pharmacological and toxicological concepts of dose–response, exposure to trace levels of chemicals represents very low risks if chemicals act via more generalised modes of action. TTCs have been developed for classes of substances with a systemic mode of toxicological action and with exposure by ingestion.

The TTC approach was for many years put forward as a pragmatic solution for addressing low concentrations of additives in food (Frawley 1967; Munro 1990; Munro et al.1996). While it was first applied in a regulatory sense by the FDA (FDA 1995; CFR 2001) and was later used by the European Commission (EC) (2003) to address chemicals migrating from plastic packaging into food, it is now applied in a number of regulatory activities (see Table 11).

Renwick (2004; 2005) has described the JECFA application of the TTC concept to the safety evaluation of flavouring agents. Since 1996 some 1,200 compounds have been assessed using the TTC concept.

The TTC concept has also been adapted for:

- deriving criteria for soil risk management for chemicals of unknown toxicological hazard or potency at contaminated sites (Wilson et al. 2000)
- judging whether ingredients at low concentration in personal and household-care products require toxicological testing (Blackburn et al. 2005)

Table 11: Current uses of the TTC approach in chemicals regulation

Regulatory agency	Use	References
United States Food and Drug Administration (FDA)	*De minimis* level (i.e. level of minimum importance) for regulation of minor contaminants (i.e. chemicals in food packaging materials that can migrate). Threshold of toxicological concern (TTC) is applied as a threshold of regulation for indirect food additives. The US FDA has dealt with 183 applications under this regulation and issued 78 exemptions using the TTC concept (Barlow 2005).	FDA (1993ab, 2006)
Joint FAO/WHO Expert Committee on Food Additives (JECFA)	Evaluation of flavouring substances. TTCs for different structural classes have been used for the safety evaluation of over 1,200 flavouring substances.	JECFA (1993; 1995; 1999), Munro et al. (1999), Renwick (2004; 2005)
European Food Safety Authority (EFSA)	Evaluation of flavouring substances	EFSA (2004)
European Medicines Agency (EMEA)	Assessment of genotoxic impurities in pharmaceutical preparations. See also Dolan et al. (2005)	CHMP (2004)
European Commission, Joint Research Centre	The TTC principle has been endorsed as a mechanism for the regulation of chemicals under draft chemical legislation reforms being considered by the European Union.	EC JRC (2005)

- establishing ADIs for chemicals likely to be present as impurities in pharmaceutical manufacturing (Dolan et al. 2005)
- use as a screening tool for risk assessment of air toxics (Drew & Frangos 2007)
- (more recently in Australia) setting concentrations that would be unlikely to pose a human health risk for chemicals likely to be present in recycled water (NRMMC, EPHC, NHMRC 2008).

The following description of the TTC approach is taken from the *Australian guidelines for water recycling* (NRMMC, EPHC, NHMRC 2008).

In establishing TTCs for chemicals that are not carcinogens, an evaluation of toxicological databases undertaken for non-carcinogenic endpoints is used (Munro et al.1996; 1999; Kroes et al. 2000; 2004). In these evaluations, some 900 non-carcinogenic organic chemicals were assigned to three 'classes' based on their chemical structure, presence of structural alerts for toxicity and known metabolic pathways, according to the classification scheme of Cramer et al. (1978). The Cramer classification scheme divides chemicals into three classes according to their predicted toxicity as judged from structural alerts and metabolism:

- *Class I:* substances of simple chemical structure with known metabolic pathways and innocuous end products that suggest a low order of toxicity
- *Class II:* chemical structures that are intermediate they are chemicals that are less innocuous they may contain reactive functional groups but do not contain the structural features suggestive of toxicity
- *Class III:* chemicals for which structural features or likely metabolic pathways permit no strong presumption of safety, or may even suggest significant toxicity.

The 5th percentile NOEL (no observed effect level) of each of the three Cramer classes was divided by an uncertainty (safety) factor of 100 to yield TTC values that are somewhat higher than those created by the FDA for carcinogens. No formal stratification of toxicological endpoints was used in establishing NOAELs for the three Cramer chemical classes. The NOAELs are:

- *Class I:* 3 mg/kg/day (equates to a TTC of 30 µg/kg bw/day)
- *Class II:* 0.9 mg/kg/day (equates to a TTC of 9 µg/kg bw/day)
- *Class III:* P 0.15 mg/kg/day (equates to a TTC of 1.5 µg/kg bw/day).

In applying the TTCs to derivation of drinking-water guidelines, a more conservative approach has been applied to reflect the use of safety factors used in the ADWG (NHMRC, NRMMC 2004). These guidelines apply a safety factor of 1,500 to organic chemicals (95th percentile). To achieve this, an additional safety factor of 15 has been applied in converting TTCs (which already include a safety factor of 100) to drinking-water guidelines.

5.13.1 Can the TTC approach be applied to carcinogens?

A generic approach has been developed for potentially genotoxic carcinogens using the TTC approach (NRMMC, EPHC, NHMRC 2008).

The FDA (1995; CFR 2001) regulatory TTC is based on a carcinogenic potency database of more than 500 chemicals examined in more than 3,500 experiments. The FDA (1995; CFR 2001) and other leading researchers (Munro et al.1996; 1999) have concluded that, if no toxicological data is available to derive a health-based guideline for a chemical, intakes of 1.5 µg/person/day (0.02 µg/kg bw/day for a body weight of 70 kg) are unlikely to result in appreciable health risk, even if the substance was later found to be a carcinogen. According to Munro (1990), a daily intake at the TTC of 0.02 µg/kg bw corresponds to a 96 per cent probability that the lifetime risk of cancer would be less than one in a million (1×10^{-6}).

The TTC that is protective of cancer endpoints is termed a 'generic TTC', to differentiate it from the TTC developed for non-cancer endpoints and using the Cramer classification. The TTC estimate of 0.02 µg/kg bw/day is conservative, erring on the side of safety, because of the numerous compounding conservative assumptions used to derive the low-dose cancer risk estimates (Barlow et al. 2001; Kroes et al. 2004).

5.14 QSAR AND READ ACROSS TECHNIQUES

Better understanding of structure–activity relationships (SAR), especially when combined with quantitative information (quantitative structure–activity relationship – QSAR) may facilitate prediction of toxicological properties of chemicals without testing, or where no testing has been done to establish the toxicological profile of a new chemical. There will also be consequent benefits in terms of lower costs, shorter testing time frames and less use of animals. QSAR may also be useful in complementing the increasing use of *in vitro* and *in silico* technologies to provide insights into toxicological properties of chemicals without using live animals. As well as facilitating chemical and drug development by industry, regulatory recognition of QSAR is also growing in importance. For example, it has been suggested that up to 10 per cent of new chemical notifications in the UK include QSAR data, and this proportion is expected to grow over time.

The implementation of the REACH (Registration, Evaluation, and Authorization of Chemicals) program in Europe, as well as the increasing number of high production volume (HPV) chemicals requiring assessment in OECD programs, should provide impetus for the further development and use of SAR and QSAR techniques.

SARs already utilised in regulatory toxicology include:

- recognition of structural elements that act as alerts for particular types of toxicological behaviour (e.g. epoxides or other reactive metabolic intermediates which that confer DNA- and protein-interactive capabilities)
- recognition of common structural elements in chemical classes that are consistent with known patterns of toxicity (e.g. organophosphonate groupings that enable phosphorylation of the active site on acetylcholinesterases)
- grouping of chemicals based on recognisable structural features that lead to common toxicological properties (e.g. dioxin-like chemicals and others that interact with aryl hydrocarbon (Ah) or peroxisome proliferator (PPAR) receptors)
- computational systems that use a combination of features of the molecule (electronic, physico-chemical, size, hydrophobicity, etc.) to predict properties (e.g. EPIWIN)
- knowledge-based or rule-based systems that compare many parameters of a dataset and enable predictions of the properties of chemicals that share common structural features. One such commercially available system is DEREK, a computer-based SAR program (Sanderson and Earnshaw, 1991), although its utility is mainly limited to predicting sensitisation and carcinogenic properties.

Another alternative approach when data on a specific chemical is lacking is to use 'read across' techniques to make informed predictions about the toxicity profile from a known, and closely related chemical. Read across is primarily useful for hazard prediction. It has limited capabilities for predicting quantitative dose–response behaviour. It relies on there being a high-quality toxicological dataset for the reference compound.

5.15 UNCERTAINTY AND SENSITIVITY ANALYSIS

5.15.1 General

At the completion of a risk assessment, it may become apparent that there are inherent limitations to the outcomes, such as:

- information gaps (e.g. effects of mixtures, low-level and variable exposures over time, relative contributions of lifestyle factors versus other environmental hazards, variations in sensitivity)
- poor exposure information (e.g. complex mixtures of hazards with complex behaviours in the environment, limited knowledge about the actual or potential population and sensitive sub-populations geographic, variations in exposure)
- limitations of toxicological and epidemiological research (e.g. small populations, limited exposure information, multifactorial causes of many diseases)
- 'background noise' affecting research into common diseases or symptoms, population heterogeneity, and the fact that it is expensive and time consuming.

Some of these limitations may be apparent before beginning the risk assessment process. For example:

- the large number of combinations of hazards, exposures and health states leading to complexity that cannot be readily resolved
- complex causality for many of the health conditions addressed in the EHRA
- confidentiality of health and commercial information preventing full disclosure
- the atmosphere of fear, antagonism and distrust being so charged that it inhibits meaningful dialogue between the stakeholders.

In formulating an EHRA report it is crucial that all uncertainties and knowledge gaps be acknowledged and guide the development of risk management options (see Chapter 7). It is also important that these uncertainties be managed in a consistent and scientifically defensible way, and that there is a clear explanation of how 'scientific judgement' may have been applied to the management of these uncertainties. This may include careful description of definitions of default parameter inputs or using more complex probabilistic approaches to defining bounding values, or intervals within which the risk assessor expects the best estimates of risk to lay.

It may be important to carry out proper sensitivity and uncertainty analyses so the level of effort expended in an EHRA can be appropriately matched to the precision of the desired outcomes (NRC 2008). If the outcomes or advice to the risk manager will not be materially affected by adopting more simplistic approaches, it may be wasteful of scarce resources to use more sophisticated methodologies (e.g. deterministic versus Monte Carlo assessment of exposures). Similarly, the sophistication of analytical techniques used to measure environmental concentrations should be matched to the level of precision required in the EHRA.

5.15.2 Uncertainty analysis

Uncertainty in health risk assessment is the lack of knowledge about the correct value such as a specific exposure measure or estimate. Uncertainty is

distinguished from variability, which refers to true differences in attributes due to diversity or heterogeneity; variability cannot be reduced by further measurement or study, although it can be better characterised (NRC 2008).

Both uncertainty and variability contribute to uncertainty in the estimation of risk and should be adequately assessed in a risk assessment. Such consideration needs to be done transparently so that all users of the risk assessment can understand the approach taken.

An analysis of the uncertainty in the risk assessment is important because of the following:

- Information from different sources carries different kinds of uncertainty and knowledge of these differences is important when uncertainties are combined for characterising risk.
- The risk assessment process, with risk management input, involves decisions regarding the collection of additional data (versus living with uncertainty). In the risk characterisation, a discussion of the uncertainties will help to identify where additional information/data could contribute significantly to reducing uncertainties in the risk assessment.
- A clear and explicit statement of the strengths and limitations of a risk assessment requires a clear and explicit statement of related uncertainties (US EPA 1995b).
- Characterising uncertainty in a risk assessment informs the stakeholders about the range of possible risks from an exposure. Risk estimates may sometimes diverge widely (NRC 2008).
- Characterising the uncertainty in a risk assessment associated with a given decision informs the decision maker about the range of potential risks that may result from the decision (NRC 2008).

Uncertainty analysis is generally a qualitative process; however, in some cases it can be semi-quantitative or quantitative.

The first step should be a consideration of the conceptual site model and what aspects of that model are uncertain and how that uncertainty has been accounted for.

The second most important part of the uncertainty assessment is an evaluation of the uncertainty and variability in the data available relating to the site, situation or activity being assessed. Data will always be limited. However, the risk estimates based on even quite limited data can be fit for purpose if the exposure concentrations are a long way below (or above) toxicity reference values which indicate that the risks are very low (or very high). Decision making based on such uncertain but quite clear results is straightforward. Where risks are close to or slightly above the relevant toxicity reference values or 'target risk' level (the 'grey' zone), the issue of the uncertainty and variability in the data becomes much more important and so the uncertainty assessment needs to be more detailed.

When assessing risks, uncertainty can arise from missing or incomplete information, be incorporated into the scientific theory behind the model used to make predictions, and factors affecting a particular parameter, for example, errors in sampling. Such uncertainty has the potential to cumulatively overestimate or underestimate risk during an assessment. An assessment of uncertainty is a part of the health risk assessment process and consequently must be addressed for each step of the risk assessment and for its cumulative effect from all of the steps.

There are three broad types of uncertainty (US EPA 1992):

- *Scenario uncertainty:* uncertainty arising from missing or incomplete information such as descriptive errors, aggregation errors, errors in professional judgement, and incomplete analysis.
- *Parameter uncertainty:* uncertainty affecting a particular parameter such as measurement errors, sampling errors, variability, and use of generic or surrogate data.
- *Model uncertainty:* uncertainty in scientific theory affecting the ability of a model to make predictions.

NRC (2008) provides a detailed evaluation of the techniques currently provided for in US EPA guidance and concludes that although a number of usable methodologies are provided, it is unclear what level of detail is required to capture and communicate key uncertainties. A further comment is that quantitative methods suffer from the difficulty in sensibly quantifying all uncertainties, and that the apparent precision of quantitative analysis for some uncertainties may distract attention from other, possibly equally important but unquantifiable, uncertainties.

In most health risk assessments, it is unlikely that quantitative uncertainty analysis will provide value given the effort required to undertake it. A clear qualitative analysis is considered sufficient in most cases to provide the communication of the effects of uncertainty that is necessary.

NRC (2008) and IPCS (2008) provide useful guidance on the principles to be adopted for uncertainty analysis; these have been adapted for specific relevance to the enHealth document.

- Risk assessments should provide qualitative (as a minimum) or quantitative description of uncertainty and variability consistent with available data. The information required to conduct detailed uncertainty analysis may not be available in many situations.
- Sensitive sub-populations should be considered to the extent that they are not covered by the selected toxicity criteria (generally they will be).
- The uncertainty analysis should seek to communicate which uncertainties are most important to the conclusions of the risk assessment.
- The level of detail in the uncertainty analysis should be commensurate with the scope of the risk assessment.
- Uncertainty analysis should be expressed in terms that can be understood by the risk manager and other stakeholders.
- Uncertainty and variability should be kept conceptually separate.

The combination of uncertainty in the scientific data and assumptions (the 'inputs') and inability to validate assessment results directly or to isolate and evaluate the impact of a resulting decision (the 'outputs') creates a situation in which decision makers, the scientific community, the public, industry and other stakeholders have little choice but to rely on the overall quality of the many processes used in the conduct of risk assessment to provide some assurance that the assessment is aligned with societal goals (NRC 2008).

5.15.3 Sensitivity analysis

Sensitivity analysis is an important final step in the risk characterisation process, especially where modelling has been used to determine important components of the EHRA (see Section 8.7.4). It provides a quantitative estimate of the effect of uncertainty and/or variability in the input parameters on the results of the risk assessment and it should be undertaken when a risk assessment is conducted using a deterministic exposure model.

While a single value must be entered for each parameter in a deterministic model, it is unlikely that reasonable inputs for each parameter can be limited to a single value. This may be due to uncertainty and/or variability. A range of reasonable values will be defined as appropriate for a given input parameter. Sensitivity analysis is the process of changing variables used as input parameters one at a time to determine how such changes influence the final output. Variables are changed within a defined range while leaving the others constant and determining the effect on the output – the risk estimate. The procedure involves fixing each uncertain quantity, one at a time, at its credible lower bound and then its upper bound (holding all other at their medians), and then computing the outcomes for each combination of values (US EPA 1992). It can be used to test the effects of both uncertainty and variability in input values. The substitution of input parameters should be informed by knowledge of the upper and lower bounds of the expected parameter distributions.

Sensitivity analyses can be used to identify the most important input variables (or groups of variables) that are critical to the outcome of the risk assessment. It follows that variation of some inputs may have inconsequential effects. Sensitivity analysis can develop bounds on the distribution of exposure or risk. A sensitivity analysis can also estimate the range of exposures or risk that result from combinations of minimum and maximum values for some parameters and mid-range values for others (US EPA 1989). Effort may then be directed to the collection of additional data for these important variables; as additional data is collected, the uncertainty in the 'true' value is reduced, and it may be possible to define a smaller range for a given parameter. The uncertainty in the results of the risk assessment may therefore be reduced.

All risk assessments where conclusions are derived using modelling should incorporate a sensitivity analysis and describe the variability in the model outputs generated by plausible variation in the inputs. Note that some input variables may be connected and unable to vary independently. Monte Carlo models, where inputs are described by probability distribution functions, provide probability distribution function outputs.

The Monte Carlo method reduces the requirement for sensitivity analysis but may not eliminate it, depending on the model used.

5.16 INTERPLAY OF SCIENTIFIC JUDGEMENT AND SCIENCE POLICY

The interplay between these processes and the importance of providing appropriate explanation of assumptions, the use of scientific judgement, and the overlay of 'science policy' considerations is illustrated in the ebb and flow of regulatory actions and interpretations surrounding the presence of chloroform in drinking water and its surrogacy as an indicator of disinfection by-products (Box 1).

BOX 1:
Chloroform – New evidence can eventually change regulatory perspective, albeit slowly

(Adapted from Hrudey 2009).

Chloroform and the related trihalomethanes (THMs) were first identified as by-products of chlorine disinfection by Johannes Rook, a Dutch water chemist (Rook 1974). Rook had consistently identified chloroform in treated, but not raw water samples. He chose not to publish the identity of the large chloroform peak until he had figured out what was causing its formation but he was not troubled about consumer health risk, noting (Symons 2001): 'Our health officer told us chloroform was a normal constituent of cough syrups and was not known to be particularly toxic.'

There was also originally little health concern about chloroform at US EPA (Symons 2001) because of the widespread use of chloroform in consumer products, but they confirmed finding higher levels of THMs with increasing chlorine contact during disinfection (Bellar et al. 1974). The concern about chloroform only started to escalate when it was recognised that THMs were being formed from the reaction of chlorine with natural organic material (NOM), a constituent that is ubiquitous in surface water supplies.

Shortly after the growing body of evidence showing chloroform appearing in chlorinated drinking water supplies, the National Cancer Institute (NCI) published results of a rodent cancer bioassay on chloroform NCI (1976). This bioassay was conducted in accordance with the practices of that day, i.e. it was designed to determine the potential for chemical substances to cause cancer in mammals and were designed to maximise the ability of the experiment to reveal any carcinogenic effect by using the maximum tolerated dose.

Dosing in this experiment was done as a daily bolus dose of chloroform dissolved in corn oil. The initial high dose in female rats of 250 mg/kg(bw)-d had to be reduced to 180 mg/kg(bw)/d after 22 weeks because of the frank toxic effects that were observed. Mice proved more tolerant to chloroform, so their initial doses of 200 and 400 mg/kg(bw)-d were increased after 18 weeks to 300 and 500 mg/kg(bw)-d. For comparison, a human dose of chloroform equivalent to the highest dose rate would correspond to more than 25,000 times the daily dose achieved by consuming 2 L per day of drinking water containing 100 µg/L of chloroform daily for a lifetime. Furthermore, delivering a slug dose once per day in a vehicle like corn oil provides a higher peak loading to the liver than consuming water with an equivalent dose of dissolved chloroform spread out over a day.

The results of this high dosing showed strong evidence of liver tumours in mice (98% of males and 95% of females at lifetime average doses of 277 mg/kg/d and 477 mg/kg/d, respectively, 36% of males and 80% of females at lifetime average doses of 138 and 238 mg/kg-d respectively) in the mouse experiments. These high dose levels were from 27 to 115% of published median lethal doses (LD_{50}) for the mouse (Hill et al. 1975), suggesting that the $B6C3F_1$ strain of mouse used in these cancer bioassays was unusually tolerant of chloroform. The rats dosed at up to 200 mg/kg/d failed to show a significant excess of liver tumours relative to controls. Rats did show a significant increase in kidney tumours, but mice did not.

Within four months of the publication of the NCI bioassay results, the US Food and Drug Administration banned the use of chloroform in cosmetics. This was a dramatic change in relation to chloroform, which had been widely used as an anaesthetic from the mid-1800s into the early 1900s.

Health concerns associated with chloroform and THMs rapidly led to the adoption of drinking water guidelines and standards; Canada was first in 1978 to adopt a guideline maximum value for THM4 of 350 µg/L. Then in 1979, the US adopted a regulatory standard for THM4 under the Safe Drinking Water Act of 100 µg/L as a running annual average of four quarterly samples. In 1984, WHO proposed a guideline for chloroform of 30 µg/L based on an estimate that this would assure less than a 1 in 100,000 lifetime cancer risk assuming a linear extrapolation to zero dose. The *Australian drinking water guidelines* established a guideline value of 250 µg/L in 1996, based on the NOAEL for kidney toxicity in a 90d rat study, concluding that 'In view of the safety factors used in the derivation of the guideline value, it is unlikely that short-term consumption of water containing significantly higher concentrations of trihalomethanes would pose a health risk.' This value was confirmed when reviewed in 2004 (ADWG 2004, 2010).

The initial extremely high dose bioassays results on chloroform (NCI 1976) provided the case that was widely cited throughout the late 1970s and early 1980s that chloroform and, by extension, THMs, were carcinogenic. The NCI (1976) results were obtained by a method (high dose of chloroform in corn oil) that was later found to be much more toxic than the equivalent dosing of chloroform in water (Bull et al. 1986). The comparison of corn oil versus water as a vehicle was undertaken to explain the results from a study providing high concentrations of chloroform (up to 1,800,000 µg/L) dissolved in drinking water (Jorgenson et al. 1985) that produced no significant carcinogenic response.

The impact of extremely high doses of chloroform in corn oil to the liver was first noted as evidence of cytotoxicity on liver cells. Larson et al. (1994, 1995) demonstrated by direct experimentation that the corn oil gavage delivery of chloroform induced cytotoxicity and cell proliferation in liver for mice and kidney and liver for rats. The mouse experiments found this effect for the corn oil gavage, but not for direct delivery of similar daily doses orally by drinking water. These findings on a plausible mechanism for chloroform carcinogenicity were supported by extensive evidence showing virtually no mutagenic activity for chloroform (Golden et al. 1997). The earlier noted distinction in mechanism of tumour formation from cytotoxicity rather than genotoxicity justifies a threshold approach to risk assessment rather than a no-threshold approach for THMs (Fawell 2000).

According to the US SDWA the maximum contaminant level goal (MCLG) is the maximum level of a contaminant in drinking water at which no known or anticipated adverse health effects would occur, and which allows an adequate margin of safety. US EPA policy for carcinogens in drinking water had required a MCLG to be zero, apparently ignoring the possibility of a non-genotoxic carcinogen having a threshold. However, the foregoing toxicological evidence on the mode of action of chloroform resulted in a US EPA expert review panel recommending the abandonment of the MCLG of zero and replacement with a limit based on an estimated threshold. Thus in 1998, the US EPA (1998a) proposed to raise the MCLG to 300 µg/L in accordance with this expert advice. Because many intervenors protested this precedent-setting measure, the US EPA (1998b) Final Rule withdrew the proposal to change the MCLG for chloroform from zero to 300 µg/L (Pontius 2000).

The Chlorine Chemistry Council sought a court review of the US EPA decision because the SDWA requires the US EPA to use the best available science in setting standards and regulations. Although the US EPA acknowledged that the best available science called for raising the MCLG above zero, it had nevertheless decided to retain the zero MCLG. On 31 March 2000, the US District Court ruled that the US EPA had violated the SDWA by failing to use the best available science. The court found that the EPA action of setting the MCLG of chloroform at zero to be 'arbitrary and capricious' and in excess of statutory authority. The US EPA withdrew the zero MCLG in May 2000, subsequently replacing it with a MCLG of 70 µg/L in 2003.

Meanwhile, WHO had changed its drinking water guideline for chloroform from 30 µg/L in its first edition of *Guidelines for drinking-water quality* (WHO 1984) to 200 µg/L in the second edition (WHO 1993a). The rationale associated with recognising that chloroform exhibited a threshold for acting as a carcinogen was invoked to justify this change.

The initial NCI (1976) carcinogenic finding on chloroform, taken together with the background expectation that substantial numbers of human cancers could be explained by environmental contamination, resulted in more than 65 epidemiology studies of varying quality from 1977 to 2008, seeking to determine if some measure of chlorination DBPs was associated with an increase in one or more types of cancer. The epidemiologic evidence regarding cancer has been reviewed at various times (IARC 1991, Mills et al. 1998, ICPS 2000, IARC 2004). Overall, the epidemiologic evidence has generally been found to be insufficient to declare chlorination DBPs to be carcinogenic in humans. The evidence for colon and rectal cancer has been suggestive of a causal association while the evidence for bladder cancer has been clearly the most consistent, providing the greatest likelihood of being causally associated with chlorination DBPs (Mills et al. 1998). There is now common understanding among experts on DBPs and health evidence that chloroform in particular and THMs in general are at best surrogates for some DBPs as yet unidentified in chlorinated drinking water that may pose a drinking water cancer risk. Despite the original focus on chloroform and THMs as carcinogens in drinking water, more than 30 years of evidence now verifies that these chemicals do not pose a cancer risk at realistic drinking water exposure levels (Hrudey et al. 2003).

Chapter 6: Community engagement

6.1 INTRODUCTION

Engaging with stakeholders as part of the EHRA risk process is a cornerstone to effective risk management, since it provides a basis for developing sustainable risk management options over which a concerned community can feel a sense of 'ownership'. This principle is enunciated throughout Chapter 1 and in all contemporary guidance on EHRA. Effective community engagement can also facilitate transfer of risk assessment and risk management information, a process referred to as risk communication.

It is important that risk communication encompass a clear description of the assumptions and underlying uncertainties inherent in the risk assessment processes. These points are further emphasised in guidance on how to write comprehensive reports (see Chapter 7)

To be effective, community engagement must occur early in the process and be sustained through a structured strategy. The strategy should provide for meetings with concerned groups as well as the dissemination of information in printed form (e.g. newsletters) or online.

There are a number of sources of guidance on community engagement that have been developed in an Australian context.

- The Victorian Department of Sustainability and Environment has published a series of monographs on risk communication as part of a community engagement strategy along with consultation and community involvement. The monograph series is: *Effective engagement: Building relationships with community and other stakeholders, Books 1–3.* See <http://www.dse.vic.gov.au/engage>.
- The NSW Department for Urban Affairs and Planning has produced a manual on community consultation (Carlson & Gelber 2001).
- The Queensland Government has published two guidance monographs on community engagement and risk communication (Queensland Government 2001a, b).
- Two community consultation guides have been published by the WA Government: one by the Department of Environment and Conservation (2006) and another by the Department of Health (2006). While the first of these guidance documents is primarily aimed at managing consultations on contaminated sites, the second provides a framework for developers of new as well as existing proposals to use with communities. Both include useful general advice on community engagement.
- enHealth has produced a national guidance document that includes pragmatic advice on how to understand and respond to community concerns about environmental health issues (enHealth 2006).

6.2 RISK PERCEPTION AND RISK COMMUNICATION IN THE CONTEXT OF COMMUNITY ENGAGEMENT

The goals of risk communication are to (Reckelhoff-Dangel & Petersen 2007):

- help residents of affected communities understand the processes of risk assessment and risk management
- improve the quality of the risk assessment by enabling community members to contribute relevant information, such as their observations and local knowledge about the risks
- enable residents to form valid and informed perceptions of the likely hazards
- provide residents with knowledge so they can participate more effectively in making decisions about how to manage the risks.

It is crucial that consultation with stakeholders occurs early in the EHRA process, and that it be done with an understanding of the factors likely to shape community perception of the risks.

It is also crucial that the risk assessor be aware that ordinary people perceive risks in quite a different way to the scientific professional. Emotional factors often dominate the way risks are perceived, especially when the risks are thought to be beyond the control of the individual and to affect themselves, family members or close relationships. Risk perception is discussed in Section 6.2.1.

The potential mismatch in language and understanding can be illustrated with a hypothetical description of the risks associated with aflatoxin contamination of peanuts, as described by an 'expert', juxtaposed against the real question that a community member may be concerned about.

Expert:

'*A lifetime exposure to aflatoxin at a concentration of 20 ppb in food, assuming an average dietary pattern, yields an estimated excess carcinogenic risk to the exposed population (at the 95 per cent confidence level) of one case in a million.*'

Concerned parent:

'*Will my children be safe if they eat peanut butter sandwiches every day?*'

Providing an appropriate answer to the real question that concerns the community can be quite challenging.

Community engagement can be derailed, particularly when a situation has been allowed to develop where obfuscation and a lack of communication have led to an elevated level of community anger, conflict and mistrust. High levels of stress, concern or controversy are bound to make the already complex task of risk communication more difficult.

This chapter does not set out to be a complete guide to community engagement and risk communication. Rather, it attempts to summarise some key points from the expanding literature on community engagement, risk perception and effective risk communication.

There are a number of books and reviews that provide information relevant to risk perception and risk communication, which are important in the community engagement process. For example:

- Paul Slovic, one of the leading 'gurus' on risk perception, has written extensively on the topic. His monograph (Slovic 2000) addresses heuristics (emotional factors) and other cognitive theories of risk perception.
- Melissa Finucane (Finucane 2004) has written a chapter on risk communication in an Australian context in a textbook on Australian EHRA (Cromar et al. 2004). This chapter also includes information on the heuristic factors that shape or influence risk perception.
- The US EPA has issued a guidance document (workbook) on risk communication that includes pragmatic advice on community engagement (Reckelhoff-Dangel & Petersen 2007). This document complements an earlier US EPA guidance manual on risk communication (US EPA 2003b).
- Various authors have explored the significance of trust and credibility as factors influencing effective community engagement (Peters et al.1997).
- The UK Health & Safety Executive recently commissioned a review of UK risk communication practices and their effectiveness (HSE 2010).
- CRC-CARE has published guidance on risk communication that specifically addressed community engagement around contaminated sites (Heath et al. 2010).

6.2.1 Risk perception

All parties, both expert and non-expert, will have perceptions of risks. Experts and non-experts alike are influenced by emotion, beliefs and their views of the world (Thomas & Hrudey 1997).

Heuristics is the psychological term used to describe the process whereby people frame their perceptions of risk. Heuristics are essentially 'rules of thumb' by which we make judgements about everyday occurrence. These rules are simple, widely applicable and often reasonably accurate. However, they can lead people into making snap judgements about risks, when a more deliberate or logical analysis could lead to a different appraisal (Finucane 2004).

There are different types of heuristic approach:

- An 'availability heuristic' describes a probability estimate based on our ability to imagine an outcome. It is influenced by the frequency with which one can recall an event, or how vivid the memory is. Intense media coverage can increase this form of heuristic by making it more memorable.
- A 'similarity heuristic' is where something is seen as representative, an instance, a set or a group. It may involve stereotyping or some other automatic process of categorisation based on preconceived ideas.
- An 'anchoring heuristic' is where people start with an original bias and one 'anchors' their beliefs based on this pre-conceived idea.

6.2.2 The social context of risk perception

The social amplification of risk framework (SARF) is a concept that explores how media and other sources of information dissemination can provide a stimulus to community engagement, but it can also lead to exacerbation of community concerns by maintaining an ongoing and spreading of narrative relating to environmental incidents or issues (Pigeon et al. 2003). SARF can be both an integrative and a predictive model for analysing the social context of risk perception and risk communication. The influence of media reports on maintaining community concerns around a series of reports of cancer clusters in Australia has been explored in the SARF context (see Stebbing et al. 2008). There are also instances where the media have raised concerns about a particular or perceived risk, and then 'walked away'. In such instances, the media may engender community concern without awareness or acknowledgement of the consequences and may use the argument when called to task, that 'it is the community's right to know'.

Critics of SARF have suggested that the framework ignores the nature of perceived risk by implying that the risk must be shown to be 'real' before it is socially amplified or socially attenuated (Busby et al. 2009). Even if no causal factor linked to the environment can be demonstrated, as in the above Australian case, the mere perception that a cancer cluster has occurred is enough to create community anxiety and calls for action to be taken. Lay and media discussions of such events are often based on the proposition that there was a real, but undefined, hazard out there somewhere that was causing the risk event (Stebbing 2010). This confirms that, real risk or not, consequences do occur because people respond to their perception of risk not to the risk itself, no matter how it is characterised. In this way, the social response to the perceived hazard may become enlarged

or expanded beyond that expected by experts, institutions and media and it indicates that risk cannot be studied or discussed in isolation from the social context of engaged stakeholders.

6.2.3 Differences in 'real' and 'perceived' risk

Ideally 'actual', 'estimated' and 'perceived' risks should be closely aligned. This presents an immediate problem, as actual risks are often unquantifiable and unknowable. The aim of risk assessment should be to achieve the alignment of actual and estimated risk and the aim of good risk communication should be the alignment of perceived and actual risk. However, if the risk messages are not properly framed, or delivered in a context of mistrust or limited engagement, the outcome may involve an exacerbation of misunderstanding and conflict.

A simple numerical estimate of risk portrayed as the 'real risk' ignores the subjectivity and multiple dimensions of risks (Thomas & Hrudey 1997). People see risk as multidimensional and not represented by a numerical value, and will judge it according to its characteristics and context. For example, trauma or death as the result of an involuntary catastrophic reaction is likely to be dreaded more than the situation where the adverse consequences are the result of a situation where the risk is assumed voluntarily and the person feels some degree of control (e.g. motor vehicle accidents).

Concerns about risk will be heightened by risks that (DOH 1998):

- are involuntary or imposed on the community
- are man-made rather than natural
- are inescapable
- are controlled by parties outside the community
- are of little or no benefit to the community
- are unfairly distributed
- are related to an untrusted source
- are exotic or unfamiliar
- affect children or pregnant women
- are ones that affect identifiable rather than anonymous people
- cause insidious and irreversible damage
- cause dreaded health effects such as cancer
- are poorly understood by science
- are subject to contradictory statements from responsible sources (or, even worse, from the same source).

Concerns about risk will be lessened when:

- the risks are voluntarily assumed
- the risks have a natural origin
- individuals or the community feel able to exert some control over the risks
- there are clear benefits from the risks
- the risks and benefits are fairly distributed
- the risks are associated with a trusted source
- the risks are familiar
- the risks only affect adults
- the risks are understood
- the process of how the risks are determined is understood.

These concepts about how concerns about risks may be heightened or lessened are drawn from the work of Sandman (1993), and they are essentially the same as the factors he describes that affect 'outrage'. While knowledge of these factors may not necessarily be helpful in planning a risk communication strategy, an awareness of the factors that influence community perception of risk could help if the planned strategy begins to fail.

6.2.4 Environmental risk perception – an Australian context

There is relatively little published information on how Australians perceive environmental health risk and the factors that influence these perceptions.

The findings of one such survey revealed the following interesting findings (Starr, Langley & Taylor 2000):

- People are very concerned about chemical pollution.
- More than 80 per cent of respondents tried to avoid chemicals in their daily life, but chemicals represent less of a health hazard than lifestyle factors.
- Two-thirds of respondents thought they had some control over risks to their health.
- Higher perceptions of health risks were associated with personal involvement (e.g. smoking, illegal drugs and sunbaking).
- Many people feel that 'natural' chemicals are less risky than man-made chemicals.
- 76 per cent of respondents believe there is more environmental contamination than previously, but most thought their community is becoming a healthier place to live.
- Respondents sought information on environmental risks from a wide range of sources.
- 60 per cent of respondents had at least a moderate degree of confidence in news media reports.
- Doctors are seen to be credible sources of information on environmental health risks.

Some of these findings may appear paradoxical such as: the apparent mismatch between perceptions that environmental pollution is becoming worse; general concerns about man-made chemicals; and perceptions that communities are somehow 'healthier'. Also interesting is the way in which people source information from the media and the trust placed in such sources. This survey was undertaken before the widespread adoption of broadband internet access. It is likely that the internet would feature as a primary information source today. This will be increasingly challenging because there is no quality control on information on the internet.

6.3 RISK COMMUNICATION – THINGS TO KNOW AND THINGS TO AVOID

Risk communication may be erroneously seen as a one-way process aimed at rectifying incongruities between the community's perceptions and the opinions of regulators. However, it should be recognised that all parties will have perceptions about a situation and the ultimate aim is to draw these perceptions about risk, the estimated levels of risks and the actual levels of risks as closely together as possible.

Risk communication should not be seen as a retrospective form of community involvement and consultation. It is an interactive process involving the exchange among individuals, groups and institutions of information and expert opinions about the nature, severity and acceptability of risks and the decisions taken to combat them.

Good risk communication and consultation should result in an outcome where there is a high level of agreement between the affected parties. It also entails knowing how to respond to public concern, and is a genuine process conducted with the community's interest in mind. Good risk communication and community involvement will enable government and industry to better understand public perceptions and to more readily anticipate community responses. It will increase the effectiveness of risk management decisions and reduce unwarranted tension. It will explain risks more effectively and constructively inform communities.

While engagement with affected communities is always an important element of effective risk communication, there needs to be an awareness of the possibility of 'consultation fatigue'. This can occur when engagement has been intense, covering a broad range of issues and over a long timescale. There must also be an awareness of what people expect as an outcome of the negotiations, and whether this might include an expectation of payment, or other support, for participation.

Communication must be always be seen as 'two-way'. Listening to and respecting the views of other parties is just as important as clearly communicating a prepared risk message. The language of risk communication is important. It must be kept simple and questions must be answered directly. Scientific language must be translated into language that is understandable by an educated layperson.

Effective communication requires mutual trust between all the parties. This trust must be earned through a commitment to open and honest interaction. Any factual information presented should be correct. Exposure of an error of fact during a consultation, whether deliberate or inadvertent, will undermine credibility, destroy confidence and potentially breach trust. Relevant uncertainty must be disclosed and be quantitatively or qualitatively expressed.

There must be a commitment to maintaining communication, including providing specific and relevant information when requested or promised. It is important to establish realistic risk communication objectives and ensure that openness is maintained throughout the processes.

Key risk communication messages should be shared with other stakeholders so that there is no apparent discord between parties attempting to deliver a consistent message.

The knowledge base of people involved in the consultation should never be underestimated. It is good practice to at least conduct a 'Google' search to find out what they have been reading about an issue.

The ultimate peril of poor risk communication is to lose control, perhaps permanently, to another group that earns a more trusted position, and which may have opposing vested interests. If possible, it is better to initiate communication with people before a crisis develops. It is much more difficult to explain the complexities of a risk scenario when people are panicking. It may be more difficult to build the necessary trust if you are only beginning to reach out to people when a problem has reached crisis point.

Since the initial public reaction is usually one of concern (perhaps strong concern or alarm), the risk communication message must address the policy context within which it will be dealt and also address the demand for immediate action. Don't try to convince a worried or sceptical public that 'there is nothing to fear'. Good risk communication is not simply a public relations exercise. Delivering a 'feel-good' message about how well your organisation is handling an issue is likely to be counterproductive, or even offensive.

The aim of the communication process need not always be to reduce concern about risks. Many public health interventions are intended to increase public concerns about risks such as smoking or excessive alcohol consumption. In communities with regional lead contamination (e.g. Port Pirie or Broken Hill), public health activities have been designed to increase concerns about what are often subtle effects, and to provide information about specific activities that can be undertaken to protect children.

Some of the key principles of effective risk communication are (adapted from US EPA 1988; DOH 1998):

- accepting and involving the public as a partner and stakeholder
- carefully planning and evaluating the nature and content of the risk communication undertaken so that it is relevant and understandable
- listening to the public's specific concerns – trust, credibility, competence, fairness and empathy are often as important to the community as statistics and scientific details (trust and credibility are very difficult to regain if lost; experts do not command automatic trust)
- being honest, realistic and open
- appreciating that intentional communication is often only a minor part of the message actually conveyed (the manner of delivery and its tone may be more important than its content)
- ensuring that information is accurate, consistent between agencies and not speculative
- effectively communicating with the media
- acknowledging the public concerns and the effects on the community
- focusing on issues and processes rather than people and behaviours.

6.4 RISK COMMUNICATION – UNDERSTANDING CONFLICTS

Even with good community consultation and risk communication there may be disagreement between parties.

A consideration of potential conflicts will assist in providing a context for effective risk assessment, risk management, risk communication and community consultation. Examples of these conflicts are:

- economic activity (e.g. jobs, property values) versus conservation and health protection
- personal experiences and perceptions versus so-called 'objective' evidence
- quality of life and aesthetics versus defined disease problems
- local control and involvement versus external control structures
- local concerns versus national/statewide/regional concerns
- monitoring and health data versus personal experience
- personal experience versus scientific literature in making causal inferences
- broad community concerns versus narrow interest groups
- urgency versus priority determination
- political activism versus incremental scientific analysis
- voluntary exposure hazards versus involuntary exposure hazards.

Communication and consultation are important so that these conflicts are resolved.

6.5 PLANNING RISK COMMUNICATION

Information about risks needs to take into account their complexities and uncertainties, and be constructed so it can result in meaningful interpretation by all parties. People's responses to risk will be strongly influenced by their wider values, so isolated facts about risks may have limited impact on risk acceptability (DOH 1991), especially when risks are perceived to have little benefit.

Planning for an appropriate risk communication strategy is particularly important for managing a public health emergency or crisis situation, since the level of community concern is bound to be heightened (Glik 2007). This review cites a number of guidance manuals relating to risk communication in emergency situations.

Chess and Hance (1994) outlined a number of issues that need to be addressed when designing community consultation and risk communication programs:

- What is the purpose of the consultation? Is it to gain information, ideas and options? Is it to build credibility? Is it to meet regulatory requirements? Is it to provide maximum opportunity for public involvement?
- Who is the audience? Those who perceive themselves to be affected should be able to participate in the process. 'The community' is diverse, with different groups regarding risk in different ways. They may need a range of messages and styles of delivery.
- How will industry be involved? What responsibility will be taken by industry versus the regulator?
- What does the community want to know? Local community leaders, environmental groups and environmental health officers may often be able to provide broader information about particular concerns.
- Have appropriate persons been identified who can represent the views of potentially vulnerable groups within the community, such as children and the elderly?
- How will communication occur? Smaller, informal meetings are often more effective than large impersonal meetings. At large meetings, some members of the community may feel apprehensive about asking questions or expressing opinions. There is a need to avoid partisan chairpeople for meetings, and written materials may have more credibility than the spoken word. Materials need to be pre-tested before they are printed and distributed, and evaluated afterwards. There is a need to determine how industry and government will listen to concerns and how information about concerns will be sought. If the community is not listened to, it will cease to listen.
- Do not seek more feedback than can be used because this will lead to community disillusion and loss of trust.
- Seeking grudging approval from the community will be far less productive than genuinely seeking feedback that will be used, asking for comments in a situation where plans can be changed.

A simple, effective and underutilised approach for identifying communication challenges is to engage non-technical family members or office staff in a discussion of the issue to gauge how those who are not driven by the science and policy of the risk assessment will react to the messages.

It is important to avoid problems by anticipating issues such as:

- lack of communication skills (by any of the parties)
- limited resources and time and staffing (by any of the parties)
- confusion between the 'risk assessment' and 'risk management' phases
- cultural differences
- legal considerations
- external politics, hidden agendas and political pressures
- conflicting interests within the varying parties concerned
- impacts from the media
- evaluation of the consultation.

Evaluation is a continuous process designed to avoid mid-course corrections and repeating failures. Evaluation may cover: whether the communication was timely; whether the communication was sufficient; whether the public was empowered; and whether the credibility and trust of the organisation was enhanced (adapted from Chess & Hance 1994).

6.5.1 Illustrative example of community engagement

There are a number of examples of community engagement that illustrate the application of principles outlined in this chapter. Equally, there may be some that illustrate how things can go horribly wrong if scant attention is paid to the principles.

The following example (Box 2) shows that a good outcome can be achieved where community engagement is pursued using appropriate consultative techniques.

BOX 2:
Aluminium smelting and the community

Public consultation and commitment to an independent study resulted in a successful resolution of public health concerns in the Portland community.

In 1994 Portland Aluminium sought approval to increase sulfur dioxide emissions by nearly 30 per cent so that it could increase production at its aluminium smelter in Portland.

Members of the community were opposed to any increase in emissions, with the central issue being the effect of sulfur dioxide on health. There was a widespread belief that asthma levels were high in the Portland area. There were also similar concerns about the levels of sore and itchy eyes and skin irritations, as well as odours and acid smells.

Portland Aluminium stated that, with increased emissions, the use of taller stacks would improve air quality at ground level by allowing sulfur dioxide to disperse higher into the atmosphere.

Many residents had concerns about the reliability of air monitoring within the Portland area and believed they were not given complete information about the potential health effects associated with aluminium production.

In response to these concerns, the Victorian Department of Human Services established the Health Professionals Advisory Committee, which included local health professionals, a respiratory physician and department representatives. The role of the committee was to organise and oversee an independent health study to assess the potential for any adverse health effects from the proposed increase in sulfur dioxide emissions from the smelter.

A proactive program of community consultation was established and local residents were interviewed and given the opportunity to raise key areas of concern. The committee then ensured these concerns were addressed in the study's terms of reference.

The Victorian EPA then granted Portland Aluminium approval to replace the low stacks at the smelter with six tall stacks and to monitor their emissions for 12 weeks. The findings of the health study and the results of monitoring emissions from the old and new stacks were to be evaluated before the application to increase sulfur dioxide emissions was granted.

The health study involved a literature review and a health survey. To determine whether there was an increase in asthma and itchy eyes in Portland, the consultants surveyed residents of Portland and Warrnambool (a similar population) using a questionnaire that covered a range of health symptoms.

The study also reviewed the measurements of ground-level concentrations of sulfur dioxide that resulted from emissions from the older low stacks and the new tall stacks after they were built.

The literature review found there was no evidence that sulfur dioxide caused people to become asthmatic, but it did cause symptoms such as wheeze to occur more often. The survey showed that other health symptoms such as itchy eyes, cough, stuffy noise, sore throat and skin rash were more common in Portland, but there was no significant difference in the proportion of people with asthma and wheeze in Warrnambool, although both cities had high rates.

Monitoring data for 1995, 1996 and 1997 showed that the one-hour 'acceptable level' for sulfur dioxide at ground level was exceeded four times over this period. However, monitoring of emissions from the new tall stacks showed much lower levels.

The monitoring results were used to predict the ground-level concentrations of sulfur dioxide that would occur with the proposed 30 per cent increase in smelter emissions. The levels in Portland and surrounding areas were predicted to be well below the standard.

The results of the study were discussed with the community at a public meeting and a report of the study was circulated. The study concluded there was no evidence that the proposed increase in sulfur dioxide emissions from the taller stacks would be detrimental to health.

The report was well received by the community. Portland Aluminium was given EPA approval to increase sulfur dioxide emissions from the smelter and ongoing monitoring of air pollutants was included as a condition of the licence.

Chapter 7: Reviewing and appraising an environmental health risk assessment report

The credibility of site-specific health risk assessments in the successful management of environmental risks depends upon coherent and logically developed reports. Such reports may be prepared by consultants for industry clients, or to meet regulatory requirements. The purpose of this chapter is to provide some guidance on how to review or appraise a report on an EHRA, and what constitutes good practice. It does not purport to outline the specific reporting requirements that may be required by individual state, territory or Commonwealth jurisdictions. However, information in this chapter may be useful to state or territory regulators in determining whether reports submitted for their review contain relevant information that has been presented and evaluated in a cogent and logical way, with appropriate attention to underlying scientific principles.

The relevance and application of the key review aspects identified in this section will depend on the complexity of the EHRA. It is not intended that they are rigidly applied to all EHRAs. Some EHRAs will be quite simple and will not need to address all aspects listed below.

7.1 PROBLEM FORMULATION AND SCOPE

There is increased emphasis on problem formulation and scoping, including upfront consideration of potential management options and the utility of the EHRA in helping discriminate between these options. It is important that these issues are addressed in the EHRA report.

An EHRA can be initiated for a variety of reasons. Regulatory requirements, community concern over an issue, and responding to accidents and emergencies are a few of the contexts that may form the basis of the need for an EHRA. For a risk assessor to understand the problem and to define the scope of the assessment, early identification of stakeholders and engagement are critical. In reviewing or assessing an EHRA, the answers to the following questions should be clearly evident in relation to problem formulation and scope of the EHRA:

- Has the problem been clearly defined?
- How will an EHRA assist with resolving this problem?
- Have the interests and concerns of affected parties been reflected in formulating the problem?
- If existing environmental conditions appear to pose a threat to human health, have options that exist for altering these conditions been described?
- Has the scope of the risk assessment been defined? Is it clear what elements will or will not be considered? Such elements may include stressors, sources, exposure pathways, exposure routes, populations, and effects or exposure endpoints to be evaluated.
- Have the assessments that are required to characterise the risks of existing conditions and the effects on risk of the proposed options been described?
- Where multiple stressors or COPCs have been identified, is the approach taken consistent with guidance suggested in Chapter 12 , or in guidance contained in the *Framework for cumulative risk assessment* (US EPA 2003a) or the recommendations of *Science and decisions advancing risk assessment* (NRC 2008).

7.2 HAZARD IDENTIFICATION

The purpose of hazard identification is to identify what adverse human health effects are associated with the agent or agents of concern.

The hazard assessment component is likely to be based on a number of studies, conducted in different species within each toxicology study type, such as acute, chronic, developmental or reproductive toxicity. The report must identify the critical studies and the way these have been used in the EHRA in a transparent, accountable and defensible manner (see Section 7.2.2).

7.2.1 Study identification

The toxicity studies (or reviews or monographs) on which the hazard identification and assessment are based should be clearly identified in the risk assessment report. This information is important for identifying the basic data (or reviews or monographs) on which the risk assessment is based.

7.2.2 Checklist for hazard identification

The following checklist is adapted with slight modification from US EPA (1995a). A summarised version can be used if tolerable intake data from WHO or NHMRC are used.

1. Toxicological studies (see Chapters 9 and 11 for more detail)
   - What are the key toxicological studies that provide the basis for health concerns? How good is the key study?
   - Is all relevant information presented and reviewed?
   - Does the report highlight critical aspects of data quality?
   - Is the data from laboratory or field studies? Is the data for single species or multiple species?
   - If the hazard is carcinogenic, has comment been included on issues such as: observation of single or multiple tumour sites; occurrence of benign or malignant tumours; certain tumour types not linked to carcinogenicity use of the maximum tolerated dose; and whether a mode of action (MoA) can be identified and supported by peer-reviewed studies?

- Has there been a weight-of-evidence (WoE) approach in presenting a judgement as to the likelihood of human carcinogenic hazard, and does this include a clear articulation of the rationale for the position taken?
- If the hazard is other than carcinogenic, what endpoints were observed, and what is the basis for the critical effect?
- Have other studies that support this finding been described? Are there any valid studies that conflict with this finding?
- Has the report identified research that would reduce uncertainty in the EHRA and increase confidence in its outcomes?
- Besides the health effect observed in the key study, are there other health endpoints of concern?
- Are there any significant data gaps, and how have these been addressed?

2. Epidemiological studies (see Chapter 10)
   - What types of epidemiological studies were used (i.e. ecologic, case-control or cohort)?
   - What was the size of the study population?
   - How large are the confidence intervals on the observed measures of risk?
   - Were exposures adequately described or categorised in the epidemiological studies?
   - Was the degree to which other causal factors could be excluded well described in the epidemiological studies?
   - How relevant is the available epidemiological evidence to the issue addressed in the EHRA
3. Assessing the relationship between a possible cause and an outcome
   - How much is known about the biological mechanism by which the agent produces adverse effects?
   - Were relevant studies on mechanisms of action discussed, including the possible impacts of species differences in metabolism?
   - To what extent does this information help to interpret the toxicity data?
   - What are the implications for potential health effects?
   - How were any negative or equivocal findings in animals or humans addressed, and to what extent was this data considered in the hazard identification?
4. Summarise the hazard identification and discuss the significance of:
   - confidence in conclusions
   - alternative conclusions that are also supported by the data
   - significant data gaps
   - major assumptions.

7.3 DOSE–RESPONSE CHECKLIST

1. What data was used to develop the dose–response curve? Would the result have been significantly different if based on a different dataset?

If animal data was used:

- What species were used? The most sensitive, average of all species, or other?
- Were any studies excluded? Why?

If epidemiological data was used:

- Which studies were used? Only positive studies, all studies, or some other combination?
- Were any studies excluded? Why?
- Was a meta-analysis performed to combine the epidemiological studies? What approach was used? Were studies excluded? Why?

2. What model was used to develop the dose–response curve? What rationale supports this choice? Is chemical-specific information available to support this approach?

For non-carcinogenic hazards:

- How was the tolerable intake (or the acceptable range) estimated?
- What assumptions or uncertainty factors were used?
- What is the confidence in the estimates?

For carcinogenic hazards:

- What dose–response model was used? What is the basis for selecting the particular dose–response model used? Are there other models that could have been used with equal plausibility and scientific validity?
- What is the basis for selecting the model used in this instance?

3. To what extent were the route and level of exposure observed in the toxicology or epidemiology studies consistent with the expected human exposures in the situation under appraisal?

- Is the available data from the same route of exposure as the expected human exposures? If not, is pharmacokinetic data available to extrapolate across route of exposure?
- Are there any potential anomalies in the toxicity data? (e.g. Was bolus dosing with a carrier like corn oil used?)
- What is the degree of extrapolation from the observed data in the toxicological or epidemiological studies to the expected human exposures in the situation under appraisal? One to two orders of magnitude? Multiple orders of magnitude? What is the impact of such an extrapolation?

Some key elements of the dose–response component of the report that should be present are as follows:

- Valid datasets and plausible models for high- to low-dose and inter-species extrapolation are presented in dose–response modelling.
- The report offers an explicit rationale for any preferred dataset(s) and model(s) used in dose–response evaluation strengths and weaknesses of the preferred datasets are discussed, and scientific consensus or lack thereof is indicated for critical issues or assumptions.
- The report reveals how dose–response relationships change with alternate datasets, assumptions and models.

Some other specific ideal considerations are:

- Have all relevant toxicological facts been checked for accuracy and currency?
- Has the adequacy of the available toxicological database been appraised?
- Have the effects on each significant body system (e.g. renal, hepatic, cardiovascular) and the types of effects (e.g. allergy, genotoxicity and carcinogenicity, reproductive and developmental) been appraised and summarised for the relevant exposure routes?
- Has the critical toxic effects and organ/body system been identified?
- Have known toxicity modifying factors (e.g. synergistic and antagonistic effects resulting from exposure to multiple contaminants) been considered?
- Have toxicologically sensitive sub-populations been identified?
- Has the toxicological basis of the guidance value or potency factor, where applicable, been discussed and the uncertainties noted?
- Have NHMRC (where applicable) or WHO toxicological assessments been considered as the primary toxicological resource?
- Where relevant, have differences between (e.g. WHO and US EPA toxicological assessments been appraised and discussed?)
- Has the dose–response relationship for agents of potential concern been appraised and discussed?
- Has the data been presented in a form amenable to efficient interpretation and review?

7.4 SELECTION OF GUIDELINE VALUES

1. What source has been used for guideline values?
   - Is it a reputable source?
   - Are the guidelines relevant for the situation being assessed?
   - Have they been transcribed correctly?

7.5 DATA PRESENTATION

Data presentation is a critical part of any report outlining an EHRA. There are a number of ways that data can be summarised (e.g. in tables and diagrams) so that it makes it easier for a reader to grasp the essential information on which the outcomes of the EHRA have been derived and to understand basis for the main outcomes or conclusions from the study.

7.5.1 Presentation of toxicological data

It would not be unusual for there to be a range of studies from which different toxicological endpoints can be identified. These studies may indicate that the critical doses (NOAEL and/or LOAEL) for each of these effects can vary over a wide range. It may be useful to utilise a graphical or tabular representation of this variability.

The NRC has developed such a graphical method for representing the variability in toxicological endpoints. An example is shown in Figure 25, extracted from the NRC review of the draft IRIS assessment of tetrachloroethylene (NRC 2010 p. 94).

An alternative tabular method of presentation could be adapted from the approach used to summarise toxicological endpoints in JMPR pesticide evaluations (IPCS 2000 – see Appendix G for an example). In such a table, the critical endpoints are listed, along with the NOAELs and LOAELs, indicating the lowest values that were used to drive the ADI development.

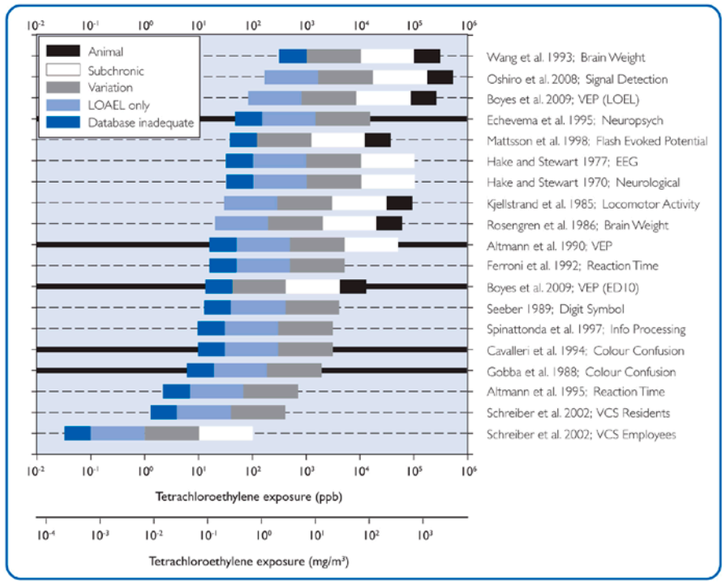
7.5.2 Presentation of epidemiological data

The burgeoning number of epidemiological studies and the need to integrate toxicological and other data creates a challenge for selecting data, analysing the data and summarising the results. The lucid presentation of the results in text, tables and figures is important for communicating these processes.

Although aimed at the assessment of interventions, the *Cochrane handbook for systematic reviews of interventions* (Higgins and Green, 2011 has a range of useful advice covering numerous topics, including selecting data, incorporating economic evidence and assessing risk of bias. There is a chapter about presenting results in text, tables and figures (e.g. using forest plots that display effect estimates and confidence intervals for both individual studies and meta-analyses).

The key challenge is to find a way to simplify the presentation of data from multiple epidemiological studies and to summarise the main findings. A common approach is to summarise the outcomes of individual studies in a consolidated table that includes some brief information on the study design, the range of variation in derived parameters (such as the adjusted odds ratios and their confidence limits).

Figure 25: Graphical depiction of the range of variability in toxicological endpoints that may be used in an EHRA



Each bar represents a single study, with the upper or right end indicating a possible point of departure for risk assessment, after conversion of dose to 'human equivalence' or adjustment for continuous exposure in animal studies. Solid horizontal lines indicate the studies considered most applicable to a risk assessment. Shadings indicate the application of various uncertainty factors (black for inter-species extrapolation; white for extrapolation from sub-chronic to chronic exposure; grey for intra-species variation to account for sensitive individuals; light blue if the study indicated a LOAEL, but not a NOAEL; dark blue for uncertainty in the study database). The left end of the bar represents the toxicological reference dose (in this case RfC), which could be derived from this database.

Reprinted with permission from NRC 2010 by the National Academy of Sciences, courtesy of the National Academies Press, Washington, DC.

The IARC commonly uses such a tabulated approach to summarise a large number of epidemiological studies of different types. The summary of epidemiological data on formaldehyde is a good example of this approach (IARC 2006).

An alternative approach is to use a diagram that illustrates the relationships between the odds ratios describing the outcomes of individual studies. Such an approach was used by Hrudey (2009) to illustrate the extent to which odds ratios for the effects of chlorinated disinfection by-products in water on birth defects clustered around the value of unity, designating no effect (Figure 26).

7.6 EXPOSURE ASSESSMENT

Exposure assessment is one of the more critical aspects of an EHRA report. It is important that the report describes all of the exposure pathways considered, and provides an explanation for any pathways not considered, or where they have been considered not relevant in the analysis. This section should also address quality aspects of the sampling plans, measures to preserve sample integrity, and analytical data quality controls.

7.6.1 Interpretation of sampling data

General issues to be considered in sampling and data interpretation are discussed in Section 8.5. This section summarises some key elements of collecting and interpreting data relating to exposure assessment.

An appraisal of data must show an understanding of:

- the context of the risk assessment
- the topography of the area affected
- the demography of the population
- environmental factors such as stratification of water bodies, movement of plumes in air or groundwater, soil structure (e.g. presence of clay or fill, and the depths of individual strata), meteorological factors, groundwater flows
- the relevant current or future human activities.

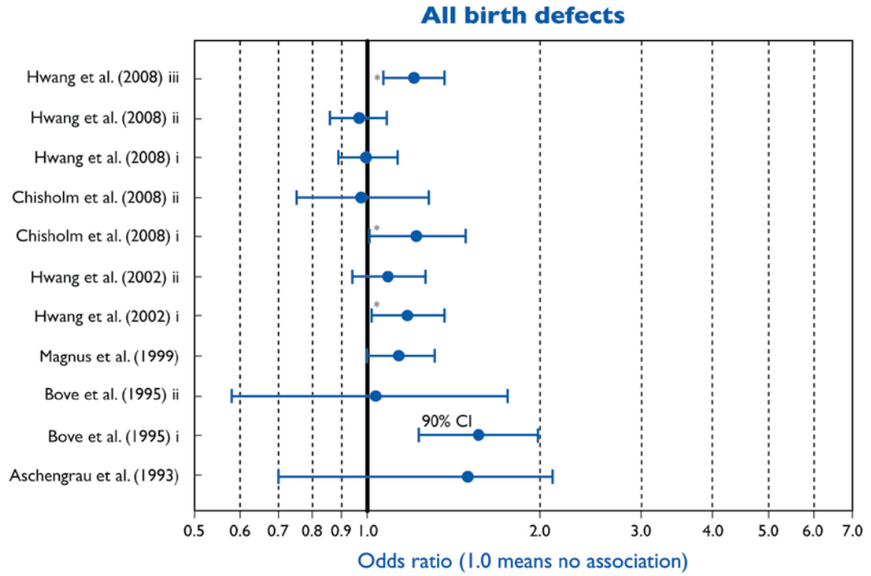
Too often numerical data is considered in isolation from other key parameters such as:

- levels of detection (and reporting)
- quality assurance for the data
- uncertainty about the data
- geographical relationship of one sample to another
- current or potential human activities.

Other key failings in numerical data analysis include:

- ignoring negative or unexceptional results by focusing on unusual or elevated results (the dataset needs to be considered in its entirety)

Figure 26: Graphical depiction of the range of variability in odds ratios describing the effects of water disinfection by-products on birth defects



Reference	Exposure comparison	Adjusted OR (95% CI)
Hwang et al. (2008) iii	Low TTHM (5 - 9 µg/L) vs. Lowest (0 - 4 µg/L)	1.21 (1.07 - 1.36)*
Hwang et al. (2008) ii	Medium TTHM (10 - 19 µg/L) vs. Lowest (0 - 4 µg/L)	0.97 (0.86 - 1.08)
Hwang et al. (2008) i	High TTHM (≥20 µg/L) vs. Lowest (0 - 4 µg/L)	1.00 (0.89 - 1.13)
Chisholm et al. (2008) ii	Medium TTHM (avg 109 µg/L) vs. Low (avg 54 µg/L)	0.98 (0.75 - 1.48)
Chisholm et al. (2008) i	High TTHM (avg 137 µg/L) vs. Low (avg 54 µg/L)	1.22 (1.01 - 1.48)*
Hwang et al. (2002) ii	Chlorination - high colour vs. no chlorination - low colour	1.09 (0.94 - 1.26)
Hwang et al. (2002) i	No chlorination - high colour vs. no chlorination - low colour	1.18 (1.02 - 1.36)*
Magnus et al. (1999)	Chlorinated water, high colour vs. no chlorination	1.14 (0.99 - 1.31)*
Bove et al. (1995) ii	TTHM >80 µg/L vs. <20 µg/L)	1.04 (0.58 - 1.76)[a]
Bove et al. (1995) i	TTHM >80 µg/L vs. <20 µg/L)	1.57 (1.23 - 1.99)[a]
Aschengrau et al. (1993)	Major malformations: chlorination vs. chloramination, surface water only	1.05 (0.7 - 2.1)

[a] 90% confidence interval
* significant with respect to 95% confidence interval

Reprinted from Hrudey, 2009 © 2009, with permission from Elsevier

- inadequately managing censored data (e.g. by assigning a zero value to results below the level of detection or reporting)
- accepting relatively high levels of detection or reporting so that the value of much data is obscured. This may have the consequence of failing to reveal gradients that will help to highlight the presence and location of environmental 'hot spots'. Examples have been seen where environmental health criteria levels have been treated as the level of reporting.

The very existence of levels of detection and reporting results in the need to censor data. Censoring of data can be particularly important when the maximum permitted criterion is close to the level of detection (e.g. with potable drinking-water standards). Data censoring must be addressed in an appropriate way (see Sections 4.6.4 and 8.7).

Given two similar results, the result that can be explained (e.g. by history or similarities with results from similar strata) will tend to be of less concern than the result that cannot be explained (Langley 1993b).

7.6.2 Exposure assessment checklist

The general components in an acceptable exposure component to a risk assessment report as follows (AIHC 1989):

- The purpose and scope of the exposure assessment and the underlying methodologies are clearly described.
- The specific populations and sub-populations that are the subjects of the assessment are clearly identified, and the reasons for their selections and any exclusions are given.
- The available data is considered and critically evaluated, and the degree of confidence in the data expressed (reasons for any data exclusion are presented).
- If models are used, their bases are described, along with their validation status.
- Potential sources, pathways and routes of human exposure are identified and quantified; the reasons why any are not included in the assessment are presented.
- Central estimates and, if possible, upper and lower bounds on exposures for the full population, and the distribution of exposures are described any preferred estimates are noted, together with supporting documentation.
- Uncertainties in the estimates are described, and the relative importance of key assumptions and data is highlighted.
- Research or data necessary to improve the exposure assessment is described.

Specific considerations are:

- Has the potentially exposed population been identified?
- Have potentially exposed, unusually susceptible sub-populations been identified?
- Have the estimates of chemical exposure for each significant exposure route and for each chemical of potential concern been adequately quantified and tabulated?
- In cases of presumed insignificant exposure, has the exposure been demonstrated to be small?
- Has the relative significance of each exposure pathway, based on the risk analysis, been discussed?

7.7 RISK CHARACTERISATION

The general components in an acceptable risk characterisation component to a risk assessment report are as follows (AIHC 1989):

- The major components of risk (hazard identification, dose–response and exposure assessment) are presented in summary statements, along with quantitative estimates of risk, to give a combined and integrated view of the evidence.
- The report clearly identifies key assumptions, their rationale and the extent of scientific consensus; the uncertainties thus accepted; and the effect of reasonable alternative assumptions on conclusions and estimates.
- The report outlines specific ongoing or potential research projects that would probably clarify significantly the extent of uncertainty in the risk of estimation.
- The report provides a sense of perspective about the risk through the use of appropriate analogy.

7.8 CONFIRMATION OF UTILITY

- Does the report address the issues described in the problem formulation and scope?
- Does the report provide guidance as to the relative effectiveness of risk management options?
- Has the report been peer reviewed or is it being audited?

7.9 GENERAL ASSESSMENT AND REPORT PRESENTATION

A checklist of points that should be considered in the overall presentation of the report and an assessment of its quality is provided below.

Report structure and data presentation

- Have all tables and figures been referred to correctly in the text of the report?
- Is the presentation of results consistent with the units for which the standards are written (e.g. using milligrams/m^3 instead of nanograms/m^3 when sampling ambient air)?
- Have SI units been used correctly throughout the report (Thompson & Taylor 2008)?
- Is there a map of testing sites that enables ready identification of sampling sites in relation to relevant environmental sources?
- Does the presentation format enable easy cross-referencing of results to maps and between different parts of a report?
- Are results presented in linear geographic sequence (e.g. either going downstream or upstream, or towards or away from a point source)?
- Has the report been presented in a suitable format to enable 'track changes', comments and cutting and pasting into a Word document? PDF formats hinder rapid transcription and can create the potential for transcription errors.

Issues of data quality

- Has the nature of the analyte been specified? This may be important when the valency (e.g. chromium) or chemical form (e.g. organic arsenic versus inorganic arsenic in fish; haem iron versus non-haem iron in animal samples) may be relevant for health risk assessment.
- Has the level of reporting (LOR) been specified for each analyte and for each batch of results?
- Have the laboratories involved in the assays been identified, along with their QA/QC procedures?
- When there have been several sampling periods, have the results been collated into a single table to enable efficient appraisal of results and trends to be readily detected?
- Is it possible to determine whether an absence of results indicates an absence of testing or that results were non-detects?Where composite samples have been used, have these been identified? Is there an explanation of compositing techniques? (Note: compositing is specified practice for some types of food sampling).
- If there is likely to be significant heterogeneity in the individual components and *n* is the number of samples, are there any reasons why guideline values should not be divided by *n* in the assessment process to help identify high concentrations in one or two individual components of a composite?
- Where censored data has been used, has the censoring been fully explained?

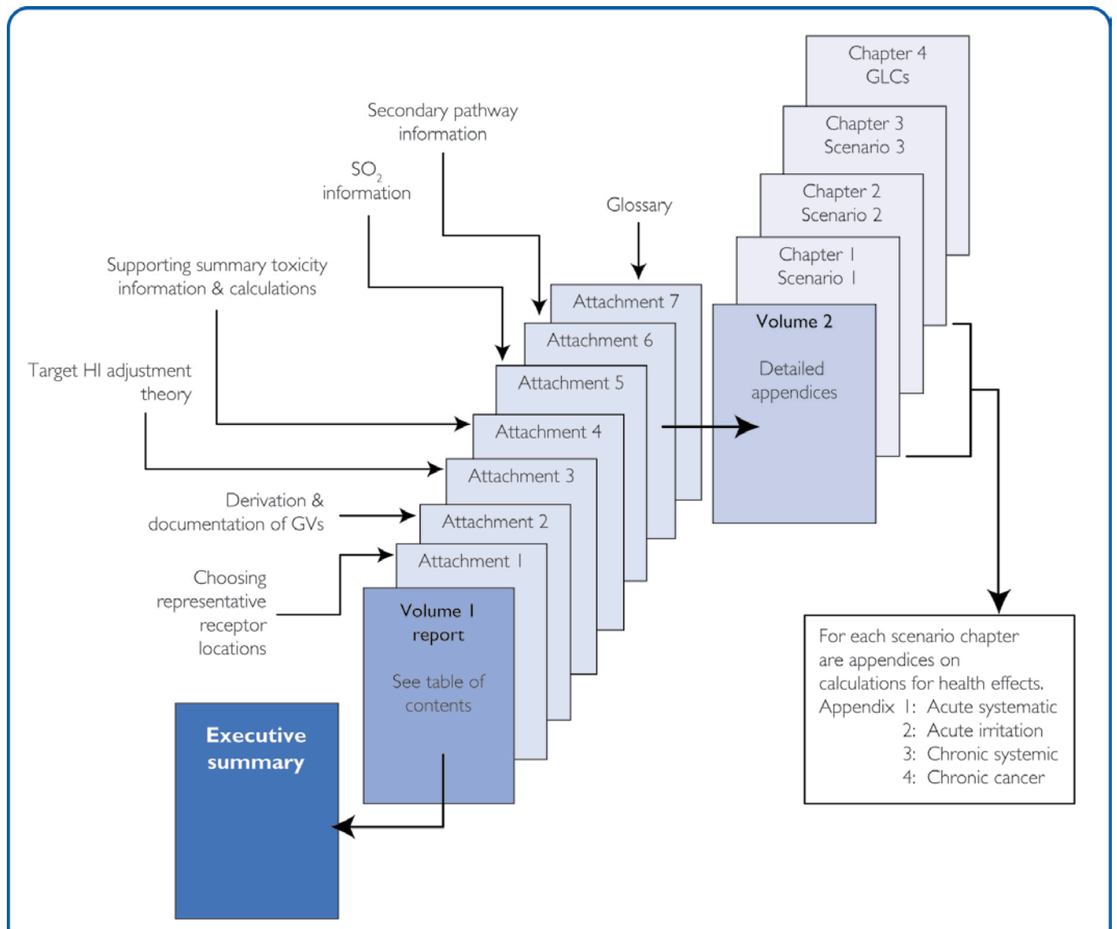
Clarity of how the data has been used, assumptions made and relevant justification

- Has information from previous reports on the situation been appropriately selected and incorporated into this report?
- Has irrelevant information from other situations been excluded from the report?
- Have all assumptions and default data been identified and justified?
- Has the analysis been based on an up-to-date literature appraisal?
- Has information been presented coherently and in an appropriate sequence, to enable efficient appraisal of the report?
- Has the rationale for the sampling program and selection of analytes (including sampling objectives) been addressed?
- Have environmental factors relevant to the choice of analytes been noted, along with a rationale if unusual analytes are included or common analytes are excluded?
- Have all conclusions been justified?
- Does the report include or enable ecological risk assessment, if required by regulatory authorities?
- If toxicological data and the exposure scenario led to the conclusion that a high concentration of agent is permissible, does the result violate ecological, aesthetic, land use or physical principles?
- Has a risk management decision been made during the course of the risk assessment and, if so, how might it have influenced the calculation of risk?
- What has been the involvement of the public?
- How has information been communicated to the public?
- What processes of community consultation have taken place?

In relation to the last two points about community engagement and consultation, it should be noted that a specific requirement for such consultation may be incorporated into EHRA guidance and/or regulations in some Australian jurisdictions. Further advice on community engagement is presented in Chapter 6.

Presentation of the report should include attention to having an orderly structure and clear delineation of its component parts. Figure 27 is presented as one example of a report structure that could be used for an EHRA report. It has been developed primarily for assessing emissions from an industrial facility, but its structure is designed to achieve the desired objectives of clarity, structure and comprehension. Various state and territory jurisdictions may have different legislative requirements for an EHRA report structure and content, and the example in Figure 27 is not meant to be prescriptive.

Figure 27: Example of a structured EHRA report for air emissions from an industrial facility



Part 2: Additional guidance on selected issues

Chapter 8: Data collection and analysis

Collecting and analysing environmental sampling data is critical to the risk assessment process. Quality assurance (QA) of the collected data is also vital, and QA processes should be well documented. Data may be available from a preliminary or detailed site investigation (either or both) and is usually available prior to commencing the risk assessment. In such cases, the assessor must determine whether the data quality objectives of any previous investigations are compatible with the objectives of the risk assessment and whether the original data quality objectives have been satisfactorily met (NEPC 2010).

The sampling plan may have been informed by a conceptual site model (see Section 4.4), which will assist with understanding the source and medium (air, water, soil) from which the samples have been obtained and information on the site history, which could assist in assessing the types and mobility of the contaminants and their likely location. More specific guidance on preparation of sampling plans and sample density requirements is available (e.g. contaminated sites NEPM schedule B(2) NEPC 2010). If the sampling data is incomplete, or the sampling plan inadequate, there should be specific comment in the risk assessment report.

The purpose of this chapter is to provide some guidance on sampling the environment to collect data suitable for exposure assessment. It addresses some issues of analytical sensitivity and quality assurance, but it does not purport to be a comprehensive review of such topics.

8.1 ENVIRONMENTAL DISTRIBUTION

It is unlikely that contaminants will be uniformly distributed in the environment under consideration. An understanding of how chemical agents move between environmental compartments and the effects of environmental partitioning is necessary for developing sampling plans for chemical agents and the process of exposure assessment. This is especially important for volatile and gaseous substances, which can move from sources such as contaminated soil or groundwater into dwellings and open spaces where human 'receptors' may be exposed. Transport pathways for gases and vapours derived from soil or groundwater contamination may be dependent on soil characteristics such as soil type, porosity, depth of the contamination and the presence of organic carbon. These characteristics should be factored into the modelling that may be used to quantify the pathways. Best practice modelling approaches would, by definition, require model parameterisation using site-specific parameters in addition to soil vapour and indoor air data collection for validation of model predictions. Transfer characteristics are also important for dusts and particulates, which can be transported over significant distances from source.

Partitioning will reflect the fact that substances will move to the environmental compartment for which they have the most affinity (Calamari 1993; 1999). Transformation may occur in any environmental compartment.

8.1.1 Meteorological data

Meteorological data will be particularly important when evaluating both point source and generalised air pollution and potential exposures of populations. When environmental monitoring is being undertaken, there is a need to have concurrent meteorological data and this may be available at minimal or no cost from the Australian Bureau of Meteorology depending on the locality.

8.2 ENVIRONMENTAL SAMPLING AND ANALYSIS

Data collection involves acquiring and analysing information about hazards on a site that may affect human health and which will be the focus for the particular risk assessment (US EPA 1989).

Sampling is often carried out to more clearly define detected or suspected contamination and, if remediation occurs, to verify that contaminated material has been removed and that any contamination remaining does not constitute a health or environmental risk.

The greatest concern in collecting samples is to ensure the samples taken adequately represent potential exposures for the situation. Consequently it is essential to be fully apprised of the context of the risk assessment, the objectives of the task, the environmental conditions at the site locations and what analytes will be tested in each sample, before sampling commences (Lock 1996).

Inappropriate sample collection procedures yield samples that are not representative of the population of interest are of little use, seriously compromise the purpose of sampling, and contribute to the uncertainty of the analytical results (Keith 1990).

Laboratory errors can and do occur and if an aberrant or an unexpected result is provided, and the potential for laboratory error should always be considered. If a laboratory error is suspected, it may be worthwhile requesting relevant QA/QC information.

An important aspect of environmental sampling and analysis is that the environment is not static and sampling results can vary over time. The methodology used for collecting samples and interpreting environmental sampling data should take this into account.

8.2.1
Strategies for sampling the environment

In sampling, the statistical considerations need to be matched to expertise in situation assessment and a knowledge of the particular situation (Lord 1987). The sampling plan and decisions regarding the number, type and location of samples need to be developed with an understanding of the potential exposure pathways and routes (US EPA 1989).

Sampling will be influenced by, and will influence, the potential risk management outcomes. The proposed human activities for the particular setting will critically affect the nature of the sampling program. Some useful guidance on sampling is reported in Heyworth (1991).

The reasons for sampling include:

- determining the nature of contamination
- determining the concentration and distribution of the agent
- monitoring site conditions to determine if remedial actions are required
- designing and implementing remedial actions
- determining if remedial actions have been effective.

There are often three phases of sampling:

- an initial assessment to determine if detailed investigation is necessary
- a detailed sampling and analysis plan
- post-remedial validation.

For any of these phases, a sampling program with multiple stages may be required, especially for large and complex situations.

8.2.2
Sampling methodologies

Numerous techniques are available for environmental sampling and the field is progressing rapidly (Keith 1988; 1990; Perkins 1997).

Further information on approaches to environmental sampling and analysis are available in schedules B(2) and B(3) of the contaminated sites NEPM, available online at <http://www.ephc.gov.au/taxonomy/term/44>, from CSIRO websites such as <http://www.csiro.au/science/Environmental-Monitoring.html> and the US EPA at <http://www.epa.gov/esd/tsc/fact-sheets.htm>.

Sampling is often most effectively done as a staged and iterative procedure, where earlier results can be used to focus later sampling stages.

Some key issues are (Keith 1990):

- When sampling water, allowance should be made for the fact that stratification can occur in bodies of water.
- Groundwater contamination is affected by 'depth to water, recharge rate, soil composition, topography (slope), as well as other parameters such as the volatility and persistence' of the substance (Keith 1990 p. 614). There is always a significant risk of cross-contamination of aquifers when sinking bores and special precautions should be made to protect against this.
- Water sample contamination is always a problem, and this is most pronounced when very low concentrations are being sought.
- Considerable variation in an environmental medium over time may occur and environmental sampling may need to be spread over a period of time to give an accurate representation of potential human exposures.

Volatile agents require specialised sampling techniques to ensure the contaminants are not lost during and after sampling so that analytical results accurately represent the concentrations present. The inhalation route will be more important than for non-volatile contaminants. Factors that will have a significant effect are: soil disturbance; the physico-chemical properties of the soil and contaminants; and whether there is a renewable source or whether the contamination will dissipate over time. This field is developing rapidly, and readers are referred to Australian and US Guidance (NEPC 2010; US EPA 2002b).

8.2.3
Sampling patterns

Sampling plans will depend on the medium being sampled. If there is sufficient information about a situation, random sampling may be inappropriate or inefficient and judgemental sampling may be more appropriate. Air and water over a small area are likely to be more homogeneous than soil.

As a starting point for sampling plan design general information is available from Gilbert (1987), Heyworth (1991) and Keith (1988; 1990).

8.2.4
Sampling density

'Statistical equations are tools to be used as aids to common sense and not as a substitute for it' (Keith 1990 p. 612). Statistical formulae for determining sampling density are usually based on the requirements that the results will be normally distributed (i.e. in a bell-shaped curve) and that a particular concentration is equally likely to occur at any point. Some analytical techniques require an estimate of the mean of the results and the standard deviation of the results before sampling density can be calculated. These requirements can rarely be met during the stages of initial and detailed investigations as sites are often heterogeneous with a highly skewed distribution of results.

Sampling is a screening process, and false positive and false negative results will occur. From a health perspective the aims of sampling are to reduce the likelihood of a false negative that could result ultimately in significant adverse health effects, and to enable contaminated sites to be sufficiently identified and adequately remedied to protect human health.

A considerable amount of expert judgement is required to determine the amount of sampling. The final amount will depend on an integrated appraisal of factors, including:

- proposed or current human activities
- the number of stages of sampling considered feasible
- the scale and distribution of potential human exposures
- potential remediation and management strategies.

The sampling density requirements will vary from medium to medium.

8.3 ANALYTICAL METHODOLOGIES

Manahan (1993), Perkins (1997) and Kim and Platt (2008) provide general overviews of analytical methodologies used in environmental sampling analysis.

Good (1993 p. 45) considers an appropriate test method must be:

- accurate: it must be shown to give results which differ little from the concentration we would accept as the 'true' value. This is generally demonstrated by comparison with other well-respected techniques;
- precise: it gives results that show acceptably small variation from batch to batch and analyst to analyst when applied as prescribed; and
- robust: results are not unduly affected by minor variations in test conditions.

If these three criteria have been measured, the method can be relied upon to provide an answer within a predictably narrow range around the accepted 'true' value for a given sample. However, for a method to be widely useful it must be available and also be:

- not too complex: A procedure so complex as to be only useable by a few highly trained persons will probably be of limited practical value;
- not expensive: The costs of site assessments are already high; and
- reasonably comprehensive: Methods should determine an appropriately wide range of compounds potentially present. This may require several methods, rather than just one method, given the range of chemistry in compounds of interest.
- available: Even the best method is of little use if its use is restricted by copyright or other instrument, or it resides in an obscure journal unknown to potential users.'

The original analytical records (e.g. traces, chromatographs) should be retained and should be reviewed when analysis of the data is about to drive a significant action.

8.3.1
Choice of analytes

The choice of analyte will be principally governed by the 'issues identification' stage for the particular situation.

In the case of metals, because of significant differences in toxicity associated with some valence and oxidation states and other chemical forms (e.g. salts), it may be necessary to further speciate the analytes. Arsenic and chromium are good examples of metals that could require further speciation. Guidance on where speciation would be useful in an EHRA is outlined further in EHC234 (WHO 2006b). Routine metal speciation analysis is commercially available in Australia for species and compounds of arsenic (As), selenium (Se), mercury (Hg), tin (Sn) and lead (Pb), while metal speciation of some other elements is routinely conducted internationally. At sites where potentially toxic metals are present, consideration should be given to whether speciated metal analysis of media is possible (NEPC 2010). The availability of guidelines for speciated metals and metalloids is limited so this factor should be considered when deciding on the usefulness of speciated analysis.

8.3.2
Field instruments

In most situations, field instruments should be regarded only as a screening tool and their results require laboratory validation. There are some types of chemicals (e.g. total residual chlorine) that may be lost rapidly from a sample and the use of field instruments may represent the only practical way of measuring them.

Field instruments require regular maintenance and calibration, and skilled and diligent use.

The accurate use of such instruments relies on factors, including:

- the method of sampling
- the nature of the contaminant
- the presence of interfering gases or vapours resulting in overestimates or underestimates of environmental concentrations
- the type and make of the instrument
- the type of calibrant used
- the length of time since the last calibration
- the cleanliness of the instrument
- the skill and knowledge of the operator.

Field instruments may be useful for assisting to identify areas where sampling should be concentrated. They do not replace analysis in a laboratory.

Examples of field instruments are photo ionisation detectors (PIDs), X-ray fluorescence (XRF), dissolved oxygen, pH/redox and temperature probes. Information on time, date and method of calibration should be provided with reports.

8.3.3
Detection limits

Ideally, the detection limit of the analytical method used should be lower than the level at which the contaminant might become a health concern. A health-based GV (e.g. HIL or Tier 1 screening level) usually provides a benchmark against which the detection limit of the analytical methodology can be measured. If the methodology cannot achieve such a detection limit, there will inevitably be some difficulty, bias or imprecision in assessing risk, and efforts should be made to find an alternative, more sensitive analytical method.

8.4 QUALITY ASSURANCE OF DATA USED IN SITE-SPECIFIC HEALTH RISK ASSESSMENT

8.4.1
Data quality objectives

Data quality objectives 'provide critical definitions of the confidence that must be inherent in the conclusions drawn from the data produced by the whole project' and determine the degree of uncertainty or error that can be tolerated in the data (Keith 1990 p. 611).

Data quality objectives that clearly specify the amount, nature and quality of the data to be collected should be detailed. Data quality objectives will be situation-specific. More detail is available from the website of the US EPA Quality System for Environmental Data and Technology at <www.epa.gov/quality>, in the US EPA guidance manual (US EPA 1996b) and in the IPCS monograph on data quality (IPCS 2008). The criteria for both accepting and rejecting data should be rigorous and well documented.

Consideration will need to made as to whether routinely collected historical data will be as appropriate to use as data collected *de novo* for the risk assessment.

8.4.2
Sample handling, storage and transport

Sample handling and transport should be done according to relevant regulatory documents or Australian Standards.

Some key issues are (Keith 1990):

- Contamination may arise from substances in the sampling devices and storage containers. PVC and plastics other than teflon tend to absorb organics and leach plasticisers and other chemicals used in their manufacture. Some pesticides, halogenated compounds and metals strongly adsorb to glass.
- The loss of volatile analytes or reduced concentrations from irreversible absorption on the walls of sampling containers can be a significant problem.
- Sample preservation can be of considerable importance. If incorrectly stored, materials can have accelerated breakdown, chemicals may be lost by volatilisation, and proliferation or diminution of microbiological organisms can occur. The nature of the storage container, its seal, and the degree of refrigeration needed should always be considered and addressed.

8.4.3
Chain of custody

The consultant's report must provide the following chain of custody information (EPA NSW 1997):

1. The sampler
2. Nature of the sample
3. Collection date
4. Analyses to be performed
5. Sample preservation method
6. Departure time from site
7. Dispatch courier(s).

AS 4482.1-1997 Appendix H provides a chain of custody form.

8.4.4 Quality assurance

The information in Sections 8.4.4 to 8.4.11 is adapted from Good (1993).

All of the actions, procedures, checks and decisions undertaken to ensure the representativeness and integrity of samples and accuracy and reliability of analysis results must be recorded.

In the field, this includes selecting appropriate sampling methods, documentation and sample storage, cleaning tools before sampling and between samples, cleaning containers, and maintaining sample environment to minimise sample contamination and analyte losses.

In the laboratory, quality assurance (QA) involves proper sample control, data transfer, instrument calibration, selecting properly trained staff and suitable equipment, reagents and analytical methods.

8.4.5
Quality control

Those parts of QA that serve to monitor and measure the effectiveness of other QA procedures by comparison with previously decided objectives. In the field, this may include checking sampling equipment cleanliness by keeping rinses for analysis, cross-checking sample identities, duplicating sampling sites and performance of 'field blanks' and 'field spikes'. In the laboratory, quality control (QC) procedures involve measuring the quality of reagents, the cleanliness of apparatus, accuracy and precision of methods and instrumentation by regular analysis of 'blanks', sample replicates, 'spiked recoveries' and standard reference materials (SRMs), with proper recording of results for these checks and immediate investigation of observed problems.

According to these definitions:

- Adequate QA is achieved when the results of QC demonstrate that agreed objectives such as freedom from contamination, method accuracy and precision can be reliably achieved. An important decision then is the correct level of QC.
- As a general rule, the level of required QC is that which adequately measures the effects of all possible influences upon sample integrity, accuracy and precision, and is capable of predicting their variation with a high degree of confidence. QC is more often performed inadequately than very well.

8.4.6
Blanks

A reagent blank (or preferably two for very low-level analysis) is prepared by processing reagents only in exactly the manner used for each sample. The aim of the blank determination is to establish the magnitude of that component of the analytical signal that can be ascribed to contributions from reagents, glassware, etc. The contribution established should be subtracted from the gross analytical signal for each analysis before calculating sample analyte concentration.

8.4.7
Replicate analysis

Repeat analysis of at least one sample or 10 per cent of the batch of samples. The variation between replicate analyses should be recorded for each batch to provide an estimate of the precision of the method.

8.4.7
Recovery check

This means checking the recovery of reference material (matrix spike). One or more replicate portions of samples from the batch should be analysed after fortifying the additional portion(s) with known quantities of the analyte(s) of interest.

Recovery check portions should be fortified at concentrations that are easily quantified but within the range of concentrations expected for real samples.

The method used to correct the reported data for recovering the analyte(s) from the media must be explicitly stated. Failure to do so may render the reported data meaningless, and significantly compromise the exposure assessment and the derived risk assessment.

8.4.8
Reference material analysis

This is analysis of a sample similar in matrix type to the samples, with accurately known concentration of the analyte(s) of interest. Results of recovery checks and reference material analyses for each batch should be recorded so that the bias of a method may be estimated, and the day-to-day method efficiency monitored.

8.4.9
Surrogate spikes and internal standards

Wherever appropriate, especially for chromatographic analysis of organics, using surrogate spikes and internal standards is highly recommended. Including this in methods requires little additional effort and greatly enhances confidence in qualitative and quantitative results obtained.

Surrogate spikes should be added to each sample, blank and recovery/reference sample. They should be similar to the analytes of interest in terms of:

- extractability
- recovery through clean-up procedures
- response to chromatographic or other measurement

but which:

- are not expected to be found in real samples
- will not interfere with quantifying any analyte of interest
- may be separately and independently quantified by virtue of (e.g. chromatographic separation or production of different mass ions in a GC/MS system).

Surrogate compounds may be alkylated or halogenated analogues or structural isomers of analytes of interest, or where GC or LC/MS is used, deuterated standards are an excellent choice.

The purpose of surrogate spikes, which are added immediately before the sample extraction step, is to provide a check for every analysis that no gross processing errors have occurred that could have led to significant analyte losses or faulty calculation.

Another term (used in schedule B(3) of the contaminated sites NEPM) is that of a 'matrix spike'. The purpose of the matrix spike is to monitor the performance of the analytical methods used, and to determine whether matrix interferences exist. Matrix spikes should be performed when validating a method by adding it to the analysis portion before extraction or digestion.

Immediately prior to instrumental analysis, each sample, blank and recovery or reference material extract is fortified with a set amount of one or more compounds to be used as internal standards. These compounds should:

- not be found in real samples
- not interfere with quantifying the analytes of interest
- be separately and independently quantified.

The purpose of internal standards in chromatograms is to provide extra peaks that serve to check the consistency of the analytical step (e.g. injection volumes, instrument sensitivity and retention times for chromatographic systems). Analyte concentrations may be determined by measuring the ratio of the analyte response to that of an internal standard, with marked improvements in quantitative precision.

8.4.10
Control charts

Nadkarni (1991) claims that the heart of a QA/QC program is a control chart. Good (1993) agrees, explaining it is a 'numerical picture (a plot) of the variation of measured QC parameter (e.g. blank and recovery values). Data is plotted in the sequence in which it was obtained, and reviewed frequently in order to detect any problem with minimal delay. The use of these charts is highly recommended'.

8.4.11
Safety plans

The safety of people assessing a situation and nearby residents must always be considered in environmental sampling. Site safety plans should be developed where there may be such risks.

Guidance on protecting the community is included in schedule B(8) of the contaminated sites NEPM, accessible at <www.nepc.gov.au>.

8.5
DATA ANALYSIS AND PRESENTATION

The information in Sections 8.5.1 and 8.5.2 is adapted from Langley 1993a.

8.5.1
Assessing summary statistic data and presenting data

Vast amounts of data can be generated about a single environmental health investigation. Enabling an efficient and accurate appraisal of a situation requires that the data be collated in a form that allows an understanding of the location, extent, trends, and likely 'behaviour' of any environmental hazards. Data mapping is essential.

An adequate understanding of what is (and will be) occurring is almost impossible to achieve from pages of raw data, especially where there are abnormal results or more than a handful of results. At its worst, sample identification numbers, sampling points, technical logs and results for each analyte will be on separate pages.

There is a constant tension between consultants who wish to maintain individuality to their reports and government agencies that seek uniform reports. A uniform approach to the location and presentation of data makes for more rapid and accurate assessments of reports.

The major problems that can occur with datasets and assessments are:

- a failure to collate data and condense it into comprehensible tables
- providing cluttered datasets, tables and graphs
- treating the sum of the data as somewhat greater than the sum of its parts (this is exemplified by elaborate contour maps based on a very limited number of data points)
- providing fairly definitive conclusions insufficiently underpinned by supporting data
- considering the numbers in isolation from other data important to interpretation (e.g. situation history and characteristics of the sampled medium)
- inappropriate 'compositing' of data
- a failure to recognise that model outputs are not 'data' in the sense described in this chapter.

8.5.2
Summary statistics

No single summary statistic (e.g. an arithmetic mean or the median) fully characterises a situation. Instead, a range of summary statistics is needed to build up a picture of potential agents and exposures.

Each summary statistic will have a contribution but will also have certain limitations. Given the complex nature of most datasets, a range of summary statistics needs to be presented, as the mix of summary statistics will be more useful than a single summary statistic.

As much sampling is judgemental rather than random, caution needs to be taken with using conventional statistical methods that usually assume the random collection of data and the use of normally distributed data. Much risk assessment data is log-normally distributed or has skewed distributions and this will require different statistical methods for analysis. More detail is given on the US EPA website, accessible at <http://www.epa.gov/quality/qa_docs.html>.

A pragmatic approach used by many risk assessors is to exclude data values that are greater than 2 × SD from the mean, designating these as 'outliers'. However, this approach is not usually appropriate and great care must be exercised in excluding such outlying data. All data should be appropriately considered and not rejected without appropriate justification. In fact, outliers may provide information of great importance to the risk assessment, for example, if the data is revealing unknown contaminant sources or 'hotspots'. There are situations relating to the way contamination is spread across a site or the way the weather interacts with chemicals being released from a facility where it would be quite normal to see occasional results that are very high compared to the average. Such results are real and should not be excluded. Rather the story they tell about the risks posed by the situation should be investigated and understood.

Statistical tests can be used to identify potential outliers for further investigation but no data should be excluded without thorough review and consideration (NEPC 2010). Statistical advice on managing outliers in datasets is also provided in Barnett and Lewis (1994), while the US EPA website includes SCOUT Version 1.00.01 software for appropriately managing such datasets at <http://www.epa.gov/esd/databases/scout/abstract.htm>.

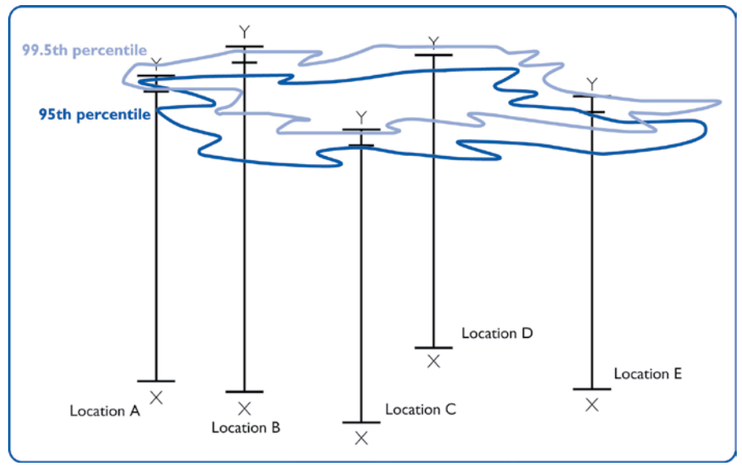
8.5.3
Contouring

Contour modelling (Figure 28) can provide a useful way of representing dispersion data (e.g. ground-level concentrations of airborne contaminants emitted from a point source, such as an industrial facility). While graphical representations of contours can provide useful information about situations such as the distribution and 'trends' of environmental hazards, poorly constructed contour maps are often based on extrapolations from an inadequate amount of data.

If the distribution of the environmental hazard is heterogeneous it is unlikely that there will be sufficient data points or sufficient associations between adjoining points for contouring to be used with any confidence in its meaning (e.g. most contaminated sites are likely to have a heterogeneous distribution of contamination). Where there is widespread or relatively homogeneous distribution of an environmental hazard, contours may provide fairly useful information on a macro scale. Examples are plumes of regional contamination such as around a lead smelter or sewer outfall.

When contouring is used (as with any model of data) there is a need to demonstrate that the model used is valid and to ensure that conclusions (hypotheses) drawn from it are tested.

Figure 28: Contour modelling of ground-level concentration (GLC) data



8.5.4
Data mapping

Mapping the results is essential, but poor design can cause clutter that obscures important data.

If there is 'too much' data available, this may be addressed by putting only significant results onto the map. However, this should be done cautiously, as 'censoring' some of the data can obscure trends.

'Normal' results are important if elevated results were anticipated and may need to be included to provide a useful comparison to the abnormal results. Other superficially unimportant data can provide surrogate information about environmental hazards.

A series of transparent overlays, each with a different dataset or a set of maps, each focusing on a different contaminant or section of the site, can be very useful to reduce cluttering.

8.5.5
Geographic information systems

Geographic information systems (GIS) allow spatial relationships between populations and hazards to be examined, and it can be useful for the hazard identification and exposure assessment phases of risk assessment.

Modern GIS tools allow visualisation of relationships between data in two or three dimensions (e.g. it can allow visualisation of certain symptoms or diseases in regard to their geographic location). The relationships may be between health, environment and socioeconomic data at many geographic scales, starting with the individual person (e.g. a person's place of residence or work). Data can be aggregated for a geographical area and patterns between geographical areas visualised.

GIS may also allow certain complex analyses to be done, such as shortest path or best path analysis. Path analysis allows predictions of population behaviour in relation to geographical variations. Path analysis will allow traffic flow patterns and densities to be predicted to assess the variations in exposures to benzene across a city.

In the case of a specific hazard, path analysis may allow exposure estimation to given pollutants, allowing opportunities for strategic public health interventions to be undertaken. For example, estimating shopping location patterns to identify the populations most likely to have been exposed to a *Legionella*-contaminated cooling tower or enabling preliminary rankings of risk for a number of towers when a case of *Legionella* is reported.

The Agency for Toxic Substances and Disease Registry (ATSDR) in the United States has used GIS to identify populations residing near hazardous waste sites.

8.6 SOME PRINCIPLES OF GRAPHICAL REPRESENTATION

Tufte (1983 p. 51) points out that 'graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space'. He goes on to say, 'Graphical excellence is the well-designed presentation of interesting data – a matter of substance, of statistics, and of design ... and consists of complex ideas communicated with clarity, precision, and efficiency.'

Tufte (1983;1990) censures cluttered tables and other failings of graphic design such as:

- excessive zeal in the use of computer software graphics packages so that bold cross-hatching and the use of wavy lines lead to 'optical art' effects
- overdoing the use of horizontal and vertical lines in tables.

Tufte also quotes Tschichol (1935): 'Tables should not be set to look like nets with every number enclosed.'

Some basic principles of graphic representation are given in Table 12.

For effective graphic presentation, Cleveland (1994) makes the following recommendations:

- Avoid excessively complicated graphs.
- Avoid pie charts, perspective charts (3D bar and pie charts, ribbon charts), pseudo-perspective charts (2D bar or line charts).
- Use colour and shading only when necessary, and then only very carefully.
- When possible, accompany graphs with tables of data.
- If probability density or cumulative probability plots are presented, present them with identical horizontal scales (preferably on the same page), with the mean clearly indicated on the curves.
- Do not depend on the audience to correctly interpret any visual display of data. Provide a narrative in the report interpreting the important aspects of the graph.
- Draw attention to any changes of scale on graphs:

Table 12: Useful versus not useful graphics

Useful	Not useful
No cryptic abbreviations	Numerous abbreviations requiring searching the text for explanation
No elaborate coding	
Words run in natural left-to-right direction	Words run vertically or in several directions Letters running vertically may be even worse
No elaborate shadings, cross-hatching and overpowering colouring	
Simple labelling or graphic means no legend or key is required	Elaborately or obscurely coded patterns require continual return to legend or key
Clearly printed	Murky or clotted printing
Enlightens and arouses curiosity	Repels interest and obscures meaning

Adapted from: Tufte 1983.

It needs to be absolutely clear when logarithmic rather than arithmetic scales are being used on the axes of graphs.

8.6.1 Cost of graphics

Graphic work is usually time-consuming and the cost of this may be significant. However, particularly for large and complex situations, some form of graphic representation is imperative for the assessor and other stakeholders to visualise accurately a model of what is occurring on a site. Without such representations, inaccurate (and probably costly) decisions will be made, and risk communication and community consultation will be much more difficult.

8.6.2 Photography

A photographic record that is well labelled for date, location and orientation is a valuable reference during the inspection (e.g. topography, soil staining, stack emissions, algal blooms, industrial processes, plant toxicity, proximity of housing) and assessment (e.g. soil strata demonstrated in test pits and soil cores). Good photography will provide considerable assistance for those unable to undertake an inspection of the situation.

8.6.3 Supplying data in electronic formats

Consultants, assessors and government agencies should have access to data on disc or other electronic formats as it:

- avoids a further source of transcription error
- facilitates the further analysis of data using other software packages.

8.7 CENSORED DATA

Where the analyte level is below the level of detection, data points are often reported as 'not detected' and referred to as 'censored' data. Therefore a transparent method of censoring such data becomes quite important. The influence of censored data on the entire dataset will depend on the proportion of data regarded as censored, and the magnitude of the level of detection compared with the levels that are of interest or concern. There are various ways of dealing with censored data, and these are described in Sections 4.6.4 and 8.7.3. Simple substitution by a number (e.g. 0, the level of detection, or level of detection divided by 2) may significantly affect any summary statistics (e.g. arithmetic means) used in the evaluation of the data. The values for the median and interquartile range generally are not affected by censored data.

8.7.1 Levels of reporting

The first step in dealing with censored data is to ensure the levels of detection or levels of reporting (LOR) are appropriate. The limit of determination (LOD) is the lowest concentration of a contaminant in the environmental medium (food, air, water or soil) that can be measured with confidence using the selected method of analysis. The limit of determination is usually different from the limit of detection (also sometimes termed LOD), so there is potential for confusion if only the acronym LOD is used without appropriate definition. The LOR must be less than the relevant criteria against which the results will be assessed. Preferably, a LOR should be no more than 10 per cent of the relevant criterion should be adopted. Where this may entail substantial costs, a higher level may be tolerable.

8.7.2 Diminishing levels of reporting

Improved analytical techniques have led to levels of reporting decreasing enormously. For example, the ability to measure benzene in water has increased by over 10,000-fold since the 1960s (Hrudey 1998). For some substances picogram concentrations (10^{-12}g) can be detected in commercial laboratories and femtogram (10^{-15}g) in research institutions.

8.7.3 Dealing with censored data

The four essential methods for dealing with censored data are outlined in Section 4.6.4.

Helsel (1990) recommends using robust methods, particularly when the data cannot be assumed to follow a defined distribution. This is also the position of the US EPA. Helsel concluded that using these methods, rather than simple substitution methods for environmental data, should reduce estimation errors for summary statistics substantially. Helsel also noted that 'simple substitution is an inappropriate method of dealing with less than detectable values as it has no theoretical basis' (Heyworth 1991 p. 25). Simple substitution methods of dealing with censored data may result in significant overestimates of risk if the level of reporting is used as a value for censored data and the concentration of the agent does not approach the level of reporting or is not present at all. Underestimates of risk can occur if, for example, a value of half the level of reporting is used but actual concentrations of the agent are actually greater than this.

Using either the distributional or robust methods is recommended, but the latter is preferred. Commonly available statistical packages readily enable the use of robust methods for dealing with censored data. Extensive information relevant to censored dataset users is also available from the US EPA website at <http://www.epa.gov/quality/qa_docs.html>.

In practice, and for convenience, many EHRAs use the simple substitution approach to data censoring. If this approach is used, substituting values less than LOR with half the LOR is most commonly used.

Irrespective of which censoring method is chosen, the EHRA report should contain a clear description and justification of the approach taken.

8.7.4 Sensitivity analysis

Sensitivity analysis is a process that enables the impact of uncertainty or imprecision in the input parameters selected in a risk assessment to be evaluated (see Section 5.15.3). In sensitivity analysis, the values of input parameters most likely to impact on the calculated outputs are varied in order to demonstrate the extent to which output parameters are changed. The results of a sensitivity analysis should be summarised in tables showing how the output variables may be altered if reasonable, relevant changes are made to input values. This should also assist with providing a 'reality check' for the input data used and on the outputs of the risk assessment.

When data censoring is done, a sensitivity analysis should be done on different methods to see how much difference the different methods cause.

Chapter 9: Evaluation of toxicity data

Toxicology studies have been designed to determine the toxic effects associated with exposure to chemical hazards. Such studies can provide information relating to toxic effects and potential health hazards likely to arise from single or repeated exposures, in terms of predicting potentially important toxicity endpoints and identifying potential target organs or systems.

This chapter on toxicological evaluation focuses on chemical hazards assessed using traditional toxicity testing methods and, in particular, on some of the problems and pitfalls that may arise during an assessment of possible compound-related changes in the parameters measured in toxicity studies. It is intended to provide guidance on the process of hazard identification and assessment.

The basic assumption is that traditional methods of toxicity assessment will continue to be the mainstay of EHRA for some time.

It is also important to note that, over time, the scientific community is gaining a better understanding of the mechanisms of toxicity, and this is leading to changes in both methodology and interpretation of toxicity data. It is inevitable that new paradigms will be introduced as science advances (see Section 9.4).

It is therefore important to acknowledge that the analysis and evaluation of toxicity studies reflects scientific consensus at the time the data is reviewed. This means that the toxicity studies underpinning many EHRAs may contain data generated during an era when toxicity testing and the interpretation of results were less well advanced.

9.1 TOXICITY TESTING – MAJOR *IN VIVO* STUDY TYPES

Hazard identification mostly relies on the results of *in vivo* toxicity studies conducted according to standard protocols. Guidance on the conduct of toxicity tests has been promulgated by the OECD (OECD 2009). There have been 53 OECD *Test guidelines* published since they were first promulgated in 1981, and many of these have been periodically updated.

The following types of studies are defined.

Acute toxicity studies are studies that investigate the effects of single doses of a substance. The LD_{50} test, or medium lethal dose test (OECD *Test guideline* TG401), which records gross toxicity and mortality data over a 14-day post-dosing period, has been commonly employed and may still be included in many data packages. However, TG401 was formally withdrawn by the OECD in 2002 in response to animal welfare concerns. Newer tests ('limit' tests and 'up-and-down' dosing methods) are now favoured as they reduce the numbers of animals required and reduce the suffering seen in the classical LD_{50} test. OECD TG420 covers acute oral toxicity determination by the 'fixed-dose method', TG 423 by the 'acute toxic class method', and TG 425 by the 'up-and-down procedure'.

The standard acute toxicity studies include tests for: acute oral, dermal and inhalational toxicity; eye irritation; skin irritation; and skin sensitisation. Such studies may serve as the basis for classifying and labelling a particular chemical or mixture, and serve as an initial guide to possible toxic modes of action and in establishing a dosing regimen in sub-chronic toxicity studies. Substantial work has been done to develop alternative tests (mainly *in vitro*) to replace skin/eye irritancy and sensitisation tests, and some of these have now been incorporated into the OECD *Test guidelines* series (e.g. TGs 429–435, and 437–438).

Sub-chronic toxicity studies are short-term repeat-dose studies. A short-term study has been defined (WHO 1990) as 'having a duration lasting up to 10 per cent of the animal's life span, 90 days in rats and mice, or 1 year in dogs', although the US EPA considers a 1-year dog study to be a chronic study. The main purpose of sub-chronic testing is to identify any target organs and to establish dose levels for chronic exposure studies.

Chronic toxicity studies, or long-term studies, are defined as studies lasting for the greater part of the life span of the test animals, usually 18 months in mice and 2 years in rats (WHO 1987; 1990). The OECD protocols for these studies may cover the investigation of chronic toxicity (TG452) or carcinogenicity (TG451), or both (TG453). All three of the OECD *Test guidelines* were updated in 2009 to better reflect developments in animal welfare and to improve dose selection.

Reproductive toxicity studies are studies designed to provide general information about the effects of a test substance on reproductive performance in both male and female animals, such as effects on mating behaviour, gonadal function, oestrous cycling, conception, implantation, parturition, lactation, weaning and neonatal mortality. These studies may also provide some information about developmental or teratogenic effects of the test substance. The conduct of and the results from these studies are very important to assess with care, since the reproductive process is critical for perpetuation of the species and factors or agents that alter or disrupt this process can have devastating consequences, both in fact and in public perception (Korach 1998). For information on study design, refer to OECD *Test guideline* 415, *One-generation reproduction toxicity study* and *Test guideline* 416, *Two-generation reproduction toxicity study*: (OECD 2009). For guidance on evaluating reproductive toxicity studies, refer to IPCS EHC 225 *Principles for evaluating health risks to reproduction associated with exposure to chemicals* (WHO 2001).

Developmental toxicity studies are studies that examine the spectrum of possible *in utero* outcomes for the conceptus, including death, malformations, functional deficits and developmental delays (Tyl & Marr 1997). Exposure during sensitive periods may alter normal development resulting in immediate effects, or may subsequently compromise normal physiological or behavioural functioning later in life. Since some developmental processes can occur perinatally or postnatally, protocols for developmental studies are being modified and extended to address developmental toxicity during the period covering major organogenesis as well as covering the perinatal and early postnatal period. This could include delayed toxicity associated with epigenetic effects during sensitive phases of foetal development. Such attention to the critical timing of exposure also accords with a growing emphasis on understanding early-life susceptibility to carcinogenesis (see Section 5.5.2).

Genotoxicity studies are designed to determine whether test chemicals can perturb genetic material to cause gene or chromosomal mutations. A large number of assay systems, especially *in vitro* systems, have been devised to detect the genotoxic or mutagenic potential of agents (IARC 1999). Most authorities consider that a minimum set of data is required to define a mutagen/non-mutagen. This data usually consists of gene mutations in bacteria and mammalian cells, and *in vitro* and *in vivo* cytogenetics. Newer assays that could provide additional information include the comet assay, mutations in transgenic animals, fluorescent *in situ* hybridisation and cell transformation. Guidance on the conduct and interpretation of *in vivo* and *in vitro* genotoxicity assays, and integration of their results, is also available in a UK Department of Health document (see COM 2000). Interpretation of the results of *in vitro* genotoxicity tests for the purposes of identifying potential human genotoxins and, by inference, potential human carcinogens, needs to be done within a well-defined science policy context (Thybaud et al. 2007).

Other tests: The OECD *Test guideline* series now includes special tests for such endpoints as neurotoxicity (TG424) and developmental neurotoxicity (TG426). It has also addressed animal welfare issues through the development of a range of validated short-term *in vivo* tests and *in vitro* tests, which may complement, or possibly substitute for, the conventional animal tests that have been used for many years. These include tests for skin absorption (TG428) and tests for endocrine-related endpoints (*in vivo* tests TG440, 441 and *in vitro* test TG455).

9.2 GUIDANCE ON EVALUATING AND INTERPRETING TOXICITY TESTS

Supplementary guidance on the evaluation and interpretation is provided in more detail in Appendix 1. This guidance is aimed primarily at experienced toxicologists who may be asked to provide a weight-of-evidence (WoE) analysis of the extent to which toxicity tests are able to define the hazard identification component of an EHRA, and to provide useful information on dose–response relationships. It may also be of value to less experienced readers seeking further detail on using conventional toxicity tests.

9.3 EVALUATING THE WEIGHT OF EVIDENCE AND CONSIDERING THE TOXICOLOGY DATABASE IN TOTO

The essential purpose of toxicity studies is detecting valid biological evidence of the hazard potential of the substance being investigated. Evaluation of the weight of evidence (WoE)[3] produced by toxicity studies is the process that considers the cumulative data pertinent to arriving at a level of concern about the potential adverse effects of a substance. It is composed of a series of judgements concerning the adequacy, validity, and appropriateness of the methods used to produce the database, and those judgements that bring into causal, complementary, parallel or reciprocal relationships, all the data considered. Because our knowledge about mechanisms of toxicity is still developing, because good epidemiological evidence is seldom available, and because animal studies are not always conclusive, the information available at a given time may provide only 'persuasive' rather than 'hard' evidence of a defensible presumption (one way or the other) about the potential health effects of a substance under given conditions of exposure. Therefore, it is necessary to succinctly discuss the rationale for judgements and conclusions contained in risk assessments together with any associated uncertainties. This becomes important when new data or new scientific knowledge requires re-evaluation of the database or a change in a previous risk assessment or regulatory action.

3 Strength of evidence' is commonly taken to mean the degree of conviction regarding the outcome of an experiment such as NTP's 'clear evidence', 'some evidence', 'equivocal evidence' and 'no evidence' of carcinogenicity. 'Weight of evidence' involves integration of all available data, not just one study.

There is no acceptable substitute for informed judgement (based on sound scientific principles) in analysing, evaluating, interpreting and weighing biological and toxicological data derived from animal toxicity studies conducted according to currently available protocols.

In addition to identifying toxic effects and the doses at which these effects do or do not occur, toxicity studies may yield insight into the mode (or mechanism) of action (MoA) of a chemical toxicant. Assessing MoA is becoming a fundamental component of carcinogenic risk assessment, especially when it comes to classifying carcinogens (see Section 11.3), and making judgements about whether a threshold or non-threshold approach to dose–response assessment is warranted (see Chapters 3 and 5). The evaluator may be able to combine information from a number of studies within the database (e.g. metabolic/toxicokinetic, acute, short-term repeat-dose, sub-chronic, chronic/carcinogenicity, developmental, reproductive and genotoxicity studies), studies using genetically modified test mice and quantitative structure–activity relationship (QSAR) analysis to adduce information about the mode or mechanism of toxic action of the substance.

It is at the point of overviewing the entire toxicology database that the WHO/IPCS *Conceptual framework for cancer risk assessment* (see Section 11.5) is intended to be applied. This 'framework' is an analytical tool providing a logical, structured approach to assessing the overall WoE for a postulated mode of carcinogenic action. Using the framework should increase the transparency of the analysis by ensuring that the facts and reasoning have been documented clearly, including any inconsistencies and uncertainties in the available data.

Note that although the conceptual framework has been developed to assist in assessing carcinogenic endpoints, the principles upon which it is based are broad, and should enable its use in analysing modes of action of non-neoplastic effects of chemicals. Irrespective of the nature of the disease process, characterising the MoA will facilitate subsequent judgements about the human relevance of the toxicological findings, the possible need for further data, risk quantification, and setting appropriate regulatory standards for the chemical.

9.4 DEVELOPMENTS IN TOXICITY TESTING

Environmental health risk assessment has traditionally been based on evaluating toxicity data generated according to the above principles, the knowledge base of which dates back several decades. However, it has long been recognised that interpreting classic toxicity studies may be compromised by the high doses used and the uncertainties in extrapolating effects to the lower doses relevant to environmental exposures. Furthermore, animal welfare issues are likely to place increasing restrictions on using live animals in toxicity testing.

Recognising these as important issues, the US EPA and the US NIEHS asked the NRC to develop new guidance on toxicity testing, which would incorporate new *in vitro* and *in silico* technologies and computational systems biology. An interim report of this project (*Toxicity testing in the 21st century: a vision and a strategy*) was published in 2006 (NRC 2006) and a final report in 2007 (NRC 2007). This has been followed by a review of some of the challenges yet to be overcome and some practical approaches to implementing this visionary program (Andersen & Krewski 2009).

There is no doubt that this program, now called Tox21, is an important driving force for the review and refinement of conventional toxicity testing. The program is a collaboration between the US EPA, National Institutes of Environmental Health Sciences/National Toxicology Program, National Institutes of Health/National Human Genome Research Institute, and the Food and Drug Administration. < http://www.epa.gov/ncct/Tox21>.

The project aims to:

- research, develop, validate and translate innovative chemical testing methods that characterise toxicity pathways
- research ways to use new tools to identify chemical induced biological activity mechanisms
- prioritise which chemicals need more extensive toxicological evaluation
- develop models that can be used to more effectively predict how chemicals will affect biological responses
- identify chemicals, assays, informatic analyses, and targeted testing needed for the innovative testing methods
- be able to provide the data generated from the innovative chemical testing methods to risk assessors to use when making decisions about protecting human health and environment.

Andersen and Krewski (2009) emphasised the importance of relating events leading to toxicity in the context of perturbations in biological functions, some of which may be reversible or capable of adaptive change. In relation to dose–response assessment, greater emphasis was placed on using computational techniques (e.g. PBPK models) to determine human exposure scenarios likely to result in target tissue concentrations associated with critical events in the transition from normal to abnormal biological function (i.e. toxicity). The use of QSAR and 'read across' techniques based on the known toxicological profiles of reference compounds is also on the increase, and can make a welcome contribution to toxicological knowledge, bypassing some of the costs and time constraints of conventional toxicity testing.

Another review (Creton et al. 2010) has focused on using acute toxicity testing, and the possibility of replacing *in vivo* tests with a range of alternative approaches. While recognising that acute oral toxicity tests are important for toxicity categorisation, Creton et al. (2010) point out that requiring additional studies of dermal toxicity rarely adds anything to the process. Furthermore, while *in vivo* tests for sensitisation potential probably remain important for categorisation purposes, developing predictive *in vitro* tests for skin/eye irritancy means that conventional *in vivo* Draize-type testing is no longer justifiable.

While it is clear that development of alternative tests has been a high priority for work and has progressed well for acute endpoints, validation of the alternative test methods for assessing eye irritancy are not as advanced as those for skin irritancy. Satisfactory tests for skin sensitisation have yet to be validated.

Morisseau et al. (2009) evaluated the efficacy of a range of high-throughput *in vitro* enzyme activation and receptor assays to predict toxicity potential. They concluded that such screening assays have some useful potential for prioritising chemicals for further testing.

The potential for incorporating new types of testing strategies for assessing genotoxicity has been reviewed by Elespuru et al. (2009) in the context of criticisms of the over-interpretation of positive results from the standard batteries of *in vitro* tests.

Chapter 10: Use of epidemiological data

10.1 INTRODUCTION

Epidemiology and toxicology are complementary in risk assessment. Epidemiology is the direct human evidence component and, if based on sound epidemiological methods, can provide the most important evidence in characterising risk. Epidemiology is the principal driver in microbiological risk assessment, but it can assist the formal framework of EHRA at all of its stages.

The term 'risk' tends to be used in a subtly different and more specific way in epidemiology than in risk assessment. In epidemiology, risk describes the 'frequency of occurrence of a disease in one population compared with another, either as a difference in rates (attributable risk) or as a ratio of rates (relative risk)' (ACDP 1996 p. 20). The feature distinguishing the two populations by its presence or distribution is referred to as a 'risk factor'. The reliance on comparisons of disease rates between populations creates substantial limitations for the sensitivity of relative risk determination for common diseases (Thomas & Hrudey 1997).

Epidemiology is 'the study of the distribution and determinants of health-related states or events in specified populations, and the application of the study to the control of health problems' (Last 1988). A more recent description of epidemiology by the same author states:

> 'Epidemiology connects the dots, the isolated bits of information that begin to form a coherent pattern when connected in the right way ... The dots can come from anywhere. Identifying them demands a broader perspective, the ability to see the Big Picture ... Sometimes the way the dots are connected is instantly apparent. Sometimes painstaking investigation and analysis are required, for instance when the problem is a common but ill-defined condition ... caused by trace amounts of a highly reactive environmental toxin …' (Last 2010)

In simple terms, epidemiological methods compare health outcomes or calculated risk estimates in an exposed population or group, with those in a non-exposed population. The types of epidemiological studies that may be useful in making such evaluations are discussed briefly in Section 10.2.

Environmental epidemiology is considered to be a subspecialty of epidemiology that addresses the effects of environmental exposures on health and disease in the population (Baker & Nieuwenhuijsen 2008).

Epidemiological methods can be used to investigate the cause of adverse health effects, the natural history of health conditions and the description of the health status of populations, and to evaluate health-related interventions (Bonita et al. 2007). They also allow for the control for possible confounding factors that may influence the results. In the context of environmental health, epidemiological methods may also be used to characterise population exposures, investigate perceived clusters of disease, to develop health surveillance programs, to establish a baseline, and to monitor the consequences of risk management activities.

At the same time, there are often unrealistic expectations of what an epidemiological study may be able to achieve.

In epidemiological studies of the potential influence of environmental factors, the exposures encountered are rarely extreme, in contrast to the high levels of exposure that can be applied in animal-based toxicity tests. Such epidemiological studies are often confounded (see Section 10.3), and any associations found are commonly weak. Furthermore in large populations substantial effects in small numbers of sensitive individuals may be 'swamped' by a lack of effect in the majority.

Epidemiological studies are often most unreliable when measuring subtle effects, although these effects may be of public health significance when large populations are exposed. Even the most rigorously conducted studies are unreliable for detecting small increases in risk. Since it is impossible to prove an absence of risk from any human study, it is often considered that the principal value of epidemiology is to exclude major health effects of an exposure (NHMRC 2006).

The purpose of this chapter is to provide a basis for understanding the strengths and weaknesses of epidemiology in supporting risk assessment. As Mundt et al. (1998) noted that if the limitations of epidemiological studies are not understood by the risk assessment team, the validity of an assessment might be compromised by including inappropriate, possibly misleading, epidemiological data. In discussing the potential for epidemiological studies to throw up false negative outcomes, particularly in relation to associations with cancer, Boffetta et al. (2008) made a plea for epidemiologists to practise 'epistemological modesty' in the types of claims or conclusions drawn from their studies.

The systematic appraisal of epidemiological studies is intended to answer the question: 'Is there any other way of explaining the set of facts before us [i.e. the study results, is there any other answer equally, or more, likely than cause and effect' (WHO 2000a). Alternative explanations may result from chance, bias and confounding.

10.2 TYPES OF EPIDEMIOLOGICAL STUDY

Broadly speaking, epidemiological activity can be either 'descriptive' (reporting and describing the distribution of exposure and effect) or 'analytical' (designed to analyse and understand the degree of association between exposure and effect). Descriptive studies include case reports, case series and cross-sectional surveys. Cross-sectional surveys examine exposure and disease prevalence at the same point in time (or over a short duration of time) and thus are unable to support causal inference because it is not possible to know if exposure pre-dated the onset of disease. Similarly, case reports and cross-sectional surveys are unable to support causal inference.

In practical terms in environmental epidemiology, there are four main categories of analytical study (from Moolgavkar et al. 1999):

- case-control studies
- cross-sectional studies
- cohort (longitudinal) studies
- ecological studies (including a special sub-group known as time-series studies).

Cohort, cross-sectional and case-control studies differ from ecological studies in that information on exposure and disease is available on an individual basis. With ecological studies this information is only available on a group basis, so the community or region is the unit of analysis.

In **case-control studies**, exposure and other attributes of cases of the disease under investigation are compared with those from a suitable control or comparison group of persons unaffected by the disease, and analysed to yield effect estimates. The approach is to start with a diseased group and look backwards (retrospective cohort) to their history of exposures, and compare these exposures with exposures among a non-diseased cohort. The selection of appropriate controls to avoid bias is a significant challenge with case-control studies. However, among their advantages, case-control studies are relatively inexpensive, ideal for studying rare diseases and useful for investigating multiple, different exposures (Gregg 1996).

Cross-sectional studies measure the prevalence of disease and measure exposure and effect at the same time. They are relatively easy and economical to conduct and are particularly useful for measuring fixed characteristics of individuals such as socioeconomic status (Bonita et al. 2007).

Cohort studies follow cohorts or groups of individuals, defined in terms of their exposures, over time to see if there are differences in the development of new cases of the disease of interest (or other health outcomes) between the groups with and without exposure. Such studies can be carried out by either reviewing past records (retrospective) or by tracking people into the future (prospective cohort). The essential feature of these longitudinal studies is that for each individual prior exposure information can be related to subsequent disease experience (Breslow & Day 1987).

Ecological studies involve investigating a group of people such as those living within a geographical area (e.g. a region, state or territory). For example, place and time of residence can provide aggregate exposures, so may be used to create surrogate measures of the real exposure of interest (Elliott et al. 1992). Rates of disease and average exposure levels to a particular agent are determined independently, and on a group basis. This may give rise to spurious apparent correlation, called the 'ecological fallacy'. Because it is not ascertained whether individuals who have been exposed to the agent are the same individuals who developed the disease, statements about causal associations are inappropriate.

However, ecological studies are relatively inexpensive for linking available health datasets and environmental information and are useful for hypothesis-generation (Yassi et al. 2001). Examples of ecological studies are the assessments of the relationship between tobacco sales in different countries and lung cancer rates, and fluoride in water supplies and dental caries.

A subset of ecological studies, known as 'time series studies', is regarded as very helpful in understanding the influence of short-term fluctuations in air pollutants on day-to-day changes in population morbidity and mortality after controlling for factors such as season and air temperature. However, disentangling the effects of individual pollutants as measured in a mixture such as urban air pollution can be quite difficult.

To strengthen the design of ecological studies, Nurminen (1995) recommended selecting areas with populations that:

- are homogeneously exposed (to minimise within-area exposure variation)
- represent different extremes of exposure distribution (to maximise between-area exposure variations)
- are comparable with respect to co-variate distributions (e.g. socioeconomic status, demography)
- use the smallest possible sampling units for ecological analysis.

The largest number of environmental epidemiology studies found in the literature are of the ecological or cross-sectional type, because they are easier to carry out and cost less (Thomas & Hrudey 1997). However, as noted above and discussed further below in relation to assessing causality, such studies may be useful for identifying potential hazards or hypothesis generation, but they cannot determine cause and effect.

Characteristics of the various study types are summarised in Table 13.

Table 13: Study designs in environmental epidemiology that use the individual as the unit of analysis

Study design	Population	Exposure	Health effect	Confounders	Problems	Advantages
Case reports, case series and other descriptive studies	Community or various sub-populations	Records of past measurements	Mortality and morbidity statistics; case registers; other reports	Difficult to sort out	Hard to establish exposure-effect relationships	Cheap; useful to formulate hypotheses
Cross-sectional study	Communities or special groups; exposed versus non-exposed	Current	Current	Usually	Current exposure may be irrelevant to current disease	Can be done quickly; can use large populations; can estimate prevalence
Case-control study	Diseased (cases) versus non-diseased (controls)	Records or interview	Known at start of study	If confounders can be identified and measured they may be addressed	Difficult to generalise; may incorporate biases; cannot derive rates Recall bias is a major problem	Relatively cheap and quick; particularly useful for studying rare diseases
Time-series study	Large community (several million people); susceptible groups such as asthmatics	Current (e.g. daily) changes in exposure	Current (e.g. daily) variations in mortality	Often difficult to sort out; e.g. effects of influenza	Many confounding factors; often difficult to measure	Useful for studies on acute effects
Historical (retrospective) cohort study	Special groups; workers, patients, insured persons	Records of past measurement	Records of past or current diagnosis	Often difficult because of retrospective nature; depends on availability of previously obtained data	Need to rely on records that may not be accurate	Less expensive and quicker than a prospective study; can be used to study exposures that no longer exist
Prospective cohort study	Community or special groups; exposed versus non-exposed	Defined at outset of study (can change during study)	To be determined during study	Usually easy to measure	Expensive and time consuming; exposure categories can change; high dropout rate possible	Can estimate incidence and relative risk; can study many diseases in one study; can describe associations that suggest cause–effect relationships
Experimental (clinical/ intervention) study	Community or special groups	Controlled or already known	To be measured during study	Can be controlled by randomisation of subjects	Expensive; ethical considerations; study subjects compliance required	Well accepted results; strong evidence for causality or efficacy of intervention

Adapted from: WHO (1991).

Epidemiological studies are rarely definitive, and a single epidemiological study cannot establish causality. Any such study is necessarily a sample of the total population of interest, so there will always be questions about being able to generalise from an individual study sample to the total population. Moreover, undetected methodological biases can only be overcome by having numerous studies by different investigators with different population samples. If consistency of outcome is demonstrated across many studies, the causal hypothesis becomes more likely. A 'weight of evidence' (WoE) approach is generally required, involving the interpretation of all available information and consideration of the strengths and weaknesses of each study.

Unfortunately, experimental interventions such as randomised controlled trials are rarely available to assist environmental health risk assessment, as it is not ethical to purposely expose populations to hazardous risks. It is appropriate, however, to expose groups to a lesser level of risk via a clean-up intervention. An example of an experimental intervention is a randomised trial of lead abatement procedures undertaken in Broken Hill (S. Corbett, personal communication). Epidemiological studies, depending on their design, may serve two purposes: hypothesis generation or assessment of a causal relationship. Their ability to evaluate a causal relationship may be limited by a lack of control of potential confounders or a lack of power (usually the result of limited sample sizes) (Samet et al. 1998).

10.2.1 Observational studies

Different observational study designs have different applications, advantages and disadvantages (see Tables 14 and 15). These comparisons assume the different types of studies are equally well designed. Even so, design variations may affect their performance and provide exceptions. Bonita et al. (2007) provides a more detailed description.

Table 14: Applications of different observational study designs

	Ecological	Cross-sectional	Case-control	Cohort
Investigation of rare disease	++++	–	+++++	–
Investigation of rare cause	++	–	–	+++++
Testing multiple effects of cause	+	++	–	+++++
Study of multiple exposures and determinants	++	++	++++	+++
Measurement of time relationships	++	–	+[a]	+++++
Direct measurement of incidence	–	–	+[b]	+++++
Investigation of long latent periods	–	–	+++	+[c]/-

Key:

+ to +++++ indicates the degree of suitability from least to most suitable; – indicates not suitable

a if prospective; b if population-based; c if retrospective or historical cohort study.

Table 15: Advantages and disadvantages of different observational study designs

	Ecological	Cross-sectional	Case-control	Cohort
Probability of:				
Selection bias	N/A	Medium	High	Low
Recall bias	N/A	High	High	Low
Loss to follow-up	N/A	N/A	Low	High
Confounding	High	Medium	Medium	Low
Time required	Low	Medium	Medium	High
Cost	Low	Medium	Medium	High

Tables 14 and 15 adapted from: Bonita et al. 2007.

10.3 INVESTIGATION OF APPARENT CLUSTERS

The assessment of an apparent cluster of non-communicable disease is a complex and resource-intensive task. It commonly involves investigation of a number of reported cases of cancer, or some other adverse health effect likely to be linked to an environmental exposure. It should be managed using a multidisciplinary approach using standardised analytical tools. The trigger for a cluster investigation is often the 'perception' that the incidence of the disease is unusually high in a region or scenario linked to a possible environmental exposure source (e.g. near a waste dump, or in an occupational setting). The assessment is usually first centred on whether the observed number of cases is consistent with that expected from the background incidence, or whether there is a sufficiently common pattern to the nature of the cancers or other health effect.

Specific guidance on cluster investigation has been developed by Queensland Health (2009), including suggested criteria for decision making. The NHMRC is working towards the development of guidelines that can be adopted nationally.

10.4 BIAS AND CONFOUNDING: KEY CONCEPTS IN ENVIRONMENTAL EPIDEMIOLOGY

There are many ways in which error can be introduced into epidemiological studies. Error may be random (due to chance alone, and potentially reduced by improving sample size) or systematic (and not reduced by increasing sample size). While this section does not attempt to deal with the subject of systematic error in any depth, the two key concepts of bias and confounding must be highlighted. The size of the statistical confidence intervals will provide an indication of the potential for random sampling error, but statistical confidence intervals do not represent uncertainty arising from bias or confounding.

Bias occurs when there is a systematic tendency by a study to produce results that diverge from the truth. There are many sources and varieties of bias, but the most important include selection bias and measurement (or classification) bias. Bonita et al. (2007) include a succinct account of bias. It may be difficult to precisely estimate the effect bias has in a study, but it is vital for risk assessors to look for and attempt to identify the potential size and direction of bias in interpreting a study's findings. Recall bias is important because case-control studies are much more common for environmental epidemiology than are cohort studies.

Confounding is the distortion of the effect of the agent of interest by an extraneous factor (Moolgavkar et al. 1999). This may occur if another exposure exists in the study population that is associated with both the disease (or outcome) and the exposure being studied, such as a third factor ('confounding variable') that independently affects the risk of developing the disease.

Three conditions are necessary for a factor to be categorised as a confounder:

- It must be a risk factor for the disease in the absence of the exposure under study (it does not have to be an actual cause but can be a marker of an actual cause).
- It must be associated with the exposure in the study population.
- It must not be affected by the exposure or disease. (In particular, it cannot be an intermediate factor in the causal pathway between exposure and disease. An intermediate factor is one that is caused by the exposure and, in turn, causes the disease.)

There are specific approaches for controlling confounding that can be used in both the design and analysis of analytic studies, providing that the confounding variables have been identified and measured. Randomisation of exposure can control confounding, however this option is not usually available because exposure tends not to be planned by the researcher. Matching groups can, however, remove confounding effects, for example, by investigating people of similar age groups or gender, or by comparing smokers with smokers.

10.5 ASSESSING THE RELATIONSHIP BETWEEN A POSSIBLE CAUSE AND AN OUTCOME

A cause is 'an event, condition, characteristic or a combination of these factors which plays an important role in producing the disease' (Bonita et al. 2007).

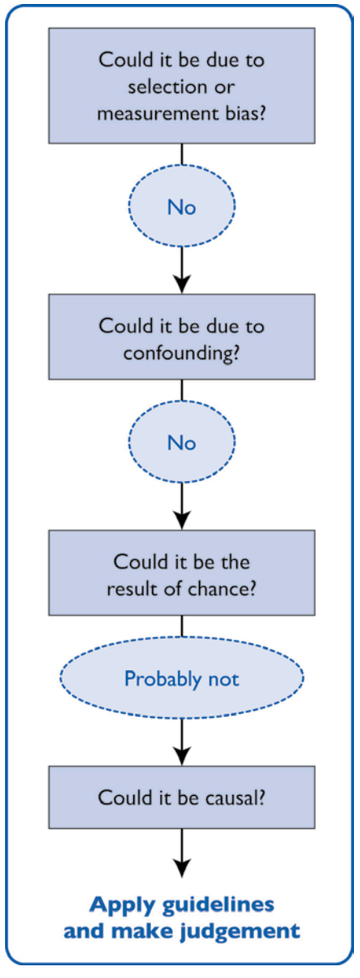
Causation of adverse health effects is affected by four types of factors:

- predisposing factors (e.g. immune deficiencies, gender and previous illness)
- enabling factors (e.g. poor nutrition and bad housing may favour the development of disease)
- precipitating factors (e.g. the exposure to a specific disease agent)
- reinforcing factors (e.g. repeated exposure may aggravate an established disease or state).

The term 'risk factor' is commonly used to describe factors that are positively associated with the risk of development of a disease but that are not sufficient in themselves to cause the disease. A 'sufficient' cause is one that inevitably produces or initiates a disease, and a 'necessary' cause is one for which a disease cannot develop in its absence (Bonita et al. 2007). In the biological sciences, there is often a constellation of components acting in concert for a cause to create an effect, and many of the components of a 'sufficient cause' may be unknown (Rothman & Greenland 1997). At the low levels of exposure commonly encountered in the environment and where there may be a range of contributory factors present, it may be difficult or inappropriate to assign this nomenclature to an agent even though the agent is accepted as causing a specific effect with high exposures.

As with other scientific disciplines, epidemiology has attempted to define a set of causal criteria to help distinguish causal from non-causal associations. In the first place, other explanations for a potentially causal association must be excluded (such as chance, selection or measurement bias, or confounding). Particularly rigorous scrutiny should be given to studies giving a positive but not statistically significant result. Figure 29 illustrates this process.

Figure 29: Assessing the relationship between a possible cause and an outcome when an association is observed



If alternative explanations such as bias and confounding can be excluded, it is then useful to systematically apply guidelines for assessing causation as shown in Table 16. The concepts in these guidelines derive from work by Hill (1965) and others. However, as Rothman and Greenland (1997) note, apart from temporality (whereby a putative cause must precede the effect), there are no necessary and sufficient criteria for determining whether an observed association is causal. Lucas & McMichael (2005) provide a useful summary of these issues and conclude that, as with statistical *p* tests, the criteria of causality must be viewed as aids to judgement, not as arbiters of reality. Thus the term 'guidelines' is more appropriate than the slightly more absolute 'criteria', and there is not necessarily an easy epidemiological roadmap to finally determine causation.

The guidelines are similar in some respects to Koch's postulates for determining whether an organism is causal of a particular disease. He required that the putative organism was to be found in every case of the disease, and that it could be isolated and grown from cases of the disease. Further, the isolate should produce a like disease in a new host and in turn the organism should be recovered from that case. Modern microbiology is now finding instances where these postulates cannot be fully complied with (e.g. the organism can be detected using molecular methods, but cannot be isolated or grown in the laboratory). Notwithstanding this, causality may be imputed by the strength and consistency of evidence.

With environmental health in particular, much decision making rests on a WoE approach rather than definitive proof of cause, which is commonly not available – hence the concept, 'Judging the evidence' in Table 16, is particularly relevant in assessment of risk.

The guidelines are arranged in a logical sequence for making judgements on causality. They are not weighted equally, and their relative contribution to a final judgement will vary from one situation to another (Thomas & Hrudey 1997). Consistency can be demonstrated if several studies give the same result, especially if a variety of designs is used in different settings since this reduces the likelihood that all studies are making the same mistake. However, other factors such as different exposure levels or study conditions may need to be taken into account, and the best designed studies should be given the greatest weight. It is important to note that in environmental epidemiology, reliance on a single pivotal study is the exception rather than the rule.

Table 16: Guidelines for assessing causation

Temporal relation	Does the cause precede the effect? (essential)
Plausibility	Is the association consistent with other knowledge? (mechanism of action; evidence from experimental animals)
Consistency	Have similar results been shown in other studies?
Strength	What is the strength of the association between the cause and the effect?
Dose–response relationship	Is increased exposure to the possible cause associated with an increased effect?
Reversibility	Does removal of the possible cause lead to a reduction of disease risk?
Study design	Is the evidence based on a strong study design?
Judging the evidence	How many lines of evidence lead to the conclusion?

Adapted from: Bonita et al. 2007.

The technique of meta-analysis grew out of the need to reduce random error in clinical trials. Meta-analysis in the context of systematic reviews can be used to pool the data from well-designed studies, each of which may deal with a relatively small sample size, in order to obtain a better overall estimate of effect. Meta-analysis has pitfalls if poor quality studies are included, and needs to be applied with caution to observational studies (which are less able to control for confounding than randomised trials). Standard methods for conducting and reporting systematic reviews have been published (Greenhalgh 1997). See NHMRC (2000) *How to review the evidence: Systematic identification and review of the scientific literature*.

The strongest evidence comes from well-designed and competently conducted randomised controlled trials. The NHMRC (1999b) places strongest emphasis on evidence obtained from systematic reviews of all relevant (and well-conducted) randomised controlled trials ('Level I').[4]

However, there are relatively few such trials available for environmental health hazards that could form the basis for a systematic review. Most apply to the effects of treatment or prevention campaigns. A rare example is the Melbourne Water Quality Study, which was a blinded study involving real and sham domestic reverse osmosis water filters and an assessment of acute gastrointestinal disease (Hellard et al. 2001).

In practice, most evidence comes from observational studies (e.g. nearly all the evidence on the health effects of smoking).

4 The NHMRC document referred to is oriented towards clinical interventions and clinical practice guideline development. At present there are no comparable endorsed 'levels of evidence' to guide assessment of epidemiological evidence for environmental health practice, although in 2009 the NHMRC published a broader approach to evidence grading for use in developing guidelines. Available at <http://www.nhmrc.gov.au/guidelines/developers.htm>.

In well-conducted cohort studies, bias is minimised. Case-control studies are subject to several forms of bias and weaknesses related to time sequence but, if well designed, may still provide useful evidence for the causal nature of an association. Cross-sectional studies are weaker as they provide no direct evidence on the time sequence of events.

Ecological studies are the least satisfactory because of the dangers of incorrect extrapolation to individuals from data derived from regions or countries. However, where certain exposures cannot normally be measured individually (e.g. air pollution, pesticides residues in food, fluoride in drinking water) evidence from ecological studies may be important in environmental health decision making. Time series studies demonstrating health outcomes associated with fluctuating air pollutant levels may be one particularly useful example.

The above principles about strength of evidence obtained from various study types are summarised in Table 17.

Table 17: Relative ability of different types of study to 'prove' causation

Type of study	Ability to 'prove' causation
Randomised controlled trials	Strong
Cohort studies	Moderate
Case-control studies	Weak/moderate
Cross-sectional studies	Weak
Ecological studies	Weak

Adapted from: Bonita et al. 2007.

The ranking in this table assumes that studies are well designed and well conducted in each case. Even the presence of a strong ability to 'prove' causation should be supplemented by mechanistic knowledge to be confident of causation.

10.6 THE STRENGTHS AND LIMITATIONS OF OBSERVATIONAL EPIDEMIOLOGY VERSUS EXPERIMENTAL TOXICOLOGY

Epidemiological studies are crucial for assessing effects directly in humans and estimating population attributable risks. However, their power of resolution is limited, mainly because of the difficulties in estimating exposure precisely and in controlling bias. Toxicological studies are necessary for elucidating causal mechanisms, which may be important for determining dose–response relations and extrapolating to low doses in risk assessment. On the other hand, direct generalisations to humans based on animal data are often uncertain (Pershagen 1999).

Epidemiological studies are often given increased weighting because they come from humans but, compared with toxicological studies of animals, may be more costly and time consuming and more likely to result in ambiguous findings (Samet et al. 1998). However, substantive findings have been obtained at times through opportunistic study of highly exposed groups – such as occupational cohorts or communities that have been inadvertently exposed to contaminants (e.g. via food or water). These can be either observational epidemiological studies or what Lilienfield and Lilienfield (1980) called 'natural experiments'.

An appropriate means of integrating information derived from epidemiological and toxicological studies would be a significant step forward in assisting the processes of risk assessment. The development of a framework for such integration has recently been proposed by Adami et al. (2011). The key elements of this framework are:

- ensuring that all relevant studies are collected and collated
- assessment of the quality of the studies
- evaluation of the weight of evidence (WoE) each brings to the conclusions; whether the study can be categorised as providing 'acceptable' results from a WoE perspective, 'supplemental' results that have limited use in a WoE approach, or 'unacceptable' results from a WoE perspective
- assignment of a scale value to the conclusions, that can be related to grid values that rank the epidemiological evidence from 'for' to 'against' a causal hypothesis, and the toxicological evidence as providing 'high' or 'low' evidence for biological plausibility of the findings
- assignment of the conclusion to a position on a causal relationship grid; with the grid categorising the evidence for a causal relationship into quadrants, labelled 'likely', 'unlikely' or 'uncertain'.

An illustration of the application of this framework to assess the strength of epidemiological and toxicological evidence linking the herbicide atrazine to cancer causation in humans was presented in a companion paper (Simpkins et al. 2011).

10.6.1 Hazard identification

Since human risk is assessed directly, epidemiology has a number of potential advantages over animal toxicology in the area of hazard identification:

- differences between humans and animals in relation to absorption, metabolism, detoxification and excretion no longer need adjustment or consideration
- sample sizes for most types of epidemiology studies are large in comparison to the number of animals used in animal studies (usually no more than 10–50 of each sex per dose)
- genetic diversity may be broad in humans compared with the more restrictive phenotypes of selectively bred animal strains used in toxicological studies
- epidemiological studies may include different groups (e.g. the young, old and susceptible) that may not be included in the usually relatively homogeneous groups used in toxicological studies
- effects on some aspects of mental function or behaviour, and more subjective effects such as nausea or headache, can be better assessed in human studies.

Differences in hazard identification based on epidemiological and toxicological data may be seen in the matter of 'site concordance'. The epidemiological data may suggest lung cancer is of concern, whereas the toxicological data may suggest liver cancer. Similar conflicts can arise where there are suggestions of a problem from epidemiological data unsupported by toxicological evidence (Samet et al. 1998).

10.6.2 Dose–response assessment

Controlled exposure studies offer the only real possibility of accurately assessing dose–response relationships in human studies. These may occasionally still be done under rigorous ethical controls, with some types of chemicals having minimal toxicity potential. Examples include air chamber studies with mildly irritant or odoriferous gases and vapours, where the objective is to assess thresholds for the onset of effects or sensory perception and for self-limiting, non-lethal microbial pathogens like *Cryptosporidium* using healthy volunteers.

Epidemiological data may still assist in assessing dose–response relationships, even when controls over exposure in human studies are unavailable.

Areas where epidemiology may help to inform dose–response relationships from animal studies may include:

- reduced uncertainty about inter-species variability in metabolism, life span, and genetic diversity
- complex temporal patterns of exposure and doses – this relates to situations where risk assessment may be impossible to replicate (e.g. animal studies), whereas some epidemiological studies may be more useful for understanding these complex dose–response relationships
- the ability to assess large numbers of people exposed to low levels of an agent. The doses from exposure to a hazardous agent in epidemiological studies are often considerably less than in toxicological studies. This may have the advantage of providing information about the exposure range of interest although, if they are the result of (prolonged) adult occupational exposures, the exposures are likely to be considerably more than those experienced by people in the general population. With appropriate tools small differences in relative risk in large populations may be able to be assessed (Roseman 1998; Samet et al. 1998).

However, it is more common that epidemiological studies are limited by the amount of data available on dose and tend to address exposure–response relationships (i.e. they are based on whether or not exposure occurred) rather than dose–response relationships. Doses are usually discontinuous and variable in epidemiological studies compared with controlled toxicological experiments. An integrated measure of exposure may need to be developed to represent the non-uniform doses.

Quantitative description of dose–response relationships may be hampered by incomplete information on exposure (especially for biologically relevant time windows), by exposure or dose misclassification, or by the use of

surrogate markers of exposure. Incorrect information about the exposure may bias the description of the exposure–response relationship. If there are wide confidence intervals around the results there can be substantially different policy endpoints depending on whether the upper bound, the lower bound or the midpoint has been chosen for policy making (Samet et al. 1998).

Commonly, too, there is insufficient epidemiological data to discriminate between alternative models that could describe the dose–response relationship. This is particularly important at very low exposure levels, and this is where both epidemiological and toxicological data is often limited. Surrogate measures of outcome (e.g. nerve conduction or tremor) and a relationship between the surrogate measures and health outcomes may need to be established in order to interpret the significance of a study, although care needs to be taken that the surrogate outcomes do relate to clinically meaningful outcomes.

Issues of sample size and whether a threshold or non-threshold can be demonstrated also need to be addressed. If assessing whole populations, the dose–response relationship is going to be different than in a smaller sub-sample. The population sample is likely to include vulnerable groups, while the smaller sample may not. This becomes an issue when standards are chosen on the basis of toxicity data that appears to show a NOAEL based on a small study population and/or a study population that only includes healthy adults or, for example, only adults with mild asthma, when an air pollutant is being assessed.

The reviewer or risk assessor should answer the basic question of whether the epidemiologic data, in an individual study or cumulatively, is adequate for use in dose–response evaluation. There is no formula or quantitative weighting scheme prescribed for making this judgement.

If epidemiologic data adequate for dose–response evaluation is not available, and a risk assessment is being developed for use in making an important regulatory decision, and if it is feasible to develop new epidemiologic data, or to extract new data from existing studies, an effort should be made to develop and provide good epidemiologic dose–response data that can be used together with, or in preference to, high-dose animal data.

The 'London principles' (Federal Focus 1996) may be used to guide the choice of studies in this critical area:

- *Principle 1:* Dose–response assessment should include a range of reasonable dose measures and an explanation why any were rejected, and provide a rationale if any particular dose metric is preferred. In evaluations of both human and animal data, several different measures of dose should be evaluated (if possible).
- *Principle 2:* In the selection of a dose–response model, the greatest weight should be given to models that fit the observed animal and human data and are consistent with the biologically relevant modes of action (genotoxic, non-genotoxic, unclassified). When mechanistic knowledge is uncertain or limited, several plausible dose–response models should be considered and the most plausible ones, based on available data and professional judgement, should generally be used in dose–response evaluation.
- *Principle 3:* When extrapolating cancer risk to exposure levels below the observable range, mechanistic data should be used to characterise the shape of the dose–response function.
- *Principle 4:* When the available epidemiologic data is not adequate to perform dose–response analyses, requiring low-dose estimates of risk to be derived exclusively from animal data, every effort should still be made to use the available human data in assessing the validity of low-dose risk estimates. To the extent feasible, heterogeneity in the human population should be accounted for. Whenever feasible, human data on metabolic biomarkers and other biological measures should be employed to adjust the risk estimates for known differences between species and between high and low doses. If possible, data on susceptibility should be included.
- *Principle 5:* When epidemiologic studies are selected for dose–response assessment, higher-quality studies should be given preference, especially those with precise and accurate exposure information. The availability of information with respect to timing of exposure and response (time/age of first exposure, intensity of exposure, time to tumour), adjustment for confounding variables, and potential interaction with other effect modifiers is particularly important.
- *Principle 6:* A properly conducted meta-analysis, or preferably an analysis based on the raw data in the original studies, may be used in hazard identification and dose–response evaluation when such combination includes an evaluation of individual studies and an assessment of heterogeneity. The combined results ought to provide, more than any single study, precise risk estimates over a wider range of doses. Before using these tools, the gains should be judged sufficient to justify potential errors in inference resulting from combining studies of dissimilar design and quality.
- *Principle 7:* When epidemiological data is used in dose–response assessment, a quantitative sensitivity analysis should be conducted to determine the potential effects on risk estimates of confounders, measurement error, and other sources of uncontrolled bias in study design.
- *Principle 8:* Scientific understanding of differentials in human susceptibility to disease (racial/ethnic/gender/genetic differences, genetic polymorphisms, etc.) should be used to refine low-dose extrapolation procedures when such phenomena are adequately understood.

10.6.3 Exposure assessment

A lack of good exposure data is a common pitfall of environmental epidemiological studies, to the extent that such studies tend only to be as good as their exposure data. The association of particular health effects and specific patterns of exposure, if in keeping with knowledge of pathophysiology, can provide strong support for causal interpretations (WHO 2000a).

Illustrating this problem, Saunders et al. (1997) reviewed the 14 best studies (judged for potential to support causal inference) selected from a shortlist of the 43 analytical studies assessing human health effects in relation to hazardous waste sites that were identified among hundreds that were published. They found that poor exposure measurement was a major factor in the overall lack of convincing evidence of causation from these studies. It is often the case that only a broad indication of the level or nature of exposure may be deduced from epidemiological studies.

Experimental toxicological studies, on the other hand, generally have the advantage of control and accurate measurement of exposures. Nevertheless, at times environmental epidemiological studies may be the only way of determining the distribution of 'real-life' exposures in terms of:

- magnitude
- duration
- temporal patterns
- routes
- size of exposed population
- nature of exposed population.

Future studies should be designed in such a way as to better capture such information.

10.7 CRITICAL EVALUATION OF PUBLISHED RESEARCH

This section is reproduced, with minor adaptation, from *Introduction to research in the health sciences* by Polgar and Thomas (1991), pp. 302–306 by permission of the publisher Churchill Livingstone. Italicised questions are from Riegelman 1981, p. 73 and *British Medical Journal* 1988, p. 50 (with minor amendments).

10.7.1 Critical evaluation of the introduction

The introduction of a paper should essentially reflect the planning of the research. Inadequacies in this section might signal that the research project was erroneously conceived, or poorly planned. The following issues are essential for evaluating this section:

- *Adequacy of the literature review.* The literature review must be sufficiently complete so as to reflect the current state of knowledge in the area. Key papers should not be omitted, particularly when their results could have direct consequences for the research hypotheses or aims. Researchers must be unbiased in presenting evidence that is unfavourable to their points of view.
- *Clearly defined aims or hypotheses.* The aims or hypotheses of an investigation should be clearly and operationally stated. If this is lacking, then how the evidence obtained in the investigation is to be used for conceptual advances in the area will be ambiguous.
- *Selecting an appropriate research strategy.* In formulating the aims of the investigation, the researcher must have taken into account the appropriate research strategy. For instance, if demonstrating causal effects is required, a survey may be inappropriate for satisfying the aims of the research.
- *Selecting appropriate variables.* The operational definition of the variables being investigated calls for selecting appropriate measurement strategies. If the selection of the variables is inappropriate to the constructs being investigated, then the investigation will not produce useful results.

10.7.2 Critical evaluation of the methods section

A well-documented methods section is a necessary condition for understanding, evaluating and perhaps replicating a research project. In general, the critical evaluation of this section will reveal the overall internal and external validity of the investigation.

10.7.2.1 Subjects

The section shows if the sample was representative of the target population and the adequacy of the sampling model used.

- *Sampling model used:* A number of sampling models can be employed to optimise the representativeness of a sample. If the sampling model is inappropriate, then the sample might be biased, raising questions concerning the external validity of the research findings.
- *Sample size:* Use of a small sample is not necessarily a refutation of an investigation, if the sample is representative. However, given a highly variable, heterogeneous population, a small sample will not be adequate to ensure representativeness. Also, a small sample size could decrease the power of the statistical analysis.

- *Description of the sample:* Was there a power-based assessment of adequacy of sample size? A clear description of key sample variables (for example, age, sex, type and severity of condition) should be provided. When necessary and possible, demographic information concerning the population should be provided. Was the population of adequate composition to answer the study questions? If not, the reader cannot judge the representativeness of the sample. Also, the readers might not be able to decide if the findings are applicable to the specific groups of patients being treated.

10.7.2.2 Instruments/apparatus

The validity and reliability of observations and measurements are fundamental characteristics of good research. In this section, the investigator must demonstrate the adequacy of the equipment used for the data collection.

- *Validity and reliability:* The investigator should use standardised apparatus, or establish the validity and reliability of new apparatus used. The lack of proven validity and reliability will raise questions about the adequacy of the empirical findings.
- *Description of instrumentation:* Full description of the structure and use of novel instrumentation should be presented so that the instrument can be replicated by independent parties.

10.7.2.3 Procedures

Full description of how the investigation was carried out is necessary for both replication and for the evaluation of its internal and external validity.

- *Adequacy of the design:* A good design should limit alternative interpretations of the data. A poor design will result in uncontrolled influences by extraneous variables, negating the unequivocal evaluation of causal effects. A variety of threats to internal validity must be considered when critically evaluating an investigation.
- *Control groups:* A specific way of controlling for extraneous effects is the use of control groups. If no control groups are employed, then the internal validity of the investigation might be questioned. Also, if placebo or untreated groups are not present, the size of the effects due to the treatments might be difficult to estimate.
- *Subject assignment:* When using an experimental design, care must be taken when assigning subjects so as to avoid significant initial differences between exposure groups. Even when quasi-experimental or natural comparison strategies are used, care must be taken to establish the equivalence of the groups.
  - Was there a satisfactory statement given of the source of subjects?
  - Was a satisfactory response rate achieved?
  - Was the assignment of people to study and control groups appropriate?
  - Could selection bias have occurred?
  - If the study was experimental, were random and blind assignment maintained?
  - Regardless of the study type, were the study and control groups comparable with respect to characteristics other than the study factors?
- *Exposure parameters:* Was exposure adequately defined? It is important to describe all the exposures experienced by the different groups. If the exposures differ in intensity or in the quality of the administering personnel, then the internal validity of the project is threatened.
- *Rosenthal and Hawthorne effects:* Whenever possible, studies should be double or single blind. If the subjects, experimenters or observers are aware of the aims and predicted outcomes of the investigation, then the validity of the investigation will be threatened through bias and expectancy effects.
- *Settings:* The setting in which a study is carried out has implications for external (ecological) validity. An adequate description of the setting is necessary for evaluating the generalisability of the findings.
- *Times of exposures and observations:* The sequence of exposures and observations must be clearly indicated so that issues such as series effects and confounding can be detected. Identifying variability in treatment and observation times can influence the internal validity of experimental, quasi-experimental or $n = 1$ designs, resulting in, for instance, internal validity problems.

10.7.2.4 Assessment of outcome

- Was the assessment of outcome properly performed in the study and the control groups?
- Was the measure of outcome appropriate to the study aims?
- Was the measure of outcome precise?
- Was the measure of outcome complete?
- Did the process of observation affect the outcome?

10.7.2.5 Critical evaluation of statistical analysis

- Was there a statement adequately describing or referencing all statistical procedures used?
- Were the statistical analyses used appropriate?
- Was the presentation of statistical material satisfactory?
- Did the analysis properly compare the outcomes in the study and the control groups?
- Were the results adjusted to take into account the effect of possible confounding variables? Common confounders are age and sex, regional differences, socioeconomic differentials, smoking, occupation, ethnic differences.
- Was a significance test properly performed to assess the probability that the difference was due to chance?
- Was a proper measure of the size of the difference presented?
- Was a proper measure of the degree of overlap of the differences presented?
- Were the confidence intervals given for the main results?

Motulsky (1995) provides a useful checklist of common pitfalls to bear in mind when reading research papers that include statistical analysis, which has been adapted as follows:

- *Look at the data:* Summary statistics may result in the loss of useful information.
- *Beware of very large and very small samples:* large samples may generate statistically significant but unimportant findings small samples have little power to detect important differences.
- *Beware of multiple comparisons:* When analysing random data, on average 1 out of 20 comparisons will be statistically significant ($p < 0.05$) by chance.
- *Don't focus on averages alone:* Variability may reflect real biological diversity, and outliers may be more important.
- *'Garbage in, garbage out':* Statistical tests do not tell whether the study was conducted properly.
- *Confidence limits:* Are as informative as p values (and may be more so, particularly when dealing with hazards).
- *Statistical significance:* Does not necessarily indicate biological importance.
- $p < 0.05$ is not sacred: It is an arbitrary cut-off value.
- *Correlation or association:* Does not imply causation.

10.7.2.6 Critical evaluation of the results

The ways in which epidemiological data are properly presented and analysed goes beyond the scope of this enHealth document in terms of complexity and depth, and reference should be made to standard texts. However, the following general points can be made.

- *The results:* Results should represent a statistically correct summary and analysis of the data. Inadequacies in this section could indicate that inferences drawn by the investigator were erroneous.
- *Tables and graphs:* Data should be correctly tabulated or drawn and adequately labelled for interpretation.
- *Complete summaries:* Complete summaries of all the relevant findings should be presented.
- *Selection of statistics.* Both descriptive and inferential statistics must be selected appropriately. Selecting inappropriate statistics could distort the findings and lead to inappropriate inferences.
- *Calculation of statistics:* Both descriptive and inferential statistics must be correctly calculated. Using computers generally ensures this, although some attention must be paid to gross errors when evaluating the data.

10.7.2.7 Critically evaluating the discussion

In the discussion, the investigator draws inferences from the data in relation to the initial aims or hypotheses of the investigation. Unless the inferences are correctly made, the conclusion drawn might lead to useless or dangerous treatments being offered to clients.

- *Drawing correct inferences from the data:* The inferences from the data must take account of the limitations of descriptive and inferential statistics. Correlations do not necessarily imply causation, or that a lack of significance in the analysis could imply a Type 2 error (see below).
- *Logically correct interpretations of the findings:* Interpretation of the findings must follow from the statistical inferences, without extraneous evidence being introduced. For instance, if the investigation used an $n = 1$ design, the conclusions should not claim that a procedure is generally useful.
- *Protocol deviations:* In interpreting the data, the investigator must indicate and take into account unexpected deviations from the intended design. For instance, a placebo/active treatment code might be broken, or 'contamination' between control and experimental groups might be discovered. If such deviations are discovered by investigators, they are obliged to report these, so that the implications on the results might be taken into account.
- *Generalisation from the findings:* Strictly speaking, the data obtained from a given sample are generalisable only to the population from which the sample was drawn. This point is sometimes ignored by investigators, and the findings are generalised to subjects or situations that were not considered in the original sampling.
- *Statistical and practical significance:* Statistical significance does not necessarily imply that the results of an investigation are applicable in practical terms. In deciding on practical significance, factors such as the size of effect, side effects, cost-effectiveness and value judgements concerning outcome must be considered.
- *Theoretical significance:* It is necessary to relate the results of an investigation to previous related

findings, as identified in the literature review. Unless the results are logically related to the literature, the theoretical significance of the investigation remains unclear.

The processes involved in comparing the findings of a set or related papers are introduced in the next sub-section.

- Was a valid interpretation drawn from the comparisons made between the study and control groups during analysis?
- Did the investigators properly reject or fail to reject the null hypothesis?
- Did the investigators consider the possibility of Type 1 and Type 2 errors in interpreting the meaning of the significance test? (Type 1 errors are the result of chance and are the rejection of the null hypothesis when no true difference exists in the larger population. Type 2 errors result from chance or too small a sample size and are the failure to reject the null hypothesis when a true difference exists in the larger population.)
- Were the size of the differences and the degree of overlap taken into consideration in the conclusions reached about the meaning of observed differences?
- In interpreting the meaning of any relationship, was the concept of cause and effect (causation) properly applied?
- Were the extrapolations to individuals not included in the study properly performed?
- Did the investigators stay within the limits of the data when extrapolating the results?
- If the investigators extrapolated from population data to individual data, was this appropriate and correct?
- Did the researchers take into consideration differences between the study population and the population to which they extrapolated their data?

10.7.3 Evaluation of meta-analyses

Meta-analysis is the process of undertaking a quantitative review of the literature, seeking consistent patterns among, and sources of discrepancies between, studies. An assessment should consider the homogeneity of the studies examined and whether summary effects estimates will be calculated and by what methods (WHO 2000a).

WHO (2000a) has recommended that the following features be considered when conducting, or assessing the findings of, a meta-analysis:

- establishing or noting a protocol specifying the objectives of the review and the methods to be employed
- having inclusion criteria that are more inclusive than exclusive, so enabling sensitivity analysis using different levels of inclusion to be undertaken
- avoiding a single quality score of studies and presenting, instead, an assessment of a range of characteristics
- weighting according to the precision of the study
- assessing and addressing the impact of publication bias
- systematically quantifying the heterogeneity of the studies that can enable the identification of sources of variability in the results of studies from factors such as the choice of methodology and the inclusion of susceptible sub-groups or unusual exposure conditions
- using sensitivity analyses of factors such as different analytic approaches, different methods of extracting results from the studies or the inclusion or exclusion of particular studies or types of studies
- appraising methods used to obtain qualitative and quantitative summary estimates from a collection of studies.

10.7.4 Common omissions and errors in relevant published research

Rushton (2000) provides a report on some of the most serious omissions and errors in papers presented in recent years to the journal, *Occupational and Environmental Medicine*. These are:

Design

- Authors unclear about type of epidemiological study
- Adequacy of sample size not considered
- Bias in selection of subjects execution
- Data collection problems and missing data not adequately reported
- Non-respondents not investigated
- Sample selection and exclusions inadequately justified

Analysis

- Parametric tests carried out on obviously skewed data
- Use of multiple paired tests
- Inappropriate analysis of repeated measures or longitudinal data
- Incorrect analysis of matched case-control studies
- Modelling incorrect (e.g. inadequate adjustment for confounders, interaction terms not included, only significant variables from preliminary analyses included)

Presentation

- Inadequate description of the methodology and statistical procedures
- Inappropriate summary statistics for non-normal data
- No presentation of risk estimates (e.g. odds ratios – and confidence intervals)

Interpretation

- Potential bias due to sample selection, no or poor response, missing values, exclusions
- Lack of statistical power not considered
  - No allowance made for multiple testing
  - Misunderstanding and misinterpretation of results from models

Table 18 summarises some potential problems that might emerge in the context evaluation of an investigation. A point that must be kept in mind is that even where an investigation is flawed, some useful knowledge might be drawn from it. The aim of critical analysis is not to discredit or tear down published work, but to ensure the reader understands its implications and limitations.

10.8 UNDERTAKING HEALTH STUDIES

The material in the following sections is adapted from ATSDR (1996).

In some situations, there will be a need to undertake health studies as part of a risk assessment. A risk assessment may have been prompted by health studies undertaken by the community. The design of health studies should be underpinned by epidemiological principles. A range of factors need to be considered before embarking on a health study.

10.8.1 Public health significance

Public health significance is a key factor in considering the merits of a proposed health study. Issues for consideration include: the toxicity of the agent; the pathways of human exposure; severity and biological plausibility of the health outcome; need for new information (beyond what is already known or what has already been done); size and susceptibility of the population affected; ability to prevent or mitigate exposure or health outcomes; and relevance to other situations with similar agents and exposure pathways.

Table 18: Checklist for evaluating published research

Problem that might be identified in a research article	Possible implications
1. Inadequate literature review	Misrepresentation of the conceptual basis for the research
2. Vague aims or hypotheses	Research might lack direction; interpretation of evidence might be ambiguous
3. Inappropriate research strategy	Findings might not be relevant to the problem being investigated
4. Inappropriate sampling method	Measurements might not be related to concepts being investigated
5. Inadequate sampling method	Sample might be biased; investigation could lack external validity
6. Inadequate sample size	Sample might be biased; statistical analysis might lack power
7. Inadequate description of sample	Application of findings to specific groups or individuals might be difficult
8. Instruments lack validity or reliability	Findings might represent measurement errors
9. Inadequate design	Investigation might lack internal validity (i.e. outcomes might be due to uncontrolled extraneous variables)
10. Lack of adequate control groups	Investigation might lack internal validity; size of the effect difficult to estimate
11. Biased subject assignment	Investigation might lack internal validity
12. Variations or lack of control of treatment parameters	Investigation might lack internal validity
13. Observer bias not controlled (Rosenthal effects)	Investigation might lack internal and external validity
14. Subject expectations not controlled (Hawthorne effects)	Investigation might lack internal and external validity
15. Research carried out in inappropriate setting	Investigation might lack ecological validity
16. Confounding of times at which observations and interventions are carried out	Possible series effects; investigation might lack internal validity
17. Inadequate presentation of descriptive statistics	The nature of the empirical findings might not be comprehensible
18. Inappropriate statistics used to describe and/or analyse the data	Distortion of data; false inferences may be drawn
19. Erroneous calculation of statistics	False inferences may be drawn
20. Drawing incorrect inferences from the data	False conclusions might be made concerning the outcome of an investigation
21. Protocol deviations	Investigation might lack internal and external validity
22. Over-generalisation of finding	External validity might be threatened
23. Confusing statistical and clinical significance	Treatments lacking clinical usefulness might be encouraged
24. Findings not logically related to previous research findings	Theoretical significance of the investigation remains doubtful

Adapted from: Polgar & Thomas 1991.

10.8.2 Community perspective and involvement

Community involvement is critical to the success of any proposed health study. Various community involvement methods can be used for health studies. Issues for consideration include: an ability to involve key community stakeholders; an understanding of community health concerns; an understanding of the approach and limitations of proposed activities; and community support for the study being conducted. It must be recognised that interested community members tend to be a more homogeneous group than a sample drawn from the general population. Significant effort must be made to identify strategies to alert individuals and interested groups of the discussion paper, and to invite their considered contribution. Any input to process is most likely to be performed in personal time, that is, out of work hours. It may also require discussions with group colleagues during internal processes that have scheduled meetings. Many dissemination processes for community groups (e.g. newsletters, also require long lead times to reach their constituent members). Community consultation time lines must therefore factor in these delays to ensure effective community engagement. Advance warning of document release can help to facilitate processes. However, a key factor to community engagement remains allowing sufficient time for the dissemination, full review, drafting and reviewing of responses.

10.8.3 Scientific importance

Scientific importance is closely related to public health significance, but they do not form a perfect match. Issues for consideration include: the ability to provide new knowledge or information about an exposure–outcome relationship; to address specific exposures or outcomes that have not been adequately studied; to allow new laboratory tests or study methods to be used or evaluated; to generalise to other situations or populations; and provide confirmation or additional support to a preliminary hypothesis or theory. The principal tenet for public health is promoting health of the population, for all sectors, and with special attention to protecting the vulnerable. The precautionary principle is now well known among the public health sector and community advocates. A tension arises when advances that promise benefits to society also bring risks (real or perceived), and this can become especially acute when the full weight of evidence cannot be known for many years. The once unquestioned faith in science has transformed to an educated and cautious scepticism among large sectors of the community.

10.8.4 Ability to provide definitive results

Since health studies may and frequently will end up with inconclusive findings, it is important to consider how definitive the study might be in providing scientifically useful results related to specific exposure–outcome relationships. Issues for consideration include the ability to: obtain appropriate measurements of exposure and to document health outcomes and exposures; use adequate control or comparison populations; obtain community support to ensure an adequate participation rate; state clearly the study objectives and specific hypothesis to be tested; have sufficient statistical power to detect predicted effects if they exist, obtain data on important potential confounders, and evaluate a dose–response relationship or gradients of exposure. Extrapolation of findings to vulnerable groups, such as children remains problematic, as involvement in these groups in studies is fraught with complexities.

10.8.5 Resources

Resources are critical to the support, conduct and completion of any proposed health study. Issues for consideration include: the availability of qualified personnel and technical support; an ability to obtain necessary data and health information; and the availability of appropriate project time lines and resources.

10.8.6 Authority and support

It is critically important that local, state, territory and Commonwealth health agencies be involved early in discussions about potential health studies. Issues for consideration include: the ability to support or provide technical assistance requested by the local, state or territory health agency; the ability of local and state or territory health agencies to address the community problem and health concerns; and the involvement of appropriate agencies with legislative and regulatory backing.

10.8.7 Nature of the health study

When the decision to conduct a health study is being considered, several criteria are used to determine the type of health study. These relate to whether the relevant research hypothesis requires:

- the characterisation of environmental contaminants by type, media, and concentration levels
- documented evidence of human exposure at a level of concern
- level of current knowledge about the relationship between exposure and specific adverse health outcomes and/or
- documented excess of an adverse health outcome, when known.

The health studies can be grouped into Type 1 and Type 2 studies.

10.8.7.1 Type 1 health studies

Type 1 health studies explore or generate hypotheses about exposure–outcome associations and address specific exposures, community health concerns, or specific information needs. Examples of Type 1 health studies follow.

- Cross-sectional studies. These are surveys of a sample of individuals to obtain information about current and past health or environmental exposures, or both. These studies can include comparison populations with demographics similar to those of the exposed (target) population.
- Pilot investigations collect additional information to assess the feasibility and value of conducting a full-scale health study. The investigation might include assessments of data completeness and quality, the level of documentation of exposures or health outcomes methods to identify and track individuals, study size, and statistical power issues and the availability of a control population or comparison.
- Cluster investigations evaluate when the reported occurrence of a specific disease or condition is above the expected number for a given geographic location and time period. These investigations can be conducted to confirm case reports, determine an unusual disease occurrence and explore potential risk factors.
- Comprehensive case reviews are medical or epidemiological evaluations of the medical status of one or more individuals through medical record reviews, interviews or biomedical testing to determine additional information about their health status or potential for exposure.
- Situation-specific surveillance is designed to assess the specific occurrence of one or more defined health conditions among a specific population potentially exposed to hazardous agents in the environment. Data collection might include using existing records of health events or records from relevant health care providers.
- Health statistics reviews use available health and demographic information to assess the occurrence of specific health effects in defined geographic areas and determine if the rates are elevated compared to similar populations elsewhere. Available information might include: death certificates, birth certificates, and census data; tumour or disease registry data; and health surveillance or disease notification data. A health statistics review may be performed in response to a reported cluster of specific diseases or conditions.
- Exposure investigations use environmental or biological testing, or both, for the hazardous agent(s) of interest. The biological test might measure the level of the hazardous agent, a metabolite of a hazardous substance, or another marker of exposure in human body fluids or tissues. The purpose of this investigation is to assess individual exposure levels to a specific agent associated with the situation. The levels identified should be compared with that of a relevant reference group or with a known standard reference level. Depending on the hazardous agent, the investigation can be used to explore for evidence of past or current exposure.
- Disease and symptom prevalence surveys are used to measure and compare the occurrence of self-reported diseases, in some instances using medical records or physical examinations to validate adverse health conditions. Addressing potential health concerns raised by the community, the survey compares an exposed population (target area) with an unexposed population (control area) with similar demographic characteristics. The purpose is to determine the need for further health studies in the target area, provided there are statistically significant excesses that are clinically important. Depending on the contaminants and circumstances, biological testing of exposure or effect, or both, might also be collected as part of the survey.

When a Type 1 health study is considered appropriate, there are several attributes that are considered necessary in order to ensure the quality of the study effort:

- a reasonable ability to document and characterise exposure in the target area
- an adequate study size for the type of study recommended
- an ability to identify and locate subjects and records
- appropriate comparisons for rates of occurrence or levels of exposure
- an ability to control confounding factors and biases (when possible).

10.8.7.2 Type 2 health studies

Type 2 health studies are specifically designed to test scientific hypotheses about the associations between adverse health outcomes and exposure to hazardous substances in the environment. The following are some examples of Type 2 health studies:

- Case-control studies are designed to collect information and compare differences in exposures and other risk factors in two groups of people: those with specific illnesses or conditions (cases) and those without the illnesses or conditions (controls). The controls are selected to represent the population from which the cases were identified. Usually the cases and controls are identified first, and then information is collected about past exposures and other risk factors.
- Cohort studies are designed to collect information from a group of people followed over a period of time, and information on the occurrence of specific illnesses or conditions is collected. Cohort studies can be prospective, meaning that individuals involved in the study are followed into the future, or cohorts can be retrospective, meaning that the cohort

is reconstructed from historical records and then followed over a specified time period. They are expensive, and require long periods of time, and large numbers of people must be followed for rarer outcomes to provide enough cases for analysis.

- Nested case-control studies are another approach that uses both of the study designs previously mentioned. The nested case-control study uses cohort individuals who have developed a specific illness or condition (case) and persons sampled from the cohort who have not developed the illness or condition (control). The case-control method is then used to collect additional information and analyse the differences between these two groups.

There are several attributes of Type 2 health studies that are considered necessary in order to ensure valid scientific findings, including:

- an ability to reasonably estimate or document individual exposures
- an ability to document or validate human health outcomes
- an adequate study size and statistical power
- an ability to identify and locate subjects and records
- availability of an appropriate control or comparison population
- an ability to control confounding factors and minimise biases
- an ability to determine influence of environmental, behavioural or other factors.

10.8.8 Ensuring the quality of a health study

To ensure a useful and appropriate outcome the following factors should be met:

- The group conducting the health study must be capable and fully responsible for conducting the health study.
- Personnel conducting the health study must be identified and have appropriate training and experience.
- The facilities and resources must be appropriate for successfully completing the health study.
- Contractors for a health study must follow written and approved work plans and their work must be carefully reviewed by the sponsoring group.
- For complex studies, a detailed study protocol should be written and undergo scientific peer review.
- Ethical issues relating to the protection of human subjects, consent and data confidentiality procedures must be addressed.
- Reports of complex health studies may need to undergo scientific peer review prior to any public release of information.
- Community involvement and knowledge of the health study are necessary: the involvement process will assist in ensuring that the community understands and supports the study focus and design, and its limitations.
- Depending on the community involvement approach, public meetings might be held to present and discuss the study methods and findings. However, final study methods must be scientifically valid before proceeding.
- All study reports, data files and related documentation should be kept in the official records.
- Any environmental sampling or biological testing must follow existing standards for collection, handling, chain of custody, storage, analysis, and reporting by an approved laboratory.

All standard quality control and quality assurance procedures must be followed and documented.

10.8.9 Contents of a health study protocol

The following components should be considered in drafting a report. Ethics approval should be outlined. Protocols for health studies might not need to contain all of the items within this outline.

The listing is more comprehensive in order to cover the wide variety of study approaches.

1. Title and identification page
2. Introduction and overview
3. Background
   - Situation description
   - Demographics
   - Contaminants and pathways
   - Community health concerns
   - Literature review
4. Purpose
5. Study objectives
6. Methods
   - Rationale for study design
   - Study description
   - Eligibility criteria
   - Selection of target area and population
   - Selection of comparison area and population
   - Sample size and statistical power estimates
   - Participant selection and definitions
   - Subject recruitment procedures
   - Locations of data and specimen collection
   - Informed consent procedure
   - Questionnaire procedures
   - Interviewer training and methods
   - Methods for measurement of exposure
   - Collection of biological specimens
   - Additional data collection or sources
   - Chain of custody and shipping
   - Laboratory methods and quality control
   - Privacy protection
   - Findings of immediate significance
   - Follow-up of abnormal lab results
   - Data analysis
   - Data entry, editing and management
   - Data transformation
7. Data analysis plan and methods
8. Study time lines
   - Key activities or milestones
9. Community involvement and notification
10. Interpretation of results
11. Limitations of the study
12. References
13. Tables and figures
14. Attachments
    - Data collection forms and questionnaire
    - Study letters of notifications and consent form
15. Specimen collection and shipping protocols

Chapter 11: Assessment of carcinogens

Cancer is an especially dreaded group of diseases. It may affect different tissues and cell types and it is multifactorial in its development. An essential feature of cancer is that it represents a process where a single cell or groups of cells lose the mechanisms that control differentiation and growth. Cancer cells may not only exhibit uncontrolled growth, but when a malignant tumour occurs, it can invade adjacent or distant tissues (metastasis).

Existing methodologies have difficulties in assessing and conveying the broad range of human health implications of exposure to environmental carcinogens. This, combined with a high 'dread factor' for cancer, has resulted in many cases in a disproportionate regulatory, political and public focus on cancer as compared with other types of adverse health outcomes associated with exposure to environmental chemicals.

Human cancer incidence generally increases with age, as a result a lifetime of exposure to various cancer risk factors, including endogenous processes independent of external environmental exposures. The interpretation of possible changes in cancer incidence in populations requires that crude cancer prevalence or incidence rates are age-adjusted to account for the propensity for cancer incidence to increase as people grow older.

Epidemiological studies have identified a variety of lifestyle, dietary and occupational factors that can appear to increase cancer risk. A history of smoking is one that has been well proven to increase the incidence of a number of types of cancer, although lung cancer is the site that is best known. The relationship between mesothelioma, lung cancer and occupational exposure to asbestos, and to some extent non-occupational exposures, is another well-known example.

However, it is rare that a single causative factor can be identified for cancer in an individual. The best that can usually be done is to identify risk factors that may contribute to an increased overall cancer rate by contributing to the initiation or promotion of the carcinogenic process. This may be via a genetic transformation of a single progenitor cell, with subsequent clonal expansion and transformation through the various stages of its progression to an established cancer.

Some commentators have put forward a view that the majority of human cancers have a purely environmental factor associated with their causation, and some appear to be still convinced that environmental factors are at the heart of the human cancer epidemic (Belpomme et al. 2007). However, determining the contribution of environmental factors on the cancer burden in humans (as opposed to dietary and lifestyle factors) is usually difficult because it is rare that the incidence of cancer in specific populations or groups can be related to exposure to any single or multiple environmental factors. This does not mean that possible impacts of environmental chemicals might be sufficiently small that they can be ignored in risk assessment. In fact, the EHRA process for carcinogens errs on the side of caution and it is quite conservative. With few exceptions, carcinogenic risk assessment is still rooted to the concept developed some 30–50 years ago that there is no threshold to a carcinogenic response because a single DNA-damaging event (usually a mutational event that survives DNA repair) could initiate a carcinogenic transformation in a single cell.

Since epidemiology provides relatively limited quantitative information on carcinogenesis associated with environmental chemicals, animal studies using controlled but relatively high levels of exposure remain the mainstay of EHRA activities where cancer is the toxicological endpoint of concern. Such tests require exposure of test animals for the majority of their life span (18–30 months in rodents). The doses used are as high as possible (up to the conceptual maximum tolerated dose – MTD) in order to maximise the sensitivity of the test to detect a carcinogenic response. As with epidemiological studies, it is important that the incidence of cancer in the rodent studies consider age- and species-specific effects, as well as inter-current mortality. In some species and strains, the incidence of tumours in untreated (control) animals can increase towards 100 per cent at some sites as the animals age, making discrimination of a treatment-related effect more difficult. There is further discussion of the broader range of issues (such as dose and species selection) that need to be taken into consideration in evaluating rodent-based carcinogenesis studies in Appendix 1.

The two issues of cancer risk assessment addressed in this chapter are:

- methods for assessing whether or not a chemical should be considered to be a human carcinogen for EHRA risk assessment purposes
- the basis for selecting either a threshold or non-threshold model for dose–response extrapolation and estimation of risk.

Conventionally, the EHRA of potential carcinogens can differ substantially from that used for other toxicological endpoints. EHRA approaches that assume a threshold (essentially all toxicological endpoints other than cancer) estimate a level of exposure (e.g. ADI or TDI) that is assumed to be without appreciable risk. No finite risk level is calculated for any dose or exposure when the ADI/TDI threshold approach is used. Consequently, there is no need to establish an 'acceptable' or target risk level for the protection of human health.

For most carcinogens, no threshold is assumed to exist for the carcinogenic response, so the EHRA approach is to extrapolate the dose–response relationship from a point of departure (POD) on the dose–response curve towards zero and estimate a finite risk level at specified doses or levels of exposure. Non-threshold methodology is applied most rigorously to chemicals that have the potential to damage DNA (genotoxic chemicals), although it may be applied also to putatively non-genotoxic chemicals, where some doubt remains about the carcinogenic mode of action (MoA). For non-genotoxic chemicals, where evidence provides more confidence about the putative MoA, there may be adequate scientific grounds to adopt a threshold approach to EHRA.

A distinction is sometimes made between the terms 'mode of action' and 'mechanism of action'. The latter implies a more detailed understanding and description of events, often at the molecular level, than is meant by mode of action (US EPA 2005a).

The difference between threshold and non-threshold approaches to quantitative EHRA is outlined in more detail in Sections 3.10, 3.11, 5.4 and 11.6.

11.1 METHODS FOR THE HAZARD IDENTIFICATION OF CARCINOGENS

11.1.1 Use of animal toxicity studies and epidemiology

As noted above, epidemiology can sometimes provide convincing evidence that a specified chemical or a cluster of environmental exposures may have a significant role in human cancer. However, in the most part, EHRA for carcinogens relies on animal studies to identify potential human carcinogens and to provide the quantitative dose–response data needed for quantitative risk assessment.

The US National Toxicology Program (NTP), a cooperative program of the National Cancer Institute and the National Institutes of Environmental Health, has had an extensive chemicals testing program for over 30 years aimed at providing standardised data on rodent carcinogenesis (Bucher & Portier 2004). The program is extensively used as a database for both qualitative and quantitative carcinogenic risk assessment. The NTP 2-year rodent bioassay program is extensively used as a basis for carcinogenic risk assessment. However, the NTP also recognises that the future lies in better understanding of mechanism-based biological observations, which can be integrated with better predictive testing regimens.

11.2 GUIDANCE ON ASSESSING CARCINOGENS

Carcinogen evaluation is subject to quite extensive (possibly even exhaustive) stakeholder consultation to use the best available science and to be consistent with other EHRA approaches. At the same time, it strives to maintain a highly conservative risk management approach for such a dreaded disease. US guidance for the assessment of carcinogenicity has been evolving through such consultative processes, although at times the evolution has been slow. The most recent US EPA guidance on carcinogenic risk assessment was issued in 2005 (US EPA 2005a) following a decade of progression through its developmental and consultative stages.

The International Agency for Research on Cancer (IARC) has traditionally held an important place in framing policies and practice in carcinogenic risk assessment. The IARC monograph series (IARC monographs on carcinogenic risk) has been an important tool for classifying carcinogens (see Section 11.3.1), but the name of the monograph series is something of a misnomer, since its assessments essentially categorise carcinogenic hazards, and do not attempt to quantify risk. This is recognised in the preamble to the IARC monographs, with statements that IARC considers that a cancer 'hazard' is an agent that is capable of causing cancer under some circumstances, and a cancer 'risk' is a probability estimate of a carcinogenic effect occurring from a defined amount, frequency and duration of exposure to a carcinogenic agent.

The International Program on Chemical Safety (IPCS) has been developing a special project on carcinogenic risk assessment as part of its broader program on harmonisation of risk assessment methodologies (see <http://www.who.int/ipcs/methods/harmonization/en/index.html>). A key feature of this program is guidance on how to assess the human relevance of carcinogenic responses in animal studies (see Sections 11.4 and 11.5).

Australian authorities have not issued any formal guidance on carcinogenic risk assessment, although most state or territory authorities adopt the US EPA linearised non-threshold approach in framing local standards and evaluating submitted EHRAs. This often includes adoption of a 10^{-5} target risk level to guide clean-up or risk management activities, although 10^{-6} is still the target risk level adopted in setting water quality guidelines.

In this update of enHealth guidance on EHRA, a case is made for 10^{-5} to become the target risk level for the non-threshold approach to risk assessment of carcinogens in most circumstances (see Section 5.10).

11.3 CLASSIFICATION OF CARCINOGENS

Advances in biological knowledge are enabling mechanistic data, pharmacokinetic data and other relevant data to be increasingly taken into account in classifying and assessing the risks of carcinogens.

Australia does not have a formal mechanism for classifying carcinogens, but most Commonwealth, state and territory jurisdictions give weight to IARC and US EPA classifications when considering risk assessment based on a carcinogenic response. The difficulty associated with reaching harmonisation of various national and international classifications schemes for carcinogens was explored by Di Marco et al. (1998).

11.3.1 IARC approach

The IARC developed the first system for qualitatively categorising chemical carcinogens (IARC 1978). Initially, the approach was to adopt a strength-of-evidence scheme to decide whether, for humans and experimental animals separately, there was sufficient or limited evidence of carcinogenicity for a substance, mixture or exposure circumstance, or whether the database was inadequate for classification (prior IARC monographs essentially only summarised existing tumorigenicity studies). Since then, the scheme has evolved whereby now all data, including human, animal and *in vitro* studies are assessed for an overall weight-of-evidence (WoE) evaluation of human carcinogenicity (Vainio & Wilbourn 1992).

Thus considerable weight is given to the animal cancer bioassays, though some scientists are not convinced of the validity of this philosophy.

Ward (2007) presented a damning criticism of the reliance on animal studies to identify carcinogens. While he acknowledges that regulatory authorities are sometimes prepared to accept that carcinogenic responses associated with high-dose studies in animals may not be relevant for human risk assessment, especially when there is evidence for a species-specific MoA, the science underlying these hypothetical mechanisms of carcinogenesis in rodents and humans remains unproven. Even the concept that genotoxicity is the key initiating step in carcinogenesis has been challenged, with observations that epigenetic events and mechanisms are also important (Trosko & Upham 2005).

Whether or not the introduction of mice genetically engineered to increase susceptibility to some classes of carcinogens will enhance the predictive science of animal studies remains to be seen (MacDonald et al. 2004). Similarly, the combination of toxicogenomic approaches to understanding carcinogenic mechanisms of action may also hold some promise for future advances in the science (Guyton et al. 2009). This may be further enhanced by combinations of classical *in vitro* genotoxicity assays with *in silico* screening techniques for genotoxicity and computational techniques for SAR analysis (Benfenati et al. 2009).

A later decision by IARC was to incorporate information on the mechanism of action of chemicals in the evaluation process (Vainio & Wilbourn 1992). In practical terms, this means that category Group 1 (sufficient evidence for carcinogenicity in humans) 'could be extended to include agents for which the evidence of carcinogenicity in humans is less than sufficient but for which there is sufficient evidence of carcinogenicity in experimental animals and strong evidence in exposed humans that the agent acts through a relevant mechanism of carcinogenesis' (Vainio & Wilbourn 1992). This aspect of the evaluation process is becoming increasingly important as understanding of mechanistic pathways improves great advances are being made, especially with the advent of sophisticated laboratory molecular techniques.

Essentially four descriptive dimensions of mechanistic data are used in the IARC process (Fitzgerald 1993):

- evidence of genotoxicity (i.e. structural change at the level of the gene)
- evidence of effects on the expression of relevant genes (i.e. functional changes at the intracellular level)
- evidence of relevant effects on cell behaviour
- evidence of time and dose relationships of carcinogenic effects and interactions between agents.

In each monograph, the 'preamble' sets out the rules and approaches used by IARC expert panels to classify substances or exposures as 'carcinogenic to humans'.

The term 'carcinogen' is used in the IARC monographs to denote 'an exposure that is capable of increasing the incidence of malignant neoplasms. The induction of benign neoplasms may in some circumstances contribute to the judgement that the exposure is carcinogenic. The terms 'neoplasm' and 'tumour' are used interchangeably.

The IARC evaluations consider whether the agent may intervene at one of more of the multiple stages of carcinogenesis or mechanisms. However, experimental evidence still maintains a high status in the assessment hierarchy. The statement in current IARC monograph preambles notes that:

> ... in the absence of adequate data on humans, it is biologically plausible and prudent to regard agents and mixtures for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to humans. The possibility that a given agent may cause cancer through a species-specific mechanism which does not operate in humans should also be taken into consideration.

The preamble goes on to define the terms 'sufficient evidence', 'limited evidence' and 'inadequate evidence'. Understanding how these terms are used in the final evaluation is critical, as is the understanding that the evaluations:

- consider only published information
- consider the route of exposure in animal studies, but not necessarily its relevance to the conditions under which human exposure may occur
- do not evaluate carcinogenic potency.

The use of these terms in reaching a final classification is embodied in the rules, as outlined in the IARC monograph preambles:

> *Group 1 – The agent (mixture) is carcinogenic to humans. The exposure circumstance entails exposures that are carcinogenic to humans.*
>
> This category is used when there is *sufficient evidence* of carcinogenicity in humans. Exceptionally, an agent (mixture) may be placed in this category when evidence of carcinogenicity in humans is less than sufficient but there is *sufficient evidence* of carcinogenicity in experimental animals and strong evidence in exposed humans that the agent (mixture) acts through a relevant mechanism of carcinogenicity (e.g. dioxins).
>
> *Group 2A – The agent (mixture) is probably carcinogenic to humans. The exposure circumstance entails exposures that are probably carcinogenic to humans.*
>
> This category is used when there is *limited evidence* of carcinogenicity in humans and *sufficient evidence* of carcinogenicity in experimental animals. In some cases, an agent (mixture) may be classified in this category when there is *inadequate evidence* of carcinogenicity in humans, *sufficient evidence* of carcinogenicity in experimental animals and strong evidence that the carcinogenesis is mediated by a mechanism that also operates in humans. Exceptionally, an agent, mixture or exposure circumstance may be classified in this category solely on the basis of *limited evidence* of carcinogenicity in humans.
>
> *Group 2B – The agent (mixture) is possibly carcinogenic to humans. The exposure circumstance entails exposures that are possibly carcinogenic to humans.*
>
> This category is used for agents, mixtures and exposure circumstances for which there is *limited evidence* of carcinogenicity in humans and less than *sufficient evidence* of carcinogenicity in experimental animals. It may also be used when there is *inadequate evidence* of carcinogenicity in humans but there is *sufficient evidence* of carcinogenicity in experimental animals. In some instances, an agent, mixture or exposure circumstance for which there is *inadequate evidence* of carcinogenicity in humans but *limited evidence* of carcinogenicity in experimental animals together with supporting evidence from other relevant data may be placed in this group.
>
> *Group 3 – The agent (mixture or exposure circumstance) is not classifiable as to its carcinogenicity to humans.*
>
> This category is used most commonly for agents, mixtures and exposure circumstances for which the *evidence of carcinogenicity* is *inadequate* in humans and *inadequate* or *limited* in experimental animals. Exceptionally, agents (mixtures) for which the *evidence of carcinogenicity* is *inadequate* in humans but *sufficient* in experimental animals may be placed in this category when there is strong evidence that the mechanism of carcinogenicity in experimental animals does not operate in humans. Agents, mixtures and exposure circumstances that do not fall into any other group are also placed in this group.
>
> *Group 4 – The agent (mixture) is probably not carcinogenic to humans.*
>
> This category is used for agents or mixtures for which there is *evidence suggesting lack of carcinogenicity* in humans and in experimental animals. In some instances, agents or mixtures for which there is *inadequate evidence* of carcinogenicity in humans but *evidence suggesting lack of carcinogenicity* in experimental animals, consistently and strongly supported by a broad range of other relevant data, may be classified in this group.

Between 1972 and 2008, the IARC monograph series (Volumes 1–100) had evaluated 935 environmental agents and exposures. The breakdown of classifications was:

Group 1 (carcinogenic to humans)	108
Group 2A (probably carcinogenic to humans)	63
Group 2B (possibly carcinogenic to humans)	248
Group 3 (not classifiable as to carcinogenicity)	515
Group 4 (probably not carcinogenic to humans)	1

The fact that only one substance has been classified in Group 4 (caprolactam in 1998) is probably because entry into the IARC program for evaluation is a selective process. There must be at least some plausible published evidence suggestive

of carcinogenicity for a substance to be selected for evaluation. There are a number of substances used as human therapeutic agents that have been classified as Group 1 (not just a range of cytotoxic anti-cancer drugs). This exemplifies the importance attached to human epidemiological data, and the relative lack of importance given to the conditions under which humans may be exposed.

11.3.2 US EPA classification of carcinogens

Current US guidance on carcinogen assessment (US EPA 2005a) also emphasises the importance of integrating and weighing all the available evidence in reaching a decision about classifying a chemical as carcinogenic to humans. This should include a consideration of the conditions under which a carcinogenic response might occur in humans, as well as its likelihood.

The US EPA policy position (US EPA 2005a pp. 1–10) is that:

> ... in the absence of sufficiently scientifically justifiable mode of action information, the EPA generally take public health-protective, default positions regarding the interpretation of toxicologic and epidemiologic data: animal tumour findings are judged to be relevant to humans and cancer risks are assumed to conform with low dose linearity.

However, in a departure from previous guidance on carcinogen classification, it recommends that a descriptive WoE narrative replace the numerical classification scheme. Such a narrative should include a summary of the information about the mode of action.

The narrative should be written around five standardised descriptors:

- carcinogenic to humans
- likely to be carcinogenic to humans
- suggestive evidence of carcinogenic potential
- inadequate information to assess carcinogenic potential
- not likely to be carcinogenic to humans.

The narrative would set out, within each of these descriptors, the conditions under which a carcinogenic response is likely to occur, and it suggests that more than one conclusion may be possible for an individual agent, for example, when carcinogenesis is dose- or route-dependent. This could occur if a chemical produces tumours at the point of contact for one route of exposure, but there is adequate evidence that no tumours occur at other sites when a different route of exposure occurs.

The narrative could also point out if there is evidence that a key event in the carcinogenic process may not occur below a certain dose range. This provides scope for evaluating carcinogenic responses in animal studies where high doses have resulted in a precursor cellular proliferative response. If the studies are further informed by toxicokinetic or mode of action data that indicates a carcinogenic response in animals may not be relevant to humans, these findings should be built into the narrative.

It is clear that a degree of consistency has emerged between the US EPA, IARC and IPCS approaches to assessing the weight of evidence that suggests a chemical could be carcinogenic to humans.

11.3.3 Other international carcinogen classification schemes

The Canadian system uses six evidence-based categories for carcinogens, based essentially on the IARC classifications system.

- *Group I:* Carcinogenic to humans – essentially the same as IARC Group 1.
- *Group II:* Probably carcinogenic to humans – essentially the same as IARC Group 2A.
- *Group III:* Possibly carcinogenic to humans – as per IARC Group 2B, but sub-categorised into four sub-groups where the relationships between inconclusive epidemiological data and positive/equivocal animal data are expanded.
- *Group IV:* Unlikely to be carcinogenic to humans – essentially like IARC Group 2B, but where there are no useful epidemiological studies, and the four sub-groups are based on the relative strengths of association using animal studies, including whether there is evidence or not of genotoxicity.
- *Group V:* Probably not carcinogenic to humans – like IARC Group 4, where adequate epidemiological studies are lacking, but where the data in animal studies is more limited, and evidence indicating a lack of genotoxicity is stronger.
- *Group VI:* Unclassifiable with respect to carcinogenicity in humans – essentially like IARC Group 3, but including a sub-category where the evidence from epidemiological and animal studies is conflicting.

In the UK, carcinogens are defined and categorised in both the Control of Substances Hazardous to Health (COSHH) regulations (see www.hse.gov.uk/coshh) and the Chemicals (Hazard Information and Packaging for Supply) regulations (CHIP), although the term 'carcinogen' has a wider meaning in CHIP than in COSHH. The classification system is relatively simple, being based on three categories only.

- *Category 1:* substances known to cause cancer on the basis of human experience
- *Category 2:* substances which it is assumed can cause cancer, on the basis of reliable animal evidence
- *Category 3:* substances where there is only evidence in animals that is of doubtful relevance to human health (i.e. the evidence is not good enough for Category 1 or 2).

11.4 WHEN IS A CARCINOGENIC RESPONSE RELEVANT TO HUMANS?

In characterising risk associated with carcinogens that have been identified using rodent studies, the carcinogenic mechanisms (preferably termed MoA) should be taken into account. For carcinogens that are unequivocally genotoxic, conservative risk assessment methodology requires extrapolation of the dose–response curve from a suitable point of departure (POD) to zero to derive a cancer slope factor (see Section 3.10.3). Genotoxicity is MoA that is considered relevant for carcinogenesis across species, including humans, so it is rare for a genotoxic carcinogen to be subjected to HRA by any other method.

Carcinogenesis is a multifactorial and multi-stage process, involving initiation, promotion and progression. The presence of multiple mutations in critical genes in cancer cells is a strong indicator that cancer is associated with accumulative irreversible DNA damage. While DNA damage or genotoxicity is probably involved at one or more stages, there may also be stages of the carcinogenesis process where epigenetic events or non-genotoxic mechanisms may act to increase cancer incidence. For example, non-genotoxic chemicals may increase cancer incidence by stimulating cellular proliferation following a cytotoxic response (tissue injury), or by promoting the further carcinogenic transformation of cells which have already undergone genetic predisposition to a pre-cancerous state. Non-genotoxic mechanisms can also include hormonal effects, immunosuppression, receptor activation (e.g. dioxins/Ah receptor peroxisome proliferators/PPAR receptor).

The extent to which non-genotoxic MoAs contribute to the universe of chemicals that have been classified positive for carcinogenicity by IARC (Categories 1, 2A and 2B) was reviewed by Hernandezet al. (2009). They found that 12 per cent (45/371) of these proven and likely human carcinogens had an MoA that included a non-genotoxic component, with 27 per cent of these (12/45) posing a potential human hazard as assessed through a margin of exposure (MoE) approach based on exposure assessment potential.

In some cases, it is difficult to differentiate the significance of genotoxic and non-genotoxic MoAs for a specific chemical. For example, formaldehyde causes nasopharyngeal cancer in rats after inhalation exposures, and while formaldehyde has some genotoxic potential, the initiation of nasal cancers in rats appears only to follow strong irritancy of the nasal mucosa, followed by tissue repair involving a cellular proliferative response. Chloroform is another example of a weakly genotoxic chemical whose carcinogenic response in the kidney and liver appears to be determined by the extent to which cytotoxicity, followed by tissue repair, has occurred in those target tissues.

A very good example of how differentiation of genotoxic and non-genotoxic MoAs can impact on the quantitative risk assessment process is found in a published risk assessment of 1,4-dichlorobenzene (Butterworth et al. 2007). See Box 3.

Butterworth (2006) provides a framework for assessing when a genotoxic MoA may be important in the carcinogenic process, and may be differentiated from cytotoxic and regenerative cell proliferative responses that occur in high-dose rodent bioassays and which may confound the MoA assessment. Butterworth proposed (pp. 19–20) the following classification criteria for a carcinogen acting via a genotoxic MoA:

BOX 3: Example of the application of MoA analysis

1,4-dichlorobenzene induced liver cancer in male and female mice after either gavage or inhalation exposure. No liver cancers were observed in rats. The inhalation route is the primary source for likely human exposure. The MoA for the mouse liver tumours appears to be via a non-genotoxic mitogenic stimulation, with a putative threshold established from the dose–response relationship. Kidney tumours seen in male rats were probably acting via deposition of alpha-2μ-microglobulin and were not included in the risk assessment. The gavage and inhalation data were used to establish a BMD_{01} (1 per cent extra risk) for both routes of exposure, as a POD for further risk assessment. Using a conventional threshold approach, and dividing the POD by a 300x uncertainty factor, yields an air concentration of 0.1 ppm below which there should be no appreciable cancer risk to humans. In contrast, if an assumption of a genotoxic MoA had led to the use of a non-threshold approach to the risk assessment, the estimate of the air concentration associated with a 10^{-6} lifetime excess cancer risk would have been 0.00004 ppm, about 2,500-fold lower (and 1,875,000 times lower than the actual inhalation NOAEL of 75 ppm).

Evidence that a chemical can induce mutagenic activity:

- *Clear evidence for mutagenic activity:* Genotoxic activity is observed consistently in several well-validated cell culture assays with different endpoints. Chemically induced genotoxic activity is expressed in whole animal assays, particularly at the tumour target site.
- *Some evidence for mutagenic activity:* Genotoxic activity is observed in several well-validated cell culture assays with different endpoints. No activity, however, is expressed in whole animal assays.
- *Equivocal evidence for mutagenic activity:* Negative response in tests that include those assays which over time have proven to be most reliable and relevant. No responses are seen in whole animal genotoxicity assays. Positive responses are seen only in unvalidated or unproven assays or in assays where the experimental procedures used were questionable.
- *Clear evidence for lack of mutagenic activity:* Negative responses in tests that include those assays which over time have proven to be most reliable and relevant. No responses in whole animal genotoxicity assays. The weight of evidence indicates that the compound is not genotoxic.
- *Insufficient data to make a conclusion regarding genotoxic activity:* No or minimal testing has been conducted with the compound. Compounds that have structural alert for potential genotoxicity, but for which there is no experimental data may be placed in this category.

Evidence that the chemical produced cancer via a mutagenic mode of action:

- *Clear evidence that cancer was induced by a mutagenic mode of action:* Clear evidence for mutagenic activity of the compound. Experimental evidence for the likelihood that genotoxic activity was the predominant driving force in tumour formation, such as the identification of DNA adducts at the target tissue site.
- *Some evidence that cancer was induced by a mutagenic mode of action:* Some evidence for mutagenic activity of the compound. There may be a question as to whether genotoxic activity was occurring in the target tissue. Other biological activity was evident that could also have driven tumour formation.
- *Equivocal evidence that cancer was induced by a mutagenic mode of action:* Equivocal evidence for mutagenic activity of the compound. Clear evidence that other biological activity was likely driving tumour formation. Such activity would include endpoints such as necrosis, cytolethality, regenerative cell proliferation, peroxisomal proliferation, and associated mitogenic and promotional activity in the rodent liver and receptor-mediated mitogenic or other activity at doses associated with tumour formation.
- *Evidence that cancer was induced by both mutagenic and non-mutagenic activity:* In many cases, tumours are formed by a combination of mutagenic and non-mutagenic activity. In such instances, it is important to approximate the relative contributions of the two different activities in order to select a more realistic risk assessment.
- *Clear evidence that cancer was induced by a non-genotoxic mode of action:* Weight of evidence indicates an equivocal or clear lack of evidence for mutagenic activity of the compound. Other biological activity known to have the potential to drive tumour formation is associated with the compound at the target site, such as cytolethality and regenerative cell proliferation in a dose-dependent manner that indicates it was likely responsible for tumour formation.

While these suggested guidelines from a research scientist may provide a useful framework for decision making on the relevance of a proposed MoA, they are not proposals that have been endorsed by any regulatory agency. However, they are reasonably consistent with guidance that has been formulated in the IPCS and ILSI programs (see Section 11.5).

Determination and evaluation of the potential MoA is a key feature of the IPCS and US programs aimed at harmonising risk assessment for carcinogens.

11.5 IPCS PROGRAM ON CHEMICAL CARCINOGENESIS

The IPCS program on harmonising risk assessment for carcinogens began in 1997. The structure of the framework and the key steps proposed for a carcinogen evaluation were described in Appendix 5 of the original 2002 enHealth EHRA guidance document, as they stood at April 2001. The following text is reproduced from that appendix.

Introduction

This section describes the cancer endpoint or endpoints that have been observed and identifies which of these is addressed in the analysis. (The nature of the framework is such that only one mode of action is analysed at a time hence, for example, tumour types associated with a different mode of action, even if recorded in the same animals, will require separate framework analyses.) However, where different tumours are induced by related mode of action, they are best addressed in a single analysis. It should also be noted that some modes of action will involve multiple contributing components.

Postulated mode of action (theory of the case)

This section comprises a brief description of the sequence of events on the path to cancer for the postulated mode of action of the test substance. This explanation of the sequence of events leads into the next section which identifies the events considered 'key' (i.e. measurable) given the database available for the analysis.

Key events

This section briefly describes the 'key events' – i.e. measurable events that are critical to the induction of tumours as hypothesised in the postulated mode of action. To support an association, a body of experiments needs to define and measure an event consistently.

Pertinent observations: e.g. tumour response and key events in same cell type, sites of action logically relate to event(s), increased cell growth, specific biochemical events, organ weight, histology, proliferation assays, hormone or other protein perturbations, receptor–ligand changes, DNA or chromosome effects, cell cycle effects.

For example, key events for tumours hypothesised to be associated with prolonged regenerative proliferation might be cytotoxicity in as measured histopathologically and an increase in labelling index. Key events for induction of urinary bladder tumours hypothesised to be due to formation of bladder stones composed primarily of calcium phosphate might include elevated urinary calcium, phosphate and pH, and formation of bladder stones followed by irritation and regenerative hyperplasia of the urothelium.

Dose–response relationship

This section should detail the observed dose–response relationships and discuss whether the dose–response for the key events parallels the dose–response relationship for tumours. Ideally, one should be able to correlate increases in incidence of a key event with increases in incidence or severity (e.g. lesion progression) of other key events occurring later in the process, and with the ultimate tumour incidence. Comparative tabular presentation of incidence of key events and tumours is often helpful in examining dose–response.

Temporal association

This section should detail the observed temporal relationships or sequence of events and discuss whether the key events precede the tumour response. One should see the key events before tumour appearance this is essential in deciding whether the data support the postulated mode of action. Observations of key events at the same time as the tumours (e.g. at the end of a bioassay) do not contribute to temporal association, but can contribute to analysis in the next section. Most often, complete datasets to address the criterion of temporality are not available.

Strength, consistency and specificity of association of tumour response with key events

This section should discuss the weight of evidence linking the key events, precursor lesions and the tumour response. Stop/recovery studies showing absence or reduction of subsequent events or tumour when a key event is blocked or diminished are particularly important tests of the association. Consistent observations in a number of such studies, with differing experimental designs increases that support since different designs may reduce unknown biases or confounding.

Consistency, which addresses repeatability of key events in the postulated mode of action for cancer in different studies, is distinguished from coherence, which addresses relation of the postulated mode of action with observations in the broader database. Pertinent observations: e.g. tumour response and key events in same cell type, sites of action logically relate to event(s), initiation-promotion studies, stop/recovery studies.

Biological plausibility and coherence

The postulated mode of action and the events that are part of it need to be based on current understanding of the biology of cancer to be accepted, though the extent to which biological plausibility as a criterion against which weight of evidence is assessed is necessarily limited, due to considerable gaps in our knowledge in this regard. One should consider whether the mode of action is consistent with what is known about carcinogenesis in general (biological plausibility) and in relation to what is also known for the substance specifically (coherence). For the former, likeness of the case to others for structural analogues may be informative (i.e. structure activity analysis). Additionally, this section should consider whether the database on the agent is internally consistent in supporting the purported mode of action, including that for relevant non-cancer toxicities. Some modes of action can be anticipated to evoke effects other than cancer, e.g. reproductive effects of certain hormonal disturbances that are carcinogenic. Moreover, some modes of action are consistent with observed lack of genotoxicity. Coherence, which addresses relation of the postulated mode of action with observations in

the broader database – for example, association of mode of action for tumours with that for other endpoints – needs to be distinguished from consistency, which addresses repeatability of key events in the postulated mode of action for cancer in different studies.

Other modes of action

This section discusses alternative modes of action that logically present themselves in the case. If alternative modes of action are supported, they need their own framework analysis. These should be distinguished from additional components of a single mode of action which likely contribute to the observed effect, since these would be addressed in the analysis of the principal mode of action.

Assessment of postulated mode of action

This section should include a clear statement of the outcome with an indication of the level of confidence in the postulated mode of action e.g. high, moderate or low.

Uncertainties, inconsistencies and data gaps

Uncertainties should include those related to both the biology of tumour development and those for the database on the compound of interest. Inconsistencies should be flagged and data gaps identified. For the identified data gaps, there should be some indication of whether they are critical as support for the postulated mode of action or simply serve to increase confidence therein.

11.5.1 Recent updates to the IPCS framework

The outcomes of the IPCS project were summarised in a review article in 2006 (Boobis et al. 2006).

There have been a series of workshops that have evaluated the application of the framework using a selected range of substances with differing modes of action. For example:

- conazoles – cytotoxicity, cellular proliferation, metabolic induction
- d-limonene – cytotoxicity, cellular proliferation, metabolic pathway
- di-ethylhexylphthalate – cytotoxicity, cellular proliferation, peroxisome proliferation
- vinclozolin – cytotoxicity, cellular proliferation, hormonal perturbation
- captan – cytotoxicity, cellular proliferation, genotoxicity.

Since 2000 there has been substantial cooperation and alignment between the carcinogenic framework programs of the IPCS, the US EPA and a similar program of the International Life Sciences Institute (ILSI) (Meek et al. 2003). A special issue of *Critical Reviews in Toxicology* (Boobis et al. 2006) addresses the integration of cancer assessment frameworks developed by these three programs. The overall approach of all these programs has been essentially the same, with emphasis on evaluating the MoA and human relevance of the carcinogenic response. The special issue of *Critical Reviews in Toxicology* provided further illustration of the framework application using three additional case studies:

- 4-aminobiphenyl – DNA reactivity
- formaldehyde/glutaraldehyde – nasal cytotoxicity
- thiazopyr – thyroid disruption.

The various projects have distilled down to focus on three key questions (Cohen et al. 2004):

1. Is the weight of evidence sufficient to establish the MoA in animals?
2. Are key events in the animal MoA plausible in humans?
3. Taking into account kinetic and dynamic factors, are key events in the animal MoA plausible in humans?

11.6 QUANTITATIVE RISK ASSESSMENT OF CARCINOGENS

The risk assessment of carcinogens presents special challenges in quantitative EHRA. In part, this is because the methodology has been driven by the concept that there is no threshold for carcinogenic risk. The consequent use of mathematical constructs of the dose–response relationship to provide a platform for extrapolation from a POD is the mainstay of quantitative HRA of carcinogens.

There is an extensive discussion in Sections 3.10, 3.11, 3.12, 5.4.2, 5.7 and 5.10 of the basis for the assumption of no threshold for a carcinogenic response, on the application of non-threshold dose–response assessment methodologies and the reliance on establishment of a policy-driven 'target risk' for the evaluation of the EHRA outcomes. While there has been some movement towards accepting a threshold approach for carcinogenic responses associated with some non-genotoxic modes of action, the dominant method of dose–response assessment in Australian EHRA practice and regulatory requirements remains the non-threshold approach taken in US regulatory agencies.

The detailed critique of quantitative cancer risk assessment provided in Hrudey (1998) is quite informative and still relevant. The continued use of a non-threshold approach for most carcinogenic risk assessments results in quite conservative, and potentially costly outcomes in many types of EHRA processes.

11.7 AN ALTERNATIVE APPROACH TO QUANTITATIVE CARCINOGEN RISK ASSESSMENT?

While there is strong precedent for both Australian and international regulatory systems to expect application of a non-threshold approach to risk assessment of carcinogens, in 1999 the National Health and Medical Research Council (NHMRC), acting on advice from a specialist working party, endorsed a set of guidelines for risk assessment of carcinogens on contaminated sites, which was based on a modified threshold approach, or modified benchmark dose (mBMD) methodology (NHMRC 1999a). While the NHMRC document has since been rescinded, it did outline an approach that could harmonise the HRA of cancer and non-cancer toxicological endpoints by moving away from the need to calculate or estimate a finite risk level for carcinogens alone, and the need to establish a 'target risk' with which this could be compared.

The essential features of this alternative approach, deriving a POD for carcinogenic risk assessment were:

- use of a range of mathematical models to explore the data 'fit' to a derived dose–response relationship
- estimation of a BMD at the 5 per cent excess risk level (BMD_{05}) to serve as a 'point of departure' (POD) for further risk assessment
- application of a range of modifying factors to the BMD_{05} dose estimate, in order to derive a 'guideline dose' (GD), which should be protective of human health.

The NHMRC mBMD methodology has many similarities with the BMD methodology used by the US EPA to derive a POD for its carcinogen risk assessments, but there are also some important differences. The NHMRC advocates estimation of a mean BMD_{05} from all the curve-fitting data, after discarding those models which have an evidently poor fit to the data. The US methodology derives an upper 95th percentile estimate of the BMD_{10} as a POD for linearised extrapolation to estimate doses associated with the target risk used in US regulations (usually 10^{-6}; see Section 5.6). The use of a linearised extrapolation to zero dose using this method does not differ conceptually from the linearised multi-stage procedure because the slope factor derived is anchored and therefore strongly influenced by passing through the origin (zero dose – zero response).

The modifying factors (MF) applied to the BMD_{05} to derive a GD include a mixture of those used to adjust a NOAEL to derive an ADI, TDI or RfD (inter-and intra-species variability adequacy of the database and data quality) as well as some which are based on the nature of the carcinogenic response (genotoxic versus non-genotoxic rare, benign or malignant tumours produced). The range of possible MF values is up to 50,000. It is noted that, application of a MF of 50,000 to the BMD_{05} would be equivalent to the derivation of a 10^{-6} excess risk estimate using a linear extrapolation approach, thus providing a similar level of conservatism where the data warrants use of such a large MF value. Refer to the NHMRC document for further guidance on how to assess the carcinogenic hazard and how to select and apply the various MFs.

A key driver for the development of the now-rescinded NHMRC mBMD methodology was its lack of dependence on defining an 'acceptable' or 'target risk' level, and driving the risk assessment towards an excess risk estimate (e.g. 10^{-6}), which may be either excessively conservative or widely misunderstood in the community, or perhaps both.

While the methodology retains an element of 'expert judgement' in relation to the curve-fitting and selection/justification of the MFs, it is largely insulated from the science policy decision of what constitutes an acceptable level of risk (see also discussion in Section 5.6). This methodology is also not captive to the extrapolation through the origin, which risks the artefact of correlation between the CSF and the MTD (Krewski et al. 1993, Hrudey 1998).

The NHMRC method also represents an alignment with the methodology used for risk assessment of virtually all other toxicological endpoints. There have been a number of commentators in the scientific literature who have argued that it is no longer justifiable to differentiate between threshold and non-threshold approaches to carcinogenic risk assessment (see Gaylor et al. 1999; Purchase & Auton 1995).

Chapter 12: Assessing multiple routes and sources of exposure

The passing of the US *Food Quality Protection Act of 1996* put the onus on the US EPA to develop scientifically sound methodologies to incorporate 'aggregate' and 'cumulative' methodologies into environmental health risk assessment programs, particularly with regard to pesticide residues in food. The program led to the development of model software termed CARES – Cumulative and Aggregate Risk Evaluation System.

In this US context, the meaning of these terms was:

- *Aggregate exposure:* the analysis of exposure to a chemical by multiple pathways and routes of exposure
- *Cumulative exposure:* the combined risk estimate where exposure occurs simultaneously, or consecutively, to multiple chemicals that exert toxicity through a common mechanism.

A key factor in US 'cumulative' HRA methodologies is the determination of whether the multiple chemicals act by a common mechanism of action. In the earliest determinations, this was relatively easy, since the methodologies focused on cumulative risk assessment of organophosphonate (OP) pesticides, which act via the common mechanism of inhibiting acetylcholinesterase (AChE) (US EPA 2001b). Subsequently, the work was extended to include other pesticides (N-methyl carbamates, triazines and chloroacetanilides), which also share a common mechanism of action (see <www.epa.gov/oppsrrd1/cumulative>).

The US EPA finalised its first guidance document on the application of cumulative risk assessment in 2002 (US EPA 2002b), with publication of a framework document in 2003 (US EPA 2003a). Guidance on aggregate risk assessment, initially promulgated in 1997, had been finalised the year before (US EPA 2001a). The European Union has also published a comprehensive review of practices used in the risk assessment of mixtures (Kortenkamp et al. 2009).

12.1 RISK ASSESSMENT OF CHEMICAL MIXTURES

The conventional approach to health risk assessment often relies on evaluating toxicity data where chemicals have been administered as single entities using fixed dose rates. However, in the real world, exposures may occur simultaneously or consecutively with multiple chemicals and doses varying over time. Assessment of such complex situations presents a real challenge to toxicologists and risk assessment professionals.

One of the more important challenges is to determine whether the combined risk is independent of the individual components whether the risk is additive (i.e. no interactions among components) more than additive (i.e. synergistic) or less than additive (antagonistic). All are theoretically possible at high doses, although there have been relatively few opportunities to test whether additive models or those based on independent actions are better predictors of mixtures toxicology. One such review (Cedergreen et al. 2008) found that the predictive capability of either concept was relatively poor when 158 datasets relating to tests with binary mixtures were evaluated. On the other hand, Pohl et al. (2009) analysed 380 binary combinations from the ATSDR database and found that only 156 (41 per cent) indicated possible additive effects, with 57 showing antagonism (less than additive) and 91 had insufficient information to be classified. Importantly, 16 combination effects suggested synergism, with mechanistic data supporting such interactions, while a further 50 combinations suggested synergism, but with incomplete supporting mechanistic data. Among the examples in this synergistic group were various combinations of arsenic, benzo(a) pyrene, chloroform, and PCBs and an atrazine/diazinon interaction based on a metabolic interference.

However, the key issue is whether interactions are likely to occur at the low doses commonly associated with environmental exposure scenarios. It has been proposed that there is an 'interaction threshold' representing a point below which interaction effects in mixtures may not be relevant (Hamm et al. 2005).

For most of the environmental scenarios that have been evaluated, the 'interaction threshold' appears to be higher than the dose or exposure where toxicity is seen, thus suggesting that interactions are unlikely to be relevant (Yang & Dennison 2007; Crofton et al. 2005). A possible exception is where the toxic effects are mediated via disruption of the endocrine system, a system potentially sensitive to very low doses of environmental chemicals (Kortenkamp et al. 2007).

While interactions between chemicals in an environmental mixture may or may not be relevant, it is generally assumed that additivity is the more normal outcome (Lambert & Lipscomb 2007). This may be expressed as either:

- *Dose addition:* Where the combined effects are assumed to be by the same mode of action (MoA) and are based on relative potency at a fixed level of response, or
- *Response addition:* Where the combined effects are assumed to be independent of the mode of action and defined by dose–response curves which may have varying slopes, and hence different potencies at defined levels of exposure.

Five basic approaches have been applied to the health risk assessment of chemical mixtures as outlined in the following sections (12.2 to 12.6).

12.2 ASSESSMENT OF REPRESENTATIVE MIXTURES

Simple binary mixtures, or in some cases more complex mixture matrices, may have been directly tested in conventional *in vivo* or *in vitro* toxicity testing systems. These tests allow an assessment of whether a fixed chemical mixture may cause toxicity that is either different in potency or not predicted by studies using the individual components.

Since the mixtures may have undergone a concentration step in order to administer doses that adequately reveal the toxicity profile, this approach presents the same high-to-low dose extrapolation issues for health risk assessment as conventional single chemical tests. The concentration step may also increase the likelihood of observing confounding interactions that may not be relevant at lower doses. Tests based on *in vitro* and *in silico* exposures may have used more relevant lower concentrations, but the interpretative problem then becomes one of extrapolating the dose–response and interactive relationships to whole organisms, including humans.

The obvious limitation in such an approach is the practical impossibility of testing more than a few possible permutations and combinations of the chemicals of interest.

While the testing of 'representative' mixtures has been applied to the risk assessment of some groundwater contaminants (Ryker & Small 2008), and fixed mixtures such as certain petroleum-based fuels, the database of such studies is relatively limited compared with that of more conventional single-chemical toxicity tests.

A further issue is that in the real environment, chemical mixtures may change in composition over time, due to differential translocation, degradation or 'weathering' effects. Such problems are not unique to the approach of testing 'representative' mixtures, since interpretation of environmental data based on single chemical toxicity data also needs to take into account time- and pathway-dependent variables in exposure.

12.3 THE TOXICITY EQUIVALENCE FACTOR APPROACH

The basic assumption in this approach is that the toxicity contribution of individual chemicals in a mixture can be summed by expressing the concentration or dose of each component in terms of a standard or reference compound. The amount of each component is adjusted by multiplication with a toxicity equivalence factor (TEF) based on relative toxicity potency. The derivation of TEFs is based on assessment of a relative potency factor (RPF) for the individual components (Chen et al. 2003).

The HRA approach sums the doses of individual components of a mixture using the formula:

$$\text{Total toxicity equivalents (TEQ)} = \Sigma\,(\text{component 1 amount/concentration} \times \text{TEF1})_n$$

It is generally used as the default methodology where it has been well established that the group of chemicals share a common MoA and where there is a reasonable basis for assigning relative toxicity potencies.

The best known example of the use of the TEF approach is with the family of 'dioxin-like' compounds (polychlorinated dibenzodioxins, polychlorinated dibenzofurans and dioxin-like polychlorinated biphenyls (PCBs). It is likely that most, if not all, of the toxicity of these families of chemicals is initiated by binding to a common intracellular receptor (the aryl hydrocarbon receptor or AhR) with subsequent downstream gene activation leading to a variety of cellular events (e.g. cellular proliferation and differentiation, growth factor and hormone regulation). For this class of AhR agonists, the relative potencies are related to 2,3,7,8-tetrachloro-dibenzodioxin (TCDD), putatively the most toxic congener and the one for which most toxicological data exists. Tables summarising the relative potencies of the dioxin-like compounds were first published during the mid-1980s and have since been updated as new information has come to hand (Van den Berg et al. 2007; US EPA 2000a; 2009c; NTP 2006).

Following an extensive review of the dioxin congener TEFs proposed by Van den Berg et al. (2006), the US EPA (2009c) has recently issued draft guidance that proposes endorsement of these TEFs as the best currently available estimates.

Another class of chemicals for which TEFs have been developed for EHRA purposes are the polycyclic aromatic hydrocarbons (PAHs). PAH TEFs developed as part of the contaminated sites workshop series in the 1990s have conventionally been used in EHRA in Australia (Fitzgerald 1993; 1998). These TEFs for selected PAHs are summarised in Table 20, which includes comparisons with TEFs developed by US EPA and the UK (from Fitzgerald 1998). Table 20 also contains estimates of RPFs for selected PAHs from an external review draft (US EPA 2010). While these values have not yet been endorsed by the US EPA, they are based on a comprehensive review of carcinogenic effects in several rodent bioassays covering the oral, dermal and inhalation routes of exposure. The US EPA 2010 review includes a table (Table 3-1) that makes a more comprehensive survey of PAH RPFs from published papers (1984–2004) and regulatory reviews. The variability shown in these estimates of RPFs indicate the difficulties in defining a consistent approach to comparing carcinogenic potencies across different site and different routes of exposure.

Table 19: Toxicity equivalence factors (TEFs) for dioxin congeners

Congener	TEF
Chlorinated dibenzo-p-dioxins	
2,3,7,8-TCDD	1
1,2,3,7,8-PeCDD	1
1,2,3,4,7,8 –HxCDD	0.1
1,2,3,6,7,8-HxCDD	0.1
1,2,3,7,8,9-HxCDD	0.1
1,2,3,4,7,8-HpCDD	0.01
OCDD	0.0003*
Non-ortho-substituted PCBs	
3,3',4,4'-tetraCB (PCB 77)	0.0001
3,4,4'5-tetraCB (PCB 81)	0.0003*
3,3',4,4'5-pentaCB (PCB 126)	0.1
3,3',4,4',5,5'-hexaCB (PCB 169)	0.03*

Congener	TEF
Chlorinated dibenzofurans	
2,3,7,8-TCDF	0.1
1,2,3,7,8-PeCDF	0.03#
2,3,4,7,8-PeCDF	0.3#
1,2,3,4,7,8-PHxCDF	0.1
1,2,3,6,7,8-HxCDF	0.1
1,2,3,7,8,9-HxCDF	0.1
2,3,4,6,7,8-HxCDF	0.1
1,2,3,4,6,7,8-HpCDF	0.01
1,2,3,4,7,8,9-HpCDF	0.01
OCDF	0.0003*
Mono-ortho-substituted PCBs	
2,3,3',4,4'-pentaCB (PCB 105)	0.00003#
2,3,4,4',5-pentaCB (PCB 114)	0.00003#
2,3'4,4',5-pentaCB (PCB 118)	0.00003#
2',3,4,4',5-pentaCB (PCB 123)	0.00003#
2,3,3',4,4',5-hexaCB (PCB 156)	0.00003#
2,3,3',4,4',55-hexaCB (PCB 157)	0.00003#
2,3',4,4',5,5'-hexaCB (PCB167)	0.00003*
2,3,3',4,4',5,5'-heptaCB (PCB 189)	0.00003#

* Adjusted 3x upwards (approx half log10) from the 1998 WHO values in the 2005 review.

adjusted downwards from the 1998 WHO values in the 2005 WHO review.

Source: Van den Berg et al. 2006, as cited in the 2005 WHO review.

Table 20: TEFs values for PAHs

PAH	Fitzgerald 1998	US EPA 1993	UK 1995[1]	US EPA 2010[2]	CCME 2010[3]
Benzo(a)pyrene	1	1	1		1
Dibenz(a,h)anthracene	4	1	1	10 (1–40) h	1
Benz(a)anthracene	0.1	0.1	0.1	0.2 (0.02–0.04) m	0.1
Benzo(b)fluoranthene	0.1	0.1	0.1	0.8 (0.1–2) h	0.1
Benzo(k)fluoranthene	0.1	0.01	0.1	0.03 m	0.1
I-pyrene	0.1	0.1	0.1	0 m	nd
Anthracene	0.001	nd	0.001	0 m	nd
B-perylene	0.1	nd	0.1	nd	nd
Chrysene	0.1	0.001	0.1	0.1 (0.04–0.2) h	0.01
Acenaphthene	0.001	nd	nd	nd	nd
Acenaphthylene	0.001	nd	nd	nd	nd
Fluoranthene	0.01	nd	0.001	0.08 (0.009–0.2) l	nd
Fluorene	0.001	nd	nd	nd	nd
Naphthalene	0.001	nd	nd	nd	nd
Phenanthrene	0.001	nd	0.001	0 h	nd
Pyrene	0.001	nd	0.001	0 m	nd

Adapted from: Table 5, Fitzgerald 1998.

nd = not determined

1 Cited in CRBE (1995).

2 Cited in an external review draft; not yet endorsed by US EPA; figures quoted are RPFs based on a review of rodent tumour bioassays, using different exposure routes; figures in parentheses indicate the range of estimated RPFs; the letters h, m and l denote high, medium and low confidence in the reported figures.

3 A set of values used by Canadian authorities in assessing contaminated sites.

The Canadian report (CCME 2010) includes comprehensive tables of TEF values for PAHs derived by a broader range of authorities, along with the source references. However, the report notes that more than a dozen sets of equivalency numbers have been proposed over the past two decades and cautions that there can only be limited confidence in the derived potency estimates. Adoption of any one set of values requires an understanding of the factors that can create uncertainty:

- prediction of mixture effects based on single-substance studies
- prediction of potency for one route of exposure based on data from another route of exposure
- possible presence of other carcinogens or promoters in the mixture that are not measured or addressed in the relative potency estimate
- uncertainty about the relevance of rodent studies in predicting human cancer risk

At this time, no one set of PAH TEFs has been recommended for use in Australia, although it is likely that the Canadian set is becoming more widely used, based on the fact that it the most recent compilation of such values.

Relative toxic potencies of some endocrine disrupting chemicals (EDCs) have also been determined towards activation of key endocrine receptors (e.g. the oestrogen receptor complex) and it is possible to use the TEF approach to construct aggregate risk assessment models based on relative potencies for this class of chemicals.

The TEF approach has some key limitations. Since it is assumed that chemicals in a TEF class have the same MoA, the toxicity is assumed to be dose-additive and that there are no significant interactions (e.g. receptor antagonism or modifications to metabolic clearance). It is also assumed that the relative potencies are based on similar test designs and endpoints. TEFs based on relatively simple or short-term endpoints may not

necessarily reflect longer-term toxicity where differences in tissue distribution, clearance and elimination may alter or confound the relative potency estimates. In the case of dioxin-like compounds, the TEFs are actually based on a composite of endpoints from short- and medium-term exposures, including some relative toxicity endpoints based on *in vitro* tests systems, not all of which are strictly relevant for assessing toxicity associated with lifetime exposures (e.g. cancer). In a recent 2-year rat bioassay study designed to assess the predictive powers of TEFs for dioxins and dioxin-like PCBs (Walker et al. 2005; NTP 2006), ternary mixtures of TCDD, PeCDF and PCB 126 were administered at dose levels representing the four TEF-equivalent doses used for TCDD alone in previous NTP studies. While the dioxin mixtures produced a range of cellular proliferative and carcinogenic responses comparable to that seen with TCDD alone, the dose–response relationships were not identical. Therefore, while this NTP study provided some support for the TEF concept, it did show that the numerical values for TEFs may not be applicable across the full range of toxicological outcomes (Gray et al. 2006). On the other hand, Smialowicz et al. (2008) were able to validate the US EPA dioxin-like TEFs to predict immunotoxic endpoints, based on their CYP inductive effects. A more comprehensive review of the methodologies and databases used to establish dioxin TEFs has been published by Haws et al. (2006).

12.4 SUMMATION OF RISK ESTIMATES

This approach is the one most commonly used in health risk assessments. The contribution of an individual mixture component is calculated as a ratio of predicted exposure compared with a health-based benchmark (ADI, TDI). This derived ratio is termed a hazard quotient (HQ). To derive an overall estimate of risk across all components of the mixture, the individual HQs are summed to derive a hazard index (HI) (Herzberg & Teuschler 2002).

$$HI = \sum_{j=1}^{n} \frac{\mathrm{E}_j}{\mathrm{R}f\mathrm{D}_j} = \sum_{j=1}^{\mathrm{n}} HQ_j$$

E_j = exposure level of chemical *j*

RfD_j = RfD of chemical *j*

HQ_j = HQ for chemical *j* (dimensionless)

RfD is US terminology equivalent to the ADI or TDI

While this pragmatic approach may be a useful screening tool, is used quite extensively in HRA practice and is generally accepted, or even required, by Australian regulatory authorities, it is based on a potentially flawed premise. As noted above, additive risk estimates imply a common MoA, or at least a common toxicological outcome. Where the toxicity of individual components does not satisfy such an assumption, the contributions to risk are theoretically independent.

The HI equation may be modified to take into account the potential for chemicals in a mixture to interact and refinement of the exposure assessment using target tissue dose estimates (Herzberg & Teuschler 2002; Pohl & Abadin 2008). Such modifications require a more extensive database and use a weight-of-evidence (WoE) approach to evaluate such interactions and to modify terms in the above equation.

The HI approach is essentially quite conservative in providing an estimate of cumulative risk, since 'safety factors of between 100 and 10,000 are commonly used to adjust the estimated no observed adverse effect level (NOAEL) to derive the ADI or TDI estimates. This means that the effects of the chemical combinations would need to significantly erode this 100–10,000-fold margin between the ADI/TDI and the level where toxic effects begin to occur. When the overall HI is less than 1, it is generally assumed that cumulative risk is within reasonable bounds and that there is no need to undertake a more refined risk assessment. Even when HI is greater than 1, it does not imply that risks are unacceptable, although there is clearly some erosion of the conservatism built into each of the processes of determining components of the HQ calculation (exposure and TDI). When the HI is greater than 10 there is more reason to undertake further investigation of the risks, including an assessment of whether addition of HQs is justified or whether the risk contribution of some of the components is independent.

Nevertheless, it is a common approach in the professional practice of EHRA in Australia (especially when dealing with government agencies) to not only require the calculation of the HI for mixtures, but to undertake further refinement of the risk estimates when the HI exceeds 1. There may be situations where additional information will be available that shows that the HI approach will overestimate the risk and scientific argument could be advanced to show this approach to be inappropriate.

Where the health risk assessment approach makes a finite assessment of the level of risk and compares it to some form of expressing an acceptable or 'target risk', as is the case for the non-threshold risk associated with carcinogens, a similar approach may be used to assess the combined risk of a mixture. In this case, the risk estimates for individual components derived from the carcinogenic slope factor and exposure estimates are simply summed to provide an overall risk estimate.

In some cases, the 'target risk' estimate for the mixture may be set at the same level as that for individual components (e.g. 1×10^{-6}). This is particularly conservative, since the addition of risks implies a common MoA and outcome, where the risk contributions of individual carcinogens in a mixture may in fact be independent. To overcome this inherent conservatism, the cumulative carcinogenic risk is sometimes adjusted to a higher level than that applied to individual components. For example, where the 'target risk' for an individual component may be set at 1×10^{-6}, the target risk for a mixture may be adjusted to 1×10^{-5} (see Section 5.10).

12.5 COMPONENT ELIMINATION OR SIMPLIFICATION

For mixtures where a common MoA or other determinants of risk additivity cannot be assumed, the default approach becomes an assumption of toxicity independence for components of the mixture. In such a case, the risk estimate (e.g. HQ or assessed risk level) associated with the most toxic components or those with the greatest exposure potential are assessed independently and may drive the risk assessment. This approach makes the assumption that overall risk is no greater than that of the riskiest component and that no additive or interactive effects alter the risk.

In order to simplify the process by eliminating those components of a mixture that make little or no effective contribution to risk, there is an increasing use of the concept that a toxicological threshold of concern (TTC) may be derived for risk assessment purposes (see Section 5.13). This concept may alternatively be called the concentration of no toxicological concern (CoNTC) and it is not really a new concept, since it has been applied to various areas of drug and food regulation in the US and Europe since 1993 (Drew & Frangos 2007). In an environmental context, the TTC concept has been embodied in recent Australian guidance for assessment of recycled water (EPHC, NHMRC, NRMMC 2008). The TTC is a concentration or amount (dose) derived by analysing the distribution of known toxic potencies for an extensive chemical toxicity database, and selecting an arbitrary low percentile (e.g. 5 per cent) of that distribution such that it is unlikely that the toxicity potential of any unknown chemical would be greater that the TTC. Where components of a complex mixture can be shown to be below the TTC, they may be disregarded in the risk assessment.

12.6 BIOMARKER APPROACH

The above approaches to risk assessment all require detailed knowledge of the components of a mixture, including their relative concentrations or exposures/doses and their toxicological characteristics. An alternative approach, which may overcome some of the gaps in this knowledge database, requires identification of a suitable biomarker of effect (Ryan et al. 2007). Such a biomarker would represent a common event in the progression between exposure and a toxicological outcome. In such a case, the biomarker becomes the measure that aggregates the effects of all components of a mixture that operate via this pathway.

Some of the areas in which such a biomarker approach may be useful in risk assessment could include:

- using DNA adducts or other cytological markers of DNA damage following exposure to a complex mixture of polycyclic aromatic hydrocarbons (PAHs)
- determining the extent of cholinesterase inhibition following exposure to organophosphonate and/or carbamate pesticides.

The main limitation on the biomarker approach to mixture risk assessment is an incomplete understanding of the extent to which the biomarker reflects the ultimate expression of toxicity.

12.7 INTERNATIONAL PROGRAMS AND REVIEWS

International risk assessment programs have been quite active in addressing the complex problems of mixture risk assessment. These programs include the following.

- US EPA (2000; 2003a) – guidance for conducting health risk assessment of chemical mixtures, which has interleaved with the development of programs for assessment of cumulative and aggregate risks. A further review of the US EPA methodology for assessing multiple chemical exposures and effects using dosimetry-adjustment using PBPK modelling finalised in 2006–07 (US EPA 2006b; 2007c).
- ATSDR guidance for the preparation of an interaction profile (2001), with the subsequent publication of mixtures profiles for a number of pesticides and environmental contaminants, a formal guidance manual for assessing mixtures (ATSDR 2004), and an overview of the ATSDR program by De Rosa (2004).
- Health Council of the Netherlands program on exposure to combinations of substances (Feron et al. 2004).
- Danish Veterinary and Food Administration (2003) program for assessing toxicological effects of exposures to mixtures of industrial and environmental chemicals.
- Bjarnason (2004): *Toxicology of chemical mixtures: a review of mixtures assessment*, for the Canadian Defence R & D as part of a NATO project NATO HFM-057/RTG-009.

- Project IRSST (2005): *Impact of toxicological interactions on the management of exposure to multiple contaminants: a Canadian program directed towards occupational exposures*. See <http://www.irsst.qc.ca/en/_projet_2973.html>
- NOMiracle (2005): *Novel methods for integrated risk assessment of cumulative stressor in Europe.* See <http://nomiracle.jrc.ec.europa.eu/default.aspx>
- ILSI HESI project – an ongoing industry consortium approach to risk assessment methodology, including approaches to assessing chemical mixtures (Smith et al. 2008).
- IGHRC (Interdepartmental Group on Health Risks from Chemicals) (2009). Chemical mixtures: a framework for assessing risks for UK regulatory authorities.
- IPCS (2009). The IPCS Program for Harmonization of Risk Assessment released a report on an international workshop on combined exposures (IPCS 2009a) and followed up with a consultation document in July 2009 (IPCS 2009b).
- Project 070307/2007/485103/ETU/D.1 – *State-of-the-art review of mixtures toxicity*, assessment for the European Commission Environment DG (Kortenkamp et al. 2009).

The Kortenkamp et al. (2009) review for the EC posed, and answered, a series of questions:

- Is an assessment of the effects of chemical mixtures necessary from a scientific viewpoint?
- Is there not sufficient protection against mixture effects if we make sure that each chemical is present individually at exposures unlikely to pose risks?
- Is it necessary to test every conceivable combination of chemicals or is it possible to predict the effects of a mixture?
- Which of the two assessment and prediction concepts – dose addition or independent action – should be utilised in practice?
- Which chemicals should be subjected to mixtures risk assessment?
- How should mixture effect assessment concepts be applied in practice?
- What knowledge gaps hamper the consideration of mixture toxicology and ecotoxicology in chemical risk assessment?

The review provides the full answers to these questions. In summary, where chemicals in a mixture have diverse modes of action, theory suggests that such independent actions should not yield a combination effect. However, concepts of dose additivity suggest that toxic effects of a mixture could be seen even when individual components are below (presumably only slightly below) their individual NOAELs. The authors propose that a dose addition approach is likely to provide a more realistic estimate of mixture toxicity than assumption of an independent action model. This is a more conservative approach than that recommended in Section 12.5.

One of the more constructive outcomes of the IPCS framework proposals is a diagram that outlines a tiered hazard and exposure flow chart for assessing chemical mixtures (see Figure 5, Chapter 1).

12.8 OTHER REVIEWS OF CHEMICAL MIXTURES HRA

Some of the earlier literature on mixtures toxicology was reviewed in special issues of journals (e.g. *Food Chemical Toxicology* Vol. 34 Nov–Dec 1996).

A selection of reviews and commentaries over the past decade reflecting the historical development and outcomes of programs for chemicals mixture assessment include: Groten et al. (2001), Haddad et al. (2001), Feron et al. (2002), Teuschler et al. (2002), Zeliger (2003), Altenburger et al. (2004), Andersen & Dennison (2004), Borgert et al. (2004), Monosson (2005), Rider & LeBlanc (2005), Suk & Olden (2005), Biello (2006), De Rosa et al. (2004, 2006), McCarty & Borgert (2006a,b), Sexton & Hattis (2006), Arnold & Price (2007), Boekelheide (2007), Boobis (2007), Callahan & Sexton (2007), Kortenkamp (2007), Lambert & Lipscombe (2007), Mason et al. (2007), Menzie et al. (2007), Mumtaz et al. (2007), Teuschler (2007), Trivedi (2007), Yang & Dennison (2007), Pohl & Abadin (2008), Rice et al. (2008).

Chapter 13: Exposure modelling

'*All models are wrong, some models are useful.*'

George EP Box (1979)

Modelling is used in exposure assessment 'as a means of forecasting human or other exposures in the absence of complete monitoring or other data' (WHO 1999a). Modelling provides 'a mathematical expression representing a simplification of the essential elements of exposure processes'. Point estimates and probability distributions are used in exposure modelling.

Especially where monitoring data is inadequate, fate models are useful for estimating chemical concentrations. These models can span a wide range of complexity in terms of spatial dimensions and temporal assumptions (i.e. steady state versus non-steady state).

Types of fate models include (WHO 1999b):

- *simple dilution models*, where a measured concentration in an effluent is divided by a dilution factor, or the chemical release rate is divided by a dilution factor or the chemical release rate is divided by the bulk flow rate of the medium
- *equilibrium models*, which predict the distribution of a chemical in the environment based on partitioning ratios or fugacity (the escaping tendency of a chemical from one environmental phase to another)
- *dispersion models*, which predict reductions in concentrations from point sources based on assumed mathematical functions or dispersion properties of the chemical with physical processes like wind or river flow
- *transport models*, which predict concentration changes over distance, and can represent dispersion, biochemical degradation and absorption.

Because direct measurement of contaminant concentrations in potential exposure media (air, water, soil) may not available, or because direct sampling and analysis may not be practical, the default position can become the use of exposure models.

Exposure models may be informed by conceptual site models or flow diagrams describing specific sites or exposure scenarios (see Section 4.4) and typically rely on computer programs. It is important that users of these models are aware of their components and understand the nature and sensitivity of data inputs that can influence the outcomes and the assumptions built into the calculations within the model. In this context, it is recommended that part of the EHRA process include a 'sensitivity analysis', where the values of model input parameters are varied to ascertain the impact of the changes on the model outputs (see also Section 5.15).

Some typical modelling techniques used in EHRA are:

- *vapour intrusion models*, which estimate the amounts of vapour concentrations in ambient air – these usually arise from transport of vapours into confined spaces (e.g. building interiors) from underlying sources in soil groundwater or outside air
- *groundwater fate and transport models*, which track the dispersion of contaminants in groundwater plumes, reservoirs or discrete water bodies, and which may eventually result in 'receptors' coming in contact with these sources
- dust and particulate dispersion and deposition models, which track the passage of particulate matter (dusts, aerosols) from source to potential 'receptors'
- *gas and vapour dispersion models*, which map the distribution of such contaminants from point sources of emissions (e.g. factories, smokestacks, ventilation stacks)
- *plant and animal uptake models*, which track the transfer of contaminants from sources of agricultural production (pastures, gardens and crop sites) into food chains
- *spray drift models*, which assess the spread of pesticide applications from aerial, boom spray or other agricultural practices, and assess the extent to which off-target contamination and bystander exposure may occur
- *worker exposure models*, which are more strictly in the province of occupational health and safety (OHS) assessment, rather than EHRA. However, the techniques (especially biomonitoring methods – see Chapter 14) may also be useful in selected EHRA scenarios.

It is beyond the scope of this enHealth document to discuss in detail all the above types of exposure models. An extensive summary and critique of exposure models used for chemical risk assessment in the UK and the EU has been published by Fryer et al. (2004; 2006). More general UK guidance on exposure assessment is outlined in IGHRC (2004) and US guidance is outlined in ATSDR (2006) and US EPA 2002b).

Some comment on possible pitfalls or misuse of the Johnson-Ettinger model for vapour intrusion into buildings is outlined in Section 4.10, in the *Contaminated sites NEPM review* (NEPC 2010) and in CRC-CARE (2009). Particular points to be considered are whether adjustments need to be made for finite or infinite sources, the extent to which biodegradation can reduce the extent of vapour intrusion (Davis et al. 2009a, b) and other factors that can cause the Johnson-Ettinger model to overestimate or underestimate indoor air concentrations.

13.1
USING POINT ESTIMATES AND PROBABILITY DISTRIBUTIONS

Point estimates are most commonly used in Australia for exposure assessments. A point estimate is a single value chosen to represent a population such as 70 kg as the weight of an adult. Point estimates are usually typical or average values for a population or an estimate of an upper end of the population's value such as 70 years as the duration of residence on a property. An upper-end value may be chosen for reasons of conservatism or to provide a 'worse case' scenario.

Where a risk assessment uses a series of upper end estimates, the result can be a worse than 'worst case' scenario due to the compounding effects of the estimates (e.g. the person with the upper-end value for weight is unlikely to also have the upper-end value for water consumption, the upper-end value for contamination, the upper-end value for duration of residence and the upper-end value for soil ingestion.

A point estimate of a median or a mean is inherently more certain than a point estimate of the level intended to represent the 95th or 99th percentile because there is more data involved in determining the estimates of central points than for estimating extremes. This will present problems if there are limited data for using point estimates if the point estimate is intended to be, for example, the 95th percentile. For the same reason, similar problems will arise if the tails of a probability distribution are to be estimated.

Increasing attention has been paid to the use of Monte Carlo-type exposure assessments and such methods have been acknowledged by the US EPA and the UK Department of the Environment (US EPA 1992; Ferguson 1994). However, this trend may be reversing because of the paucity of good distribution data needed for probabilistic techniques such as Monte Carlo. More recently, development of UK soil guidance values (SGVs) using the contaminated land exposure assessment (CLEA) methodology has reverted to using single value (deterministic) data inputs rather than probabilistic techniques (NEPC 2010).

'While methods using probability distributions are 'more informative and inherently more representative' [Ruffle et al. 1994 p. 403] than point estimates, if applied appropriately point estimates still have a major role in exposure assessment as they are readily understood and applied, and may incorporate safety factors that could be lost with Monte Carlo-type exposure assessments' (Langley & Sabordo 1996).

Hattis and Silver (1994) propose that there will be greater uncertainties in estimates for the variability of a parameter value (i.e. the standard deviation) than the estimate of the parameter value (i.e. the mean).

In standard-setting, this is important if one is using point estimates, as a point estimate of a mean will be more certain than a point estimate of the level intended to represent the 95th or 99th percentile.

13.2
MONTE CARLO ANALYSIS

Monte Carlo-type exposure assessments rely on the use of probability distribution functions. This may be described as a 'distribution of possible values for each of the parameters [is] described along with the probability of occurrence of each value' (Alsop et al. 1993 p. 407). Using standard mathematical formulae, several thousand iterations of the output parameter are performed.

For each iteration, values for each parameter are selected randomly from each distribution based upon the probability of occurrence. The estimated risk values are combined to provide a frequency distribution of the output parameter, such as the consolidated estimate of exposure (from Langley et al. 1998).

Figure 30 demonstrates the process of using the Monte Carlo method to estimate the probability distribution of exposures in a population.

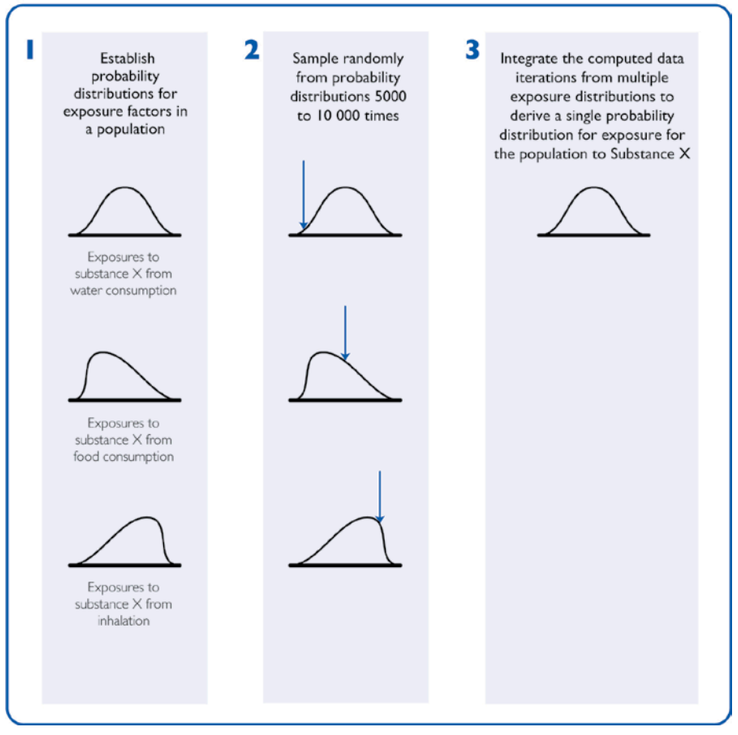
Monte Carlo analysis may add value to a risk assessment (US EPA 1997b) when:

- exposures and risks are likely to approach or be above levels of concern
- screening assessments using conservative point estimates fall above levels of concern
- it is necessary to disclose the degree of bias associated with point estimates of exposure
- exposures and exposure pathways need to be ranked
- there is a need to appraise the relative values of collecting different types of further information (Cullen & Frey 1999)
- the costs of action are likely to be high and the gains are likely to be marginal
- the outcomes of action affect different exposure pathways and the benefits need to be ranked (Cullen & Frey 1999)
- sensitivity analysis is needed to apprise the impact of default values and key pathways
- the consequences of simplistic exposure assessments are likely to be unacceptable.

Monte Carlo analysis may not add value to a risk assessment (Cullen & Frey 1999) when:

- exposures and risks are likely to be negligible
- the costs of reducing the exposure and risk are smaller than the cost of probabilistic analysis

Figure 30: Principles of the Monte Carlo method



Adapted from: Ferguson 1994.

- safety is an urgent concern and action must be taken rapidly
- probability distributions are so uncertain or incomplete that detailed probabilistic judgements are unreasonable
- there is little variability or uncertainty in the data.

If a Monte Carlo assessment is performed the methodology must be 'transparent' or problems will arise in community consultation. As with any form of risk assessment, the basic principles of the method must be able to be understood by the affected community.

For small-scale situations, the use of Monte Carlo methods is likely to be too complex or costly, and it may be more appropriate to use direct measurements of exposure. The exposures of high-end exposure 'outliers' must always be acknowledged in risk assessments, and ways of identifying and accommodating them must be considered. This is particularly important in the assessment of an existing situation (e.g. a contaminated site where housing has already been developed), rather than a forecast exposure scenario, as the presence of an 'outlier' will severely test the credibility of any risk communication exercise (Langley & Sabordo 1996).

13.2.1
Weaknesses with the Monte Carlo technique

Limitations of the Monte Carlo technique include the following:

1. *Complexity:* While the Monte Carlo method has a very general applicability, changing one variable may mean large amounts of recalculation because of the extent of the iterative process when using this model. The complexity reduces the 'transparency' of the method. This may create difficulties in community consultation and risk communication it obscures errors, and creates difficulties for checking by both the modellers and administering authorities.

2. *Loss of factor distinctions:* The method does not indicate 'which variables are the most important contributors to output uncertainty' (US EPA 1992 p. 22928).

3. *Unrealistic probability assessments:* US EPA (1992) notes that simulations such as that found with the Monte Carlo model often 'include low probability estimates at the upper end that are higher than those actually experienced in a given population, due to improbability of finding these exposures or doses in a specific population of limited size, or due to non-obvious correlations among parameters at the high ends of their ranges'. This results in overestimation of exposure dose or risk. The Science Advisory Board of the US EPA has noted: 'For large populations, simulated exposures, doses and risks above the 99.9 percentile may not be meaningful when unbounded log-normal distributions are used as a default' (US EPA 1992 p. 22922).

4. *Assessment endpoints:* With Monte Carlo-type assessments there is still a need to determine what is an acceptable level of exposure. Smith (1994 p. 48) considers that 'the level of exposure exceeded by 1 in 20

exposed persons would seem to be an appropriate reasonable maximum'. This would allow 5 per cent of the population not to be included in the exposure assessment.

5. *Variability–uncertainty confusion:* Smith (1994) highlights the need to distinguish between 'variability' (measurable factors that differ across populations such as height) and 'uncertainty' (unknown, difficult to measure factors such as frequency of trespassing on a site). Currently available software packages do not distinguish between variability and uncertainty. An administrator reviewing a Monte Carlo risk assessment will, however, need to appreciate the differences between variability and uncertainty and the nature and extent of both (Smith 1994). More recent developments of Monte Carlo analysis, such as two-dimensional simulations, may assist with differentiating uncertainty and variability in the input parameters (Simon, 1999).

6. *Limited exposure data:* Limited information is available about many variables for the exposure assessments. As a consequence of this, many input variables are described as triangular distributions. Smith (1994) stresses the need to collect and verify distributions from many currently undescribed input assumptions to improve accuracy. The use of Monte Carlo methods may be inappropriate where the predictions of exposure are so dominated by uncertainties. McKone (1994) gives the example of benzo(a)pyrene, where information on benzo(a)pyrene exposure is 'not readily available' so that the use of Monte Carlo methods to assess variability in population exposures is somewhat redundant.

7. *Simplification of complex situations:* Exposure assessments are comprised of combinations of modelling, sampling, and modelling/sampling combinations (McKone 1994). Even the use of complex models still provides a static picture of a dynamic world albeit a more elaborate representation of reality (McKone 1994) and such a picture must be placed within a sound theoretical framework.

8. *Misleading precision:* The use of more complex models does not necessarily increase precision (McKone 1994). The costs of collecting and analysing data, and constructing new models must be balanced by the value of the information obtained. There is a need to appraise the value of information along with its uncertainties in defining the capabilities and limits of exposure models (Langley & Sabordo 1996).

9. *Characterisation of extreme values:* The 50th percentile can always be estimated with less uncertainty than the 99th percentile (Finley et al. 1994). Problems in estimating the extreme percentiles can come from limitations in the measurement techniques (e.g. incorrect and implausible estimates of dietary consumption may be accepted into the survey) the duration over which exposure data was collected (see short-term and long-term variation below) and whether there are sub-populations that may have unusual exposures (e.g. vegetarians, subsistence fishers) (Finley et al. 1994). Estimating extreme percentiles can be a very time-consuming process.

13.2.2 Monte Carlo versus Latin hypercube

Monte Carlo uses 'random (or pseudo-random) numbers to sample from the input distribution ... [so that] ... samples are more likely to be drawn from values that have higher probabilities (e.g. near the mode)' (AIHC 1994 p. 3.3). This could be important if there is concern about exposures represented by the tails of the distributions (e.g. 99th percentile exposures). Large numbers of iterations are required to overcome this. It is more likely to result in unduly frequent combinations of model exposure scenarios. Latin hypercube techniques use random sampling within equiprobable intervals of the distribution so that there will not be clustered sampling near the mode. It also 'maintains complete independence of the variables' (AIHC 1994 p. 3.3). but this also means if correlations are intended between variables appropriate mathematical actions must be taken.

13.2.3 Estimating distributions for exposure factors

Factors affect the choice of distributions (Finley et al. 1994) include:

- *Variability and uncertainty:* Variability, as an inherent characteristic of a population, will not be reduced with additional data but will be more accurately characterised. Uncertainty, however, will be reduced with additional data.

  Uncertainty may arise from factors intrinsic to the available data (e.g. limitations of study design and analytical techniques) or from the application of data to non-sampled populations (e.g. extrapolating Scandinavian data to an Australian population) (Finley et al. 1994).

  The characterisation of uncertainty related to exposure factors has been developed further than two other areas of uncertainty that may in fact be more significant: the relationship between the absorbed dose and the ultimate delivered dose to a target organ and the uncertainty about the response to the dose (Finley et al. 1994).

- *Factor interdependence:* Some factors, such as body weight and skin surface area, are interdependent, and this needs to be considered. Age-specific data should be used, as the factor may be strongly related to age (e.g. inhalation rates).

- *Short-term and long-term variation:* Interpersonal variability will be decreased if the length of time over which a factor is measured is increased. Short-term data tends to overestimate inter-individual variation (Finley et al. 1994). For example, the 95th percentile of dietary intakes from studies taken over one- to three-day periods will be significantly higher than for studies taken over longer periods such as 1 month to 1 year. This has been seen in the studies of tap water consumption and fish consumption (Finley et al. 1994). It can be particularly marked for rare exposures (e.g. rarely eaten food such as shellfish).

  Studies of shellfish consumption taken over short periods of time may suggest only a very small proportion of the population consumes the foodstuff and, if the common practice of excluding all non-consumers is undertaken, there will be a poor characterisation of the variability in the general population (Finley et al. 1994). Short-term data tends to overestimate inter-individual variation.

- *Parametric versus non-parametric distribution characterisation:* For data to meet parametric distributions (e.g. normal or log-normal), appropriate statistical tests must be met. Theoretically normal or log-normal density distributions do not have an upper-bound limit yet for many factors (e.g. height, weight, fluid consumption) there are obviously physiological limitations to the factors. Some of the currently available software enables such factors to be set within the model.

- *Shapes of distributions:* Triangular shape distributions are often used in Monte Carlo-type assessments but may be viewed as conservative characterisations of truncated normal or log-normal distributions (Finley et al. 1994).

13.2.4 Selecting appropriate datasets

When establishing probability distributions, the distributions should be determined, where possible, from relevant datasets. If there is a need to estimate a probability distribution, it should be appreciated that that many environmental health factors are likely to be log-normally distributed rather than symmetrically distributed. Examples of risk variables that have been characterised by log-normal distributions are (Murphy 1998):

- body weight (each sex)
- bioaccumulation
- breathing rate
- cancer potency factors
- concentrations in air, soil, tissue, water
- drinking-water rate
- exposed skin
- fish consumption
- lifetime
- residence time
- shower duration
- shower water use
- soil ingestion rate
- surface area/body weight
- total water use
- toxic susceptibility.

Much environmental data is log-normally rather than normally distributed. Table 21 gives some examples of output variables that can be represented by probability distributions.

For describing a probability distribution, the relevant studies and the quality of the data produced may vary considerably. Unless datasets are rigorously scrutinised the resulting uncertainty in the range of risk estimates could be greater than obtained using point estimates (Finley et al. 1994).

Finley et al. (1994) recommend the following criteria for assessing data:

- consistency with other studies
- relevance of the survey population to the general population or the population being appraised as part of a risk assessment
- minimisation of confounding variables
- whether there is sufficient data to adequately characterise variability and the extremes of the distribution.

Haimes et al. (1994) propose several approaches to developing distributions when objective data is missing or scarce or not quite relevant:

- When data is sparse but relevant, expert judgement can be used to propose percentiles using available data as 'collaborators' of the expert judgement.
- Where data is not quite relevant to propose a distribution for a parameter, expert judgement again can be used collaborating the judgement with analogous data.
- Where there is an absence of data the formal elicitation of expert judgement to construct a distribution can be used.

If there are a variety of studies then the purposes, designs and methodologies that are similar maybe able to be combined (Finley et al. 1994).

Haimes et al. (1994) highlight the need to examine the tails of probability distribution functions and submit them to a 'reality check' and examine the combination of factors that resulted in the extreme values. They highlight the extreme sensitivity of these tail values to assumptions and reinforce the need to assess the sensitivity of the tails to the assumptions. The assumptions need to be examined as to whether they are mutually consistent.

Table 21: Some key variables for which probability distributions might be needed

Model component	Output variable	Independent parameter variable
Transport	Air concentration	Chemical emission rate Stack exit temperature Stack exit velocity Mixing heights
	Meteorological factors	Wind speed Wind direction
Deposition	Deposition rate	Dry-deposition velocity Wet-deposition velocity Fraction of time with rain
Overland	Surface-water load	Fraction of chemical in overland runoff
Water	Surface-water concentration	River discharge Chemical decay coefficient in river Mixing depth
Groundwater	Groundwater concentration	Predictions of plumes
Soil	Surface-soil concentration	Surface-soil depth Exposure duration Exposure period Cation-exchange capacity Decay coefficient in soil
Food chain	Concentration in animal products	Soil ingestion rates Plant to animal bioconcentration factors
	Plant concentration	Plant interception fraction Weathering elimination rate Crop density Soil-to-plant bioconcentration factor
	Fish concentration	Water-to-fish bioconcentration factor
Dose	Inhalation dose	Inhalation rate Body weight
	Ingestion dose	Plant ingestion rate Soil ingestion rate Body weight
	Dermal absorption dose	Exposed skin surface area Soil absorption factor Exposure frequency Body weight

Adapted from: NRC 1994, p. 169; Seigneur et al. 1992.

13.2.5 Monte Carlo exposure assessment datasets

Increasing amounts of data are becoming available to enable the use of Monte Carlo-type assessments.

The *Exposure factors sourcebook* (AIHC Taskforce 1994) presents extensive documentation of probability distributions from a variety of sources intended for the US population. Descriptions of the probability distributions are provided to enable easy use of @ RISK® software.

The probability distributions for soil ingestion by children use the studies of Calabrese and Stanek (1991) and Burmaster et al. (1991). These distributions vary markedly and the updated values of Calabrese and Stanek (1995) are not included.

13.2.6 Principles for the use of Monte Carlo-type techniques

The purpose and scope of the risk assessment should be clearly articulated in the issues identification section. Burmaster and Anderson (1994) stress that any method of exposure assessment must have a clearly defined assessment endpoint and provide all relevant information so that the assessment can be reproduced and evaluated. Burmaster and Anderson (1994) detail 1-14 principles for good practice in Monte Carlo assessments.

These are as follows:

- Detail all formulae.
- Detail point estimates of exposure where these are demanded by regulatory agencies.
- Detail sensitivity analyses to enable the identification of relevant and important input variables. Those variables that will drive risk assessment must obviously be included in the Monte Carlo analysis but reasons for excluding insignificant variables must also be detailed.
- Use probabilistic techniques (which may be demanding in terms of time, money and other resources) only where exposure pathways are likely to be significant.
- Provide detailed information about input distributions. The minimum stated by Burmaster and Anderson is:
  - a graph showing the full distribution and the location of the point value used in the [point estimate] risk assessment
  - a table showing the mean, standard deviation, the minimum (if one exists), the 5th percentile, the median, the 95th percentile and the maximum (if one exists) (p. 478). There needs to be a sufficient justification of the selected distribution, which should be based on adequately referenced sources and the statistical, physical, chemical, and biological mechanisms relevant to the distribution.
- Detail how the input distributions capture and represent both the variability and the uncertainty in the input variables (p. 478) so as to enable both variability and uncertainty to be described and analysed separately.
- Use measured data to test the relevance of the input distribution to the population, place and time of the exposure assessment. Further data may need to be gathered to supply missing information or supplement incomplete information.
- Describe the methods by which measured data was used to derive a probability distribution.
- Detail any correlations between data where there are relatively high correlations. Sensitivity analysis may be necessary to determine the effects of correlations between variables on the exposure analysis.
- Provide detailed information and graphs for each output distribution. Burmaster and Anderson suggest the following as a minimum:
  - a graph of the variable with administratively set allowable risk criteria as annotations and point estimates of risk using the administratively set point estimates of exposure
  - a table of the mean, the standard deviation, the minimum (if one exists), the 5th percentile, the median, the 95th percentile, and the maximum (if one exists).
- Provide records of sensitivity analyses and their impact that will enable the determination of the most important input variables (or groups of variables).
- Assess the numerical stability of the central moments (mean, standard deviation, skewness and kurtosis) and the tails of the output distributions. The latter are particularly sensitive to the nature of the tails of the input distributions and, as they stabilise very slowly, sufficient iterations are required to demonstrate the numerical stability. Burmaster and Anderson suggest that commonly more than 10,000 iterations are required. Software that enables Latin hypercube sampling results in more rapid stability of these output tails. Burmaster and Anderson state that the changes in the tails of only a few input distributions contribute strongly to changes in the upper tail of the output distribution.
- Detail the name and statistical quality of the random number generator used. Some generators are inadequate because of short recurrence periods.
- Interpret the results and detail the limitations of the methodology such as the effects of biases not elsewhere interpreted.

Burmaster and Anderson state that 'the principles are not mutually exclusive nor collectively exhaustive' (Langley & Sabordo 1996 pp. 140–141).

13.2.7
Administrative requirements for the use of Monte Carlo methods

The range of total acceptable exposures and risk will need to be defined on a situation-specific basis after consultation with stakeholders. Depending on how it is applied, the Monte Carlo method may lose much of the conservatism usually inherent in point estimates.

Regulatory authorities in Australia are likely to require the following of assessments using Monte Carlo methods:

- meeting the 14 principles of good practice detailed above
- providing adequate information to the authority to enable review of the assessment – this may require providing the software (and underlying formulae) and data
- a demonstration of the relevance of the exposure data to the site (data from other countries or cultural backgrounds may not be relevant)
- an explanation of the data and method that will be able to be understood by the relevant community (usually the most difficult aspect)
- using data that accounts for age and gender differences and takes into account susceptible populations.

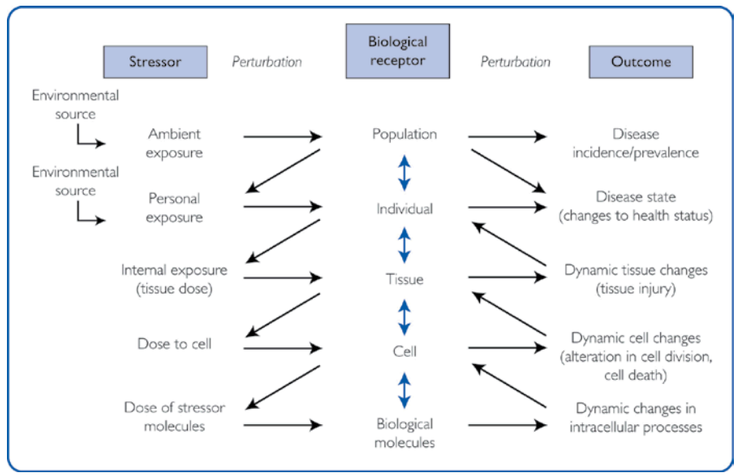
On a large site divided into housing lots, the results for specific housing lots that may be affected by atypically elevated concentrations should not be obscured by averaging or Monte Carlo techniques applied to the entire site. In many instances, Monte Carlo methods will only be relevant to large sites or sites where direct measurements of exposure are not practicable. Before the use of Monte Carlo is commenced for any situation being assessed, the assessor should check with the relevant regulator or government authority about whether such use is considered appropriate. Most regulators are likely to discourage the use of this technique, in the main due to the difficulty in explaining it to the affected community and the lack of robust probability distributions for parameters of interest.

Since the outputs of Monte Carlo analyses are distributions of risk estimates and other parameters, some guidance will be needed on where to define the cut-offs for risk assessment purposes. This is likely to fall into the realm of policy settings to be determined by government authorities. As noted above, UK guidance on establishing guidance values (GVs) for contaminated land exposure assessment (CLEA) is already showing signs of 'back-pedalling' on the use of probabilistic approaches, such as Monte Carlo analysis.

13.3 INTEGRATION OF EXPOSURE WITH EHRA OUTCOMES

As part of the NRC review of toxicity testing developments for the 21st century (NRC 2007), Hubal (2009) commented on the role that developments in exposure sciences must play in developing new paradigms of EHRA. In particular, the development of models that could be used to define exposures at levels ranging from environmental to cellular (target tissue doses) would be important for integrating animal testing data with the new generation of scientific tools using genetic, *in vitro* and *in silico* techniques for profiling chemical toxicity and individual susceptibility. A depiction of the interrelationships in such a model is shown in Figure 31.

Figure 31: Proposals for integrating exposure with outcomes of EHRA



Hubel 2009. Reproduced with permission of Oxford University Press.

Chapter 14: Biomonitoring

14.1 BIOMARKERS

The term 'biomarker' has been used in recent times to describe the measurements used in biological monitoring. The term refers broadly to almost any measurement reflecting an interaction between a biological system and an environmental agent, which may be chemical, physical or biological (WHO 1993b, 2001). Three classes of biomarker are identified by WHO (2001):

- *Biomarker of exposure:* an exogenous substance or its metabolite or the product of an interaction between a xenobiotic agent and some target molecule or cell that is measured in a compartment within an organism
- *Biomarker of effect:* a measurable biochemical, physiological, behavioural or other alteration within an organism that, depending upon the magnitude, can be recognised as associated with an established or possible health impairment or disease
- *Biomarker of susceptibility:* an indicator of an inherent or acquired ability of an organism to respond to the challenge of exposure to a specific xenobiotic substance.

For many environmental pollutants, the flow of events between exposure and health effects is not well understood. Biomarkers help address this problem by improving the sensitivity, specificity and predictive value of detection and quantification of adverse effects at low dose and early exposure (Fowle 1989; Fowle & Sexton 1992; NRC 1992). Sensitive sub-populations can be better pinpointed by biomarkers that measure increased absorption rate or a more severe biological response to a given environmental exposure (Fowle & Sexton 1992; Hemminki 1992; Lauwerys 1984; NRC 1992).

14.1.1
Why biomonitoring?

Biological monitoring is a measuring procedure whereby validated indicators of the uptake of contaminants, or their metabolites, and people's individual responses are determined and interpreted. Whereas environmental monitoring measures the composition of the external environment around a person, biological monitoring measures the amount of contaminant absorbed into the body.

Biological monitoring may be direct (e.g. the measurement of lead in blood) or indirect (e.g. the measurement of the breakdown product of nicotine and cotinine in urine). Biological monitoring may measure a biological effect, such as enzyme depression, or a physiological effect, such as tremor. The monitoring may be used to identify whether exposure has occurred at all, or the amount of exposure.

If biological monitoring is practicable, it will be more valuable than environmental monitoring in determining the level of risk from an environment, as it will measure whether exposure is occurring and the level of exposure (Langley 1991b). It can be useful in identifying highly exposed individuals or sub-populations.

The prerequisites for biological monitoring (Aitio et al. 1988) are as follows:

- The substance and/or metabolites need to be present in a tissue, body fluid or excretion suitable for sampling.
- Valid, accurate and practicable methods of sampling and analysis are available.
- The results of testing can be interpreted in a meaningful way for individuals and groups.
- An appropriate management strategy has been devised for sampling, analysis, collation of results, interpretation of results, and follow-up.

The use of biomonitoring data in environmental risk assessment was reviewed at an international biomonitoring workshop in 2004, at which six case studies illustrated the applications and utilities of this technique in environmental health surveillance (Albertini et al. 2006).

Further reviews of the application of biomonitoring to risk assessment have been presented by Doerrer (2007), Angerer et al. (2006) and Swenberg et al. (2008).

One of the difficulties traditionally associated with the interpretation of biomonitoring data has been the absence of validated values representing specific levels of exposure or linking to levels of effect. This issue was addressed by an international panel convened to develop a series of biomonitoring equivalents (BEs) (Hays et al. 2008). This panel established some guidelines on what should be taken into consideration in establishing BEs, including consideration of toxicokinetics and internal dose metrics, integration of human and animal data, and the choice of suitable tissues and analytes. This expert group devised a series of flow charts (Figure 32) illustrating how animal and human data could be integrated, depending on the extent to which pharmacokinetic data in either species are well understood.

Biological monitoring should not be commenced before:

- the objective of the biological monitoring is clearly defined
- a reference range of results that is applicable for the population under study is established – this is often not available (or a control group is not available to establish a reference range); the relationship of body burden levels and exposure (or risk) are unavailable for many substances
- consideration has been given as to how results are to be managed – significant anxiety may be caused by factors such as delays in providing information and

an inability to explain the meaning of measured levels or to take action if the person is distressed by elevated levels, perceives that any measure of exposure is unsatisfactory or equates exposure to a health effect may cause

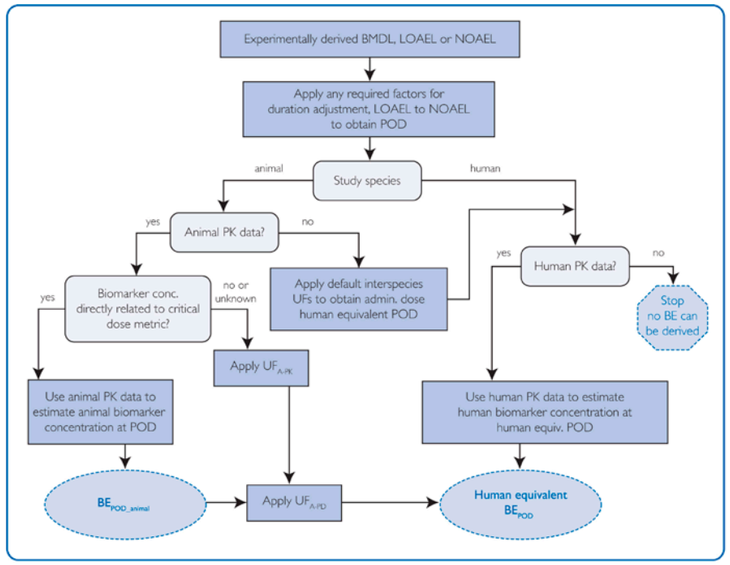
- the correct timing of sampling has been established – correct timing is critical for substances with short biological half-lives or a particular exposure is of concern
- a process has been established to enable consistent analysis and epidemiological appraisal of results
- the ethical and confidentiality aspects of collecting, maintaining and distributing information and results are fully considered
- a centralised collection point for results has been established to enable consistent analysis and epidemiological appraisal of results.

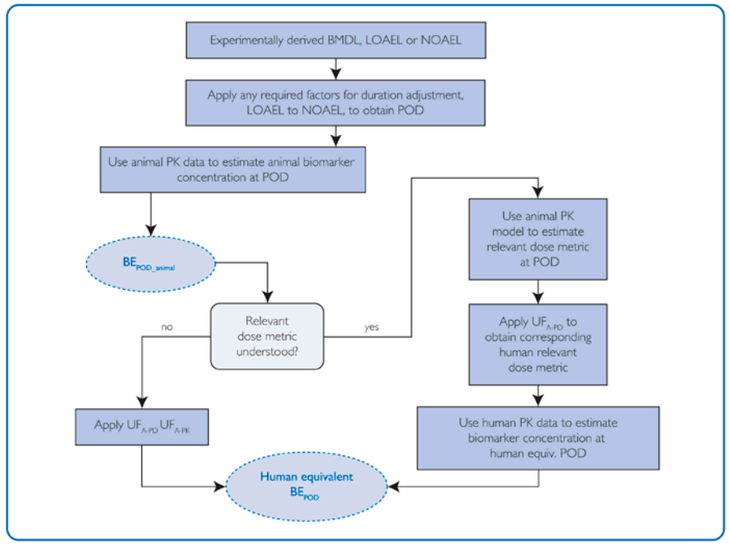
The reasons for biological monitoring include to:

- detect whether exposure has occurred
- quantitate exposure
- enable the risk of health effects to be assessed
- determine changes in exposure over time or to assess the effects of interventions such as health education or soil remediation (if serial measurements are done) or to determine exposure pathways and their relative importance, such as occupational versus domestic exposures ingestion of soil versus inhalation of dust (if combined with environmental monitoring)
- determine segments of the population at greatest risk, such as particular age groups or those living in particular locations or circumstances (if conducted as part of widespread population studies).

Results should always be available to participants in biological monitoring combined with a meaningful explanation of the results.

Figure 32: Flow charts for deriving BEs for chemicals pharmacokinetic data are available for both animals and humans along with toxicity data from both species (a) or only from animals (b)





Reproduced from Hays et al. 2008 ©2008 with permission from Elsevier.

Several aspects must be considered:

- A good biological monitoring test result may not correlate well with environmental levels (mainly because of human factors).
- The number of substances that can be used reliably for biological monitoring is still small.
- Irritative, locally or rapidly acting substances are usually unsuitable as the systemic absorption may be minimal and/or irrelevant to the level of local reaction (e.g. SO_2, ammonia, direct skin exposure to PAHs causing skin cancer).
- The substance must be in some tissue or fluid suitable for sampling.
- Accurate, valid and practical measuring methods must be available.
- The result should be interpretable in terms of health risk.
- The results are likely to have more value for a group than an individual.

The advantages of biological monitoring are:

- the exposed person is his or her own sampler, so that many 'samples' are taken over a 100 per cent sampling time
- the evaluation of absorption can be performed over a prolonged period of time
- the sampling takes into account all the person's movements within and outside the domestic environment, and accidental and illicit exposures
- the amount absorbed by various routes is considered (not only via the respiratory route as is presumed by monitoring of atmospheric concentrations), for example, oral absorption of lead compounds or in situations where skin absorption is important
- it may show exposures where past environmental monitoring is unavailable, for example, PCBs where the persistence of the substances acts as a long-term marker of exposure
- it enables an individualised assessment taking into consideration age, sex, personal hygiene, biotransformation and elimination.

The disadvantages and difficulties are:

- the relatively wide range of individual response to a substance and the wide 'normal' range that may have to be considered
- the lack of simple specific analytical methods of sufficient sensitivity (in many instances)
- difficulties in sample collection, for example, 24-hour urine collections
- unsuspected exposure can be shown but the source cannot be pinpointed – this will require detailed environmental monitoring
- inferences caused by occupational exposure, for example, lead exposure in battery makers and radiator repairers
- there must be a clear relationship (if only on a group basis) between the chosen biological indicator and the health risks of the substance.

Some analyses require specialised laboratories:

- There may be laboratory inaccuracy.
- If the substance has a short biological half-life, rapidly changing concentrations in body samples complicate interpretation and the body burden may be under-predicted or over-predicted.
- Transient periods of high exposure may not be detected.

Having decided a test for a substance is appropriate, further questions arise:

- Which compound should be measured? The substance, a metabolite or both?
- Which biological fluid or tissue is to be sampled?
- In relation to what period of exposure?
- How frequently should sampling be done?

14.2 CHOICE OF TISSUE OR FLUID

The biological samples used for monitoring may be (Fao & Allesio 1983):

- blood, urine, fat, saliva, sweat, faeces
- hair, nails, teeth
- expired air.

Physiological response to the exposure may be estimated by determining changes in:

- the amount of a critical biochemical constituent
- the activity of a specific enzyme
- a particular physiological function such as lung function.

Choice of biological tissue or fluid for a hypothetical substance is represented in Figure 33.

Figure 33: Choice of biological tissue or fluid for a hypothetical substance

Tissue or target organ	Increasing	Decrease in
Blood which has just perfused target organ Mixed venous blood Urine Exhaled air	convenience, accessibility and co-operation from subject ↓	potentilal validity

14.2.1 Blood

Depending on the biological half-life of a substance, blood analysis may provide an indication of exposure from recent hours to several years. Levels are often transient if the half-life is not prolonged. The process of blood-taking may be unacceptable for some people, including children.

When the volume of distribution is high, concentrations in blood are often too low to be measured. Samples may require careful procedures, such as plasma separation and freezing. Substances measurable in the plasma may not be responsible for the toxic effect which, instead, arises from a metabolite.

14.2.2 Urine

Only a limited number of substances can be measured in urine because of degradation of the parent substance to breakdown products. Urine samples, in general, provide a more integrated assessment of exposure than blood for periods of recent hours or days. Twenty-four-hour sample collections may be more appropriate than spot samples but many people find these collections onerous. First morning urine samples have been found to be effective for representing 24 hour urine samples. (Froese et al. 2002, Bader et al. 2004, Zhang et al. 2009). Urine samples require rapid processing and cooling.

14.2.3 Hair and toenails

Hair and toenails can provide an integrated measure of exposure over a more prolonged period than blood or urine. They are only useful for chemicals known to accumulate in those tissues and they are inappropriate tissues for biological monitoring on or near contaminated environments. External contamination of the hair cannot be adequately removed during sample preparation and an accurate measure of excretion via hair cannot be performed. Hair analysis may be useful for assessing intake from purely dietary sources when there is no general environmental contamination.

14.2.4 Breast milk

Collecting breast milk is usually easy and acceptable to nursing mothers. Breast milk provides an integrated exposure for very lipid soluble compounds for time periods related to the biological half-life of the substance. Breast milk measurements of PCBs, organochlorine pesticides and dioxins have been used for exposure assessments. The concentrations must be standardised for fat content and may vary according to the period since breastfeeding first commenced.

14.2.5 Expired air

Expired air is used to determine exposures to ethanol (e.g. traffic breathalyser) and some solvents and can be correlated to blood concentrations based on the Henry's Law constant of the substance being measured.

14.3 CHOICE OF A TEST

Optimally, a biological monitoring test would give a result that reflected the exposure, the concentration of the substance in the target organ and the risks of adverse effects (Friberg 1985). Few tests are available that approach this ideal (Langley 1991a). Furthermore, what is of most importance is 'not so much the choice of medical test as much as the way the testing program is organised, the way the results are evaluated and communicated, and the way abnormalities are pursued' (Silverstein 1990).

In Australia, exposures from contaminated soil for example will be generally low, creating problems in accurate measurement at low levels and the possibility of results being overwhelmingly influenced by other sources of exposure (e.g. the influence of cadmium in food, tobacco smoke and the occupational environment will generally be far greater than the influence of cadmium contamination of soils).

For many substances, biological monitoring is impracticable because:

- analytical techniques are not available or are inaccurate at low levels or in the tissues or fluids being tested
- insufficient information is available on inter- and intra-individual toxicokinetics and thresholds of health effects to enable risk assessment of results
- insufficient epidemiological studies have been done to determine normal ranges.

The correct choice of biological tissue or fluid is important. Rarely can the concentration in the critical organ be measured and compared with concentrations that give rise to effects.

Attempts have been made for such direct measurement, for example, *in vivo* neutron activation analysis can directly measure renal or liver concentrations of cadmium but requires specialised equipment and provides a dose of ionising radiation to the subject.

For biological monitoring based on urine analysis, simple measurements of concentrations can provide sufficient information on exposure, but in many instances, measurements of elimination rates provide more precise information. Urinary concentrations related to creatinine, or urinary flow rates may provide more accurate information, but creatinine has not been found to be worthwhile in some evaluations (Zhang et al. 2009)

Substances for which biological monitoring of general environmental exposures is practicable are detailed in Table 22.

Table 22: Substances likely to be suitable for biological monitoring

Substance	Fluid or tissue	Comments
Lead	Blood	Urinary lead does not accurately reflect either recent exposures or burden. Substantial data available on level of risk for particular blood lead ranges. Numerous Australian studies provide comparison data. levels of concern available for both general population and groups (for example, children).
Cadmium	Urine or blood	Urinary levels tend to reflect body burden; blood levels reflect recent exposures. Urinary levels need to be adjusted for changes in urinary flow rates (results often given as µg Cd/g creatinine or µg Cd/24 hour). Laboratory inaccuracy has always been a major problem, particularly prior to 1980. Limited Australian studies to provide comparison data. Most international studies have concentrated on occupational exposures. Very limited data on children, especially for those less than 5 years. WHO (cited in Mueller et al. 1989) has set levels of concern. General diet and smoking will tend to have a major influence on levels.
Arsenic	Urine	Short biological half-life; study must be done during exposure (or at most within 1–2 days afterwards). Considerable interference from organic sources of arsenic (for example, seafood). Dietary sources from the environment not under study need to be excluded and testing for inorganic arsenic undertaken. Limited comparison data and no set levels of concern.
Mercury	Blood or urine	At equilibrium, the concentration of mercury in the blood reflects daily intake and is probably the best indicator of exposure. Total measured mercury will also include methyl mercury from fish, so that a fractionated analysis of mercury salts and alkylated mercury compounds may be required (Aitio et al. 1988). Methyl mercury exposure will not affect urinary mercury levels although urinary levels show significant diurnal variation. Some international comparison data is available.
Polychlorinated biphenyls (PCBs)	Blood, adipose tissue (fat), breast milk	Long biological half-life so that historical exposures (i.e. body burden) may be able to be monitored. Different PCBs will have different behaviours in the body and different biological half-lives. Some comparison data is available. It is difficult to obtain adipose tissue samples and blood sampling is usually preferred.
Organochlorine (OC) pesticides	Blood, adipose tissue (fat), breast milk	Long biological half-life so that body burden can be assessed. Some comparison data is available, especially for blood. It is difficult to obtain adipose tissue samples and blood sampling is usually preferred.
Organophosphonate (OP) pesticides	Blood	Plasma butyrylcholinesterase or erythrocyte acetylcholinesterase (AChE) may be monitored to assess recent exposures. Depressed AChE activity may better reflect a level where a physiological response may occur. Wide range of values reflect 'normality', so individual baseline values assist interpretation.

Adapted from: Langley (1991b).

There are a range of other substances for which biological monitoring may be available – the tests should be assessed and used on their individual merits for a particular situation. Biological monitoring has been applied to a range of situations: tobacco use (polycyclic aromatic hydrocarbons, aromatic amines and specific nitrosamines), dietary exposures (e.g. aflatoxins, N-nitrosamines, heterocyclic amines), medicinal exposures (e.g. cisplatin, alkylating agents, 8-methoxypsoralen, ultraviolet photoproducts), trichloroacetic acid for chlorinated disinfection by-products in drinking water and occupational exposures (e.g. benzene, ethylene oxide, styrene oxide, vinyl chloride, aromatic amines, polycyclic aromatic hydrocarbons).

Besides the pesticides mentioned in Table 22 specialised tests may be available from some laboratories for pesticides such as glyphosate.

Most organic contaminants are not amenable to biological monitoring in general environmental situations because of the low levels of exposure and the lack of comparison data compared with occupational situations. Specialised studies may make biological monitoring for some inorganic substances practicable (e.g. manganese, radioactive isotopes).

A good knowledge of the toxicokinetics of a substance is required for the correct choice of method and interpretation of results. The duration of persistence of the agent will be important as is the volume of distribution (e.g. many very lipid soluble substances with a very high volume of distribution have such low blood levels that they can't be measured in blood but can be identified in breast milk). Individual results may be distorted if there is not constant exposure or equilibrium within the body (Langley et al. 1998).

Cytogenetic testing may occasionally be of value but is often difficult to interpret as only small numbers of cells are usually examined so that there is the potential for considerable confidence limits around the results and because there can rarely be a link made to specific agent (one exception is aflatoxin). Tests such as sister chromatid exchange and micronuclei are non-specific tests. There are problems with confounding, distinguishing recent from historical exposures, quantifying exposures and dealing with a finite background incidence of chromosomal abnormalities.

Under the National Model Regulations for the Control of Workplace Hazardous Substances (adopted by the states and territories), health surveillance is required for specified substances. Biological monitoring methods developed for some of these methods are detailed in the NOHSC series *Guidelines for health surveillance*.

14.3.1 Accuracy

Laboratory accuracy has always been a problem because of the low levels of the substance being tested and analytical problems, including those caused by the biological matrix and the risk of contamination. Gross analytical errors have occurred in the measurement of blood lead, and blood and urinary cadmium (Elinder 1985; Vahter 1982). Friberg (1985) reports that 'normal' values for aluminium in plasma and serum 'decreased' during the 10 years 1975–1985 'from several 100 µg/l to a few micrograms the only reason for this being improved analytical technique'. Aitio et al. (1988) provide a further example for the values regarded as normal average serum chromium concentrations for occupationally unexposed men. Papers published between 1956 and 1984 showed a decrease in 'normal' values from 3,600 mmol/l to 2.1 mmol/l; Aitio et al. (1988) attributed the decline to better techniques that avoided chromium contamination.

Aberrant results may need to be repeated before being accepted as 'high'. Choice of a laboratory should be governed by the presence of stringent internal and external quality control measures.

Contamination during sample collection is likely to be a significant problem unless specialised collection protocols are rigorously followed. One example is skin contamination affecting blood samples (especially capillary prick samples).

Twenty-four-hour urinary collections are likely to be impracticable during general community studies and present significant risks for contamination during collection. Inappropriate sample containers can be a significant source of inaccuracy from leaching or contamination. Without appropriate selection of containers and storage conditions, some heavy metals will adsorb to some container materials giving falsely low readings. A single laboratory is preferred for studies to minimise problems arising from inter-laboratory variations and to enable a single body of data.

14.3.2 Indicator analytes

Where there are multiple contaminants uniformly distributed in the environment and with similar environmental and biological behaviour, the measurement of one contaminant (the indicator analyte) may be a surrogate measure for other contaminants. The indicator analyte may be chosen for the ease (or accuracy) of analysis or its toxicity relative to the other contaminants. For example, if lead and cadmium are uniformly present, lead may be chosen for the ease and relative accuracy of analysis as well as the availability of levels of concern and comparison data. Alternatively, lead may also be chosen because it is the predominant contaminant. In such instances, if the blood lead results are not elevated, elevated levels of cadmium would not be expected. If high blood lead results are demonstrated, cadmium levels may need to be assessed to determine whether there may also be a significant risk from cadmium exposure.

14.4 INFLUENCES ON BIOLOGICAL MONITORING RESULTS

Factors apart from environmental contamination to be considered in interpreting biological monitoring results include (American Conference of Governmental Industrial Hygienists – ACGIH 1990):

- changes induced by strenuous physical activity
- changes induced by environmental conditions (including heat, diet and cigarette smoking)
- changes induced by water intake
- changes in physiological functions induced by pregnancy, disease or diurnal rhythms
- changes in metabolism induced by congenital variations of metabolic pathways or induced by simultaneous administration of another chemical (induction or inhibition of activity of a critical enzyme by medication or by pre-exposure or co-exposure to another chemical).

14.5 EXPOSURE AND BIOLOGICAL MONITORING RESULTS

For toxicokinetic reasons, the relationship between exposure and biological monitoring results is often not linear. For example, with air lead levels there appears to be a greater influence on the rate of change of blood lead levels with changes at lower air lead levels than moderate air lead levels (Friberg 1985).

This is one of the reasons why monitoring blood lead is a more common approach to lead EHRA and risk management (see Section 14.8).

The physico-chemical properties of the contaminant will have a crucial influence on the bioavailability of the contaminant and hence biological monitoring results. A further crucial influence will be the characteristics of the exposed population (e.g. age, behaviours).

The physico-chemical properties of the contaminant and the characteristics of the exposed population usually will be more important predictors of biological monitoring results than a statement of the concentration of the contaminant in the soil.

14.6 ABNORMAL RESULTS

If the accuracy of an abnormal result can be confirmed (this may require repeat testing), the health risks should be assessed and medical assessment may be required. The reason for the high result should be determined, that is, the relevant exposure pathways.

There should be a clear understanding of the basis of how the 'normal' range was derived. (e.g. What populations were studied? Were they comparable to this population?). If the range is derived from normally distributed results in a general population survey and the range is two standard deviations each side of the mean, 5 per cent of this population will have 'abnormal' results. If results are being compared with health standards, how were these standards set? Do the standards incorporate a safety factor and, if so, how large is that safety factor?

14.7 HEALTH MONITORING

Health monitoring is the organised medical assessment of individuals and groups of people. The medical assessment will consist of history taking and clinical examination, and, where indicated, particular tests (e.g. lung function testing where there is a concern about the effect of air pollutant). The epidemiological aspects of health surveys are covered in Chapter 10.

In Australia, health effects are likely to be found in only a limited number of situations of environmental contamination. Subtle effects may only be able to be determined on a group basis rather than on an individual basis (e.g. subtle neurodevelopmental effects determined by sophisticated testing in groups of children with different lead exposures). Similar problems of causation relating to individual findings rather than group findings arise if the putative effects are common in the general population (e.g. headache or fatigue). Health effects are rarely as specific to an exposure as chloracne with PCB or dioxin exposure.

Health monitoring for specific health effects is warranted where environmental or biological monitoring has indicated a significant risk of effects (e.g. specific tests of renal function if urinary cadmium levels above the levels of concern are detected in biological monitoring).

When health monitoring is done, it should rarely be done in isolation from environmental and/or biological monitoring. Clearly defined health effects should be sought with specific case-definition criteria. Records of other symptoms and clinical findings should also be kept to enable epidemiological assessment of other potential health effects (Langley 1991a).

Before health monitoring is undertaken, the following issues should be considered:

- how to ensure all parties involved do not have unreasonable expectations about the ability of health monitoring to resolve issues of causation or to detect any subtle effect (the studies rarely provide such evidence because of their size and biases)
- confidentiality of information
- how and when information will be made available to participants (the information must be released to participants)
- access to information (by whom and through what mechanisms)
- interpretation of information (at an individual and group level and on what evidentiary basis)
- release of findings (which should be at a group rather than individual level for reasons of confidentiality if the results are made public)
- how the information will be used to address the relevant environmental health issues.

14.8 BIOMONITORING AND BLOOD LEAD

Since the absorption and retention of lead from various environmental matrices can be variable, biomonitoring (blood lead levels) has become the method of choice for data inputs into health risk assessments and for managing environmental health risks associated with lead, particularly in children.

14.8.1 Adult lead exposures

These may be estimated using the US EPA adult lead model methodology (US EPA 2003c). This model focuses on adult women and incorporates lead exposure, uptake into the body and biokinetic transfer into the blood and developing foetus.

$$PbB_{adult} = PbB_{background} + \frac{Pb_{intake} \times BKSF \times EF}{AT}$$

PbB_{adult} = total adult blood lead concentration that have site exposures to soil lead (µg/dL)

AT = averaging time (days/year)

EF = exposure frequency (days/year)

$PbB_{background}$ = background adult blood lead concentration (µg/dL)

Pb_{intake} = total lead uptake from all media (g/day)

BKSF = biokinetic slope factor (µg/dL per µg/day)

This approach allows for protection of the most sensitive receptor in the adult scenario, which is an unborn child carried by a pregnant mother, and is the approach taken for determination of the HIL for soil lead in the contaminated sites NEPM review for the adult exposure scenario (NEPC 2010).

14.8.2 Lead exposures in children

The predominant modelling system for assessing lead exposures in children is the Integrated Exposure Uptake Biokinetic Model for Lead in Children (IEUBK model, Version 1.1 Build9, released June 2009. See <www.epa.gov/superfund/lead/products>). The IEUBK model comprises separate components for exposure, absorption and the biokinetic transfer of lead to all tissues of the body and calculates age-specific blood lead concentrations for children aged between 0 and 7 years.

The components of the IEUBK model can be summarised as follows (NEPC 2010):

- The exposure component estimates intake from soil, dust, water, air and food. The estimate is based on data input by the user. The model provides default estimates for circumstances where site-specific information is not available. Where Australian values are available (e.g. lead concentration in drinking water, dietary lead ingestion rates), these should be adopted.
- The uptake component models the process by which the lead intake is transferred to blood plasma. The amount of lead that is taken up is controlled by the bioavailability of the lead, which can be specified separately for soil, water and food.
- The biokinetic component models the balance of lead in the body between uptake and excretion. A central estimate of blood lead concentration is output from this component.
- The variability component applies a log-normal distribution to the output of the biokinetic component using a geometric standard deviation of 1.6. This value is based on empirical studies where blood lead concentrations of young children and environmental lead concentrations were measured. It models the predicted variability likely to apply to the population.
- The model contains 100 variables, of which 46 can be modified by the user. Those that cannot be modified are based on considerable research, and are detailed in the model user guide (US EPA 2007b).

This approach was taken for determination of the HIL for soil lead in the contaminated sites NEPM review for all residential and public health exposure scenarios (NEPC 2010).

Chapter 15: Microbiological risk assessment

Microbiological risk assessment (MRA) is an especially important part of EHRA associated with food and water contamination. Recent Australian initiatives to make more efficient use of scarce water resources through water recycling and stormwater harvesting have placed a particular emphasis on MRA (NRMMC, NHMRC, AHMC 2006; NRMMC, EPHC, NHMRC 2008; NRMMC, EPHC, NHMRC 2009a, b). These initiatives required the development of HRA methodologies that predict disease impacts associated with the pathogens likely to be encountered in such water sources, as well as informing the setting of microbiological standards for reclaimed water. The Phase 1 recycled water guidelines (NRMMC, NHMRC, AHMC 2006) introduced the concepts of quantitative MRA (QMRA), while subsequent documents outlined how these principles could be adopted into water management.

The aim of this chapter is to provide a brief description of the principles of MRA and QMRA in the context of EHRA and to briefly outline the processes. A more detailed description of MRA can be obtained from the cited references.

Appendix 2 expands on the concepts of QMRA in the context of managing the safety of drinking-water supplies. It contains extracts from a document on health-based targets prepared for a 2010 consultation on draft revisions of the ADWG (see <http://www.nhmrc.gov.au/guidelines/consult/consultations/draft_adwg_guidelines.htm>.

15.1 INTRODUCTION

The aim of microbiological risk assessment is to estimate the level of disease associated with a particular pathogen in a given population under a specific set of conditions and for a certain time frame.

There is much support for the application and development of MRA (ACDP 1996). To date, MRA has predominantly been applied to two exposure sources, food and water, and much of the conceptual development of MRA has resulted from the application of MRA to these media.

MRA concepts and methodologies (particularly QMRA) are somewhat less well developed than comparable methodologies in chemical risk assessment. For example, with QMRA, vast datasets need to be developed, modelling needs to be improved (e.g. secondary transmission) and analytical techniques need to be refined. Methods for extrapolation from animals to humans are still being developed, with different approaches being proposed. QMRA does have an advantage over chemical risk assessment in some cases where human volunteers' dose–response data is available. This obviates the need for animal-to-human extrapolation which introduces considerable uncertainty in most cases of chemical risk assessment.

The development of models for MRA have generally been developed based on conventional risk assessment frameworks (e.g. NRC 1983), which incorporate the conventional step-wise processes of hazard identification, exposure assessment, hazard characterisation and risk characterisation.

15.2 DEFINITIONS

MRA has been defined by various scientific organisations/committees as follows.

The Codex Alimentarius Commission (for microbiological hazards in foods):

- A scientifically based process consisting of the following steps hazard identification, exposure assessment, hazard characterisation and risk characterisation (Codex 1999).

The International Life Sciences Institute – Risk Science Institute (in conjunction with the US EPA):

- A process that evaluates the likelihood of human health effects occurring after exposure to a pathogenic micro-organism or to a medium in which pathogens exist (ILSI 2000).

The Advisory Committee on Dangerous Pathogens (UK):

- A formal structured procedure for identifying and characterising microbiological hazard and determining the risk associated with it (ACDP 1996).

The WHO and FAO guidelines on MRA:

- A tool used in the management of the risks posed by food-borne pathogens, including the elaboration of standards for food in international trade. Quantitative MRA is recognised as a more resource-intensive task (WHO 2009).

QMRA has been defined as follows:

- The application of principles of risk assessment to the estimate of consequences from a planned or actual exposure to infectious micro-organisms (Haas et al. 1999).
- Quantitative risk assessment can be either deterministic (meaning single values like means or percentiles are used to describe model variables) or probabilistic (meaning that probability distributions are used to describe model variables) and most of the literature, guidance and the worked examples in QMRA are probabilistic quantitative risk assessments. This approach offers many distinct advantages over deterministic risk assessment (WHO 2009).

15.3 GENERAL PRINCIPLES

The Codex principles of MRA, as applied to food, are listed below. These principles can also be generalised to the other media: water, air, soil and the surfaces of inanimate objects. Most of the principles listed are similar to established risk assessment principles, except for item 7, which is unique to MRA (Codex 1999, 2003).

1. MRA should be soundly based upon science and conducted according to a structured approach that includes hazard identification, hazard characterisation, exposure assessment and risk characterisation.
2. A MRA should clearly state the purpose of the exercise, including the form of risk estimate that will be the output.
3. The conduct of a MRA should be transparent.
4. Any constraints that impact on the risk assessment such as cost, resources or time should be identified and their possible consequences described.
5. The risk estimate should contain a description of uncertainty and where the uncertainty arose during the risk assessment process.
6. Data should be such that uncertainty in the risk estimate can be determined.
7. Data and data collection systems should, as far as possible, be of sufficient quality and precision that uncertainty in the risk estimate is minimised.
8. An MRA should explicitly consider the dynamics of microbiological growth, survival and death in foods and the complexity of the interaction (including sequelae) between human and agent following consumption as well as the potential for further spread.
9. Wherever possible, risk estimates should be reassessed over time by comparison with independent human illness data.
10. An MRA may need re-evaluation as new relevant information becomes available.

15.4 MICROBIOLOGICAL RISK ASSESSMENT – PARADIGMS AND FRAMEWORKS

Micro-organisms are living entities and are very different to chemicals and physical hazards by their nature. MRA requires additional methods and terminology that are particular to microbiological risks (e.g. methods of estimating secondary transmission), and infective doses need to be developed. However, the enHealth model can, in general, be applied to MRA.

Haas et al. (1999) have produced the most comprehensive attempt at describing the methods used in QMRA and the particular needs of MRA. However, they have not developed a modified framework that attempts to encompass these different needs. Instead, their approach to MRA and QMRA loosely follows the National Academy of Sciences (NAS) framework proposed for chemical risk assessment (NRC 1983).

By contrast, the International Life Sciences Institute and the US EPA (ILSI 1996; 2000) have explicitly adapted the NAS framework to suit the unique challenges presented by MRA. Like the QRMA process described by Haas et al. it essentially follows the standard description provided by the NAS paradigm but uses synonymous terms for each part of the process.

15.5 EXPRESSION OF MICROBIOLOGICAL RISK

In common with other types of EHRA, estimating microbiological risk may be expressed in numerical notation or qualitatively, using terms such as low/medium/high. Risk may also be characterised by a narrative description of the risk, or whether it breaches standards or guidelines. In practice, however, a continuum exists from a fully quantitative through to a wholly narrative expression of risk.

An MRA cannot always practically achieve numerical expression of microbiological risk (ACDP 1996). This can be due to, for example, lack of dose–response data or a lack of understanding of the route of entry of a pathogen. Semi-quantitative or qualitative MRA can be applied in these situations.

The *Australian guidelines for water recycling* (AGWR) (NRMMC, EPHC, AHMC 2006) include a detailed discussion of risk assessment metrics using disability-adjusted life years (DALYs) and the reference pathogens used in their derivation. The use of DALYs and reference pathogens is based on the approach described in the World Health Organization's *Guidelines for drinking-water quality* (WHO 2006a). The DALY concept for microbiological risk assessment was not included in risk assessment advice in the *Australian drinking water guidelines* (NHMRC, NRMMC 2004). However, it was adopted for risk management of water in Australia with publication of the AGWR.

In brief, the DALY concept is a measure of the years of life lost by premature mortality (YLL) and years of healthy life lost in states of less than full health (years lived with a disability – YLD). The YLD parameter is weighted according to the severity of the disability.

> DALYs = YLL (years of life lost) + YLD (years lived with a disability/illness)

The advantage of using DALYs as a method of expressing QMRA risk is that DALYs include a measurement of the severity of impacts on human health arising out of infection and illness. They differentiate between relatively mild impacts, such as diarrhoea, and severe impacts, such as haemolytic uremic syndrome and even death. In terms of waterborne disease, the most commonly recognised illness is gastroenteritis (involving symptoms such as diarrhoea and vomiting) following ingestion of enteric pathogens. However, a number of these pathogens can cause more severe and long-lasting symptoms in a small percentage of infected people, for example:

- diabetes, associated with Coxsackie B4 virus (Mena et al. 2003)
- myocarditis, associated with echovirus and *Coxsackievirus* (Mena et al. 2003)
- reactive arthritis and Guillain-Barré syndrome, associated with *Campylobacter jejuni* (Havelaar et al. 2000; Nachamkin et al. 2001)
- haemolytic uraemic syndrome, associated with haemorrhagic *Escherichia coli* (Teunis et al. 2004)
- reactive arthritis, associated with *Salmonella* (Rudwaleit et al. 2001).

Determining DALYs for individual hazards includes considering acute impacts (e.g. diarrhoeal disease or even death) and chronic impacts (e.g. reactive arthritis, haemolytic syndrome). Calculation of DALYs includes consideration of each of the symptoms caused by a particular pathogen and the relative frequency of occurrence (NRMMC, EPHC, AHMC 2006).

The tolerable risk adopted in the AGWR is 10^{-6} DALYs per person per year, which is consistent with the WHO *Guidelines for drinking-water quality* (WHO 2006a). This is approximately equivalent to an annual diarrhoeal risk of illness of 10^{-3} (i.e. one illness per 1,000 people). In comparison, the reported rate of diarrhoeal illness in Australia is 0.8–0.92 cases per person per year (NRMMC, EPHC, AHMC 2006).

However, there are also problems with using DALYs. The methodology is relatively complex, potentially costly, and inherently conservative. It requires validated knowledge of infectious doses of selected reference pathogens, which may or may not reflect the full range of pathogenic organisms likely to have health impacts in an EHRA. It requires estimation of likely human exposure doses of a microbial population that may change rapidly over time. The calculation of YLD is based on application of severity factors for various health outcomes that are necessarily value-laden and may not reflect the values of affected stakeholders. Like the processes of non-threshold chemical EHRA, it requires the policy-driven establishment of acceptable levels of risk. Again, like chemical EHRA, such target risk levels carry a perception of exactitude. Some of the problems associated with using DALYs are acknowledged and discussed in most references, but not always. While there may not be a better tool available currently, it would be useful if these shortcomings were discussed more widely and appreciated by those applying the standards and methodologies.

Chapter 16: Guidance on route-specific EHRA

General guidance in this updated document is relevant to EHRA for all sources and routes of environmental exposures. There are, however, some specific EHRA guidelines available in Australia relating to specific environmental sources, such as contaminated sites, air pollution, contaminants in food, and contaminants in potable and recycled water. Where available, these specific guidelines take precedence over this enHealth document. A brief description of some of these guidelines is given below.

16.1 *ASSESSMENT OF SITE CONTAMINATION* NATIONAL ENVIRONMENTAL PROTECTION MEASURE

The *Assessment of site contamination* National Environmental Protection Measure (NEPM) establishes a nationally consistent approach to assessing site contamination to ensure sound environmental management practices by the community, which includes regulators, site assessors, contaminated land auditors, land owners, developers and industry.

The main outputs of the NEPM are:

- guidance on how to undertake assessment of contaminated sites (soil and groundwater)
- a table of health investigation levels (HILs) covering four types of exposure scenarios (low-density domestic dwellings, including home gardens; high-density domestic dwellings, without home gardens; open space, including parklands; and recreational areas and industrial/commercial premises).

HILs are presented in Schedule B(7) of the NEPM. While HILs are developed using a conservative EHRA approach designed to be protective of the health of the more sensitive or vulnerable 'receptors' on or near the site, it is important to emphasise that they are not intended to represent clean-up levels or targets for clean-up. Levels found to be marginally in excess of the HILs do not imply unacceptability or that a significant health risk is likely to be present. Exceeding a HIL means simply that further investigation is needed and that it should trigger a requirement for a more detailed 'Tier 2' risk assessment. This caveat on the interpretation of HILs is widely misunderstood, or at least overlooked in some circumstances.

Subject to an appropriate investigation and assessment process, a decision not to take further action may be justifiable. The decision on whether clean-up is required, and if so to what extent, should be based on site-specific assessment. Human health risk assessment is one aspect of making the decision; the NEPM also contains guidance on how to undertake an assessment of possible impacts on the ecology of a site. Other considerations, such as practicality, timescale, effectiveness, cost and durability are also important.

The NEPM contains two schedules:

- *Schedule A*, which is included in the NEPM, identifies the recommended process for assessing site contamination
- *Schedule B*, which comprises 10 general guidelines for assessing site contamination (Schedules B(1) – (10)).

A review of the NEPM commenced in 2004. In June 2007, NEPC agreed to initiate a process to vary the NEPM based on recommendations made in the NEPM review.

The proposed variation will ensure the NEPM remains the premier document for assessing site contamination in Australia by drawing on the latest methodologies for assessing human and ecological risk from site contamination, and updating guidance on site assessment methods in line with technological changes in Australia and overseas (EHPC 2010).

The *assessment of site contamination* NEPM can be accessed at the Environmental Protection and Heritage Council (EHPC) website at <www.ehpc.gov.au> and hard copies of the NEPM can be purchased by emailing <exec@ephc.gov.au>.

16.2 AIR POLLUTANTS

Poor air quality can have a significant bearing on the causes and exacerbation of respiratory disease. For example, asthma may be exacerbated by air pollution and more than two million or 11 per cent of Australians have asthma, including one in four primary school children, one in seven teenagers and one in 10 adults (AIHW 2000).

There are several issues that differentiate the risk assessment of air pollutants from pollutants found in other environmental media.

Exposure to air pollutants occurs in all activities, while indoors, in motor vehicles, while at work and during recreation. It is important that all sources of air pollutants are considered, noting that for some pollutants, the indoor and occupational environments may contribute the most to exposure. In addition, the surface area of the internal lining of the lungs is 50–70 square metres (about the size of a tennis court) compared with 1–2 square metres for the surface area of the skin). There are 300 million alveoli in adult human lungs and the air–blood barrier (consisting of the aqueous surface, epithelial lining and thin interstitial space) is 0.36 to 2.25 mm thick, indicating a much larger area for biological interaction to occur (Hrudey et al. 1996).

While the fundamental principles of risk assessment remain the same, different exposure assessment scenarios and assumptions are available when assessing ambient air pollution from diffuse and point-source regions or large areas, localised air pollution from point sources, and indoor air pollution such as may occur in the home or workplace. Where large populations are involved, different epidemiological methodologies such as time-series analysis may be able to be used. The risk assessment of a site-specific situation will differ from that for the development of a guideline as the former will usually relate to a specific, defined population while the latter will need to take into account a broader, more diverse population.

Currently, air quality in Australia is addressed in two separate NEPMs. The ambient air quality NEPM addresses a group of six 'criteria' air pollutants. These are:

- carbon monoxide (CO)
- sulfur dioxide (SO_2)
- nitrogen dioxide (NO_2)
- particulate matter 10 and 2.5 µm (PM_{10}, $PM_{2.5}$)
- photochemical smog (measured as ozone)
- lead (Pb).

Criteria pollutants are those that are common air pollutants found in relatively high concentrations. They are typically monitored via a network of monitoring stations. These networks are usually located to meet environmental management objectives. Monitoring stations are selected for a range of reasons, including monitoring of emissions from industrial facilities, major roadways and where high concentrations of secondary pollutants may be found. Ambient air exposures to pollutants are highly dependent on meteorological factors.

Another group of air pollutants is addressed in the air toxics NEPM. These comprise hazardous air pollutants and other specific substances that are found in trace concentrations, are specific to a particular setting or activity, and are monitored on a needs basis.

Currently, the air toxics NEPM provides monitoring advice and health-based air quality guidelines for selected volatile organic compounds (VOCs), semi-volatile compounds, polycyclic aromatic hydrocarbons (PAHs), heavy metals and aldehydes.

Irritant effects are often the critical health effect with criteria pollutants, and may occur from short exposures with negligible systemic absorption. Other non-irritant health effects, such as carcinogenicity (e.g. for benzene), mutagenicity and neurotoxicity, are receiving increasing attention.

The ambient air quality NEPM and the air toxics NEPM can be accessed at the EHPC website at <www.ehpc.gov.au>. This includes a recently released technical document supporting the setting of air quality standards (NEPC 2009).

16.2.1 Air quality EHRA – illustrative example

Application of the principles of EHRA for air pollution are well illustrated in the ongoing Clean and Healthy Air for Gladstone project being undertaken by Queensland Health and the Queensland Department of Natural Resources (see Box 4).

16.2.2 Managing odours

Odour and sensory irritation are effects that occur with very short-term exposures. This is primarily due to the fact the effects are receptor mediated and have very rapid onset at effective air concentrations. They are also likely to be factors associated with industrial emissions and polluted air, which contribute to decreased sense of wellbeing. Even the perception of odour can be considered as being the 'Trojan horse' for sinister toxic compounds in industrial emissions or polluted air (NHMRC 2006).

The issue of whether the negative impacts of odours that affect quality of life should be classified as adverse health effects is controversial. While it is undoubtedly important for these matters to be considered in a standard-setting process, it is debatable whether the appropriate place is within the scientifically rigorous steps of risk assessment outlined in this guidance document, or during the consultative processes that accompany risk management.

Specific guidance on the management of odour issues in air quality risk assessment is available in NSW DEC (2006) and in New Zealand guidance (NZ Ministry for the Environment 2003). These guidance documents address:

- how to assess the effects of odour, including how to determine whether 'objectionable or offensive odour' is causing adverse effects
- how to monitor the effects of odour through community surveys, odour diaries and council investigations
- when to use dispersion modelling for odour assessment
- how to manage odour emissions, including some basic information on suitable mitigation options
- an odour impact assessment checklist
- references to relevant legislative or regulatory instruments that impact on odour assessment.

The NZ guidance also includes a background technical report and draft good practice guide for odour management.

The NSW DEC document is quite pragmatic. It includes a recognition that avoiding odour impacts is a shared responsibility between operators and local land-use planners but that the operator of the facility that emits odour must ultimately be responsible for managing odour impacts beyond its boundaries.

It also recognises that odour emissions may not be preventable from some activities, and that 'no odour' may not be a realistic objective. It is a reasonable objective that may be exceedingly difficult to achieve in some cases.

16.3 FOOD CONTAMINANTS

Food-related risks can occur because of a range of factors, and in many cases the interdependence of these factors needs to be considered when assessing risk.

Some of the risk factors associated with food are:

- agricultural and veterinary chemical residues
- biological agents, including micro-organisms, viruses and parasites
- cooking and process-related artefacts
- food additives, processing aids and packaging materials
- mycotoxins, plant and marine toxins
- novel foods and ingredients
- radionuclides
- other environmental contaminants.

Some of these may occur as a result of commercial growing or processing of the food, some of them may be because of localised pollution at the point of planting, harvesting or animal husbandry. Where food is home-grown (e.g. in backyard vegetable patches), the contamination may be associated with uptake from locally contaminated soil.

Where the contamination occurs as the result of some commercial activity that can be controlled by regulation (e.g. specifying allowable food additives or migration from packaging material specifying good agricultural practice to ensure food residues do not exceed regulated maximum pesticide residues – MRLs), these aspects of food regulation fall within the jurisdiction of the Australian Pesticides and Veterinary Medicines Authority (APVMA) and/or Food Standards Australia New Zealand (FSANZ). MRLs can be accessed at <http://www.apvma.gov.au/residues/standard.php> or in the Food Standards Code which, can be accessed at <http://www.foodstandards.gov.au/foodstandards/foodstandardscode>.

Further information on the respective roles of the APVMA, FSANZ and the Department of Health and Ageing may be found in Chapter 17.

Where the contamination occurs as a result of environmental contamination from natural sources (e.g. marine or plant toxins, mycotoxins), the role of the food regulator is to specify maximum levels (MLs) that are protective of human health. Published MLs can be found in the Food Standards Code.

A key source of information is the series of Australian Market Basket Survey publications that are published biannually by FSANZ. These provide information about certain substances in a range of foods across Australia.

16.4 WATER CONTAMINATION

There are a wide range of water types, water uses and possible routes of transmission of waterborne hazards to humans. In undertaking health risk assessments, the characteristics and potential uses of water bodies need to be determined. Water sources include fresh, estuarine, marine and waste waters. Water uses can include supply of potable water for drinking and bathing, recreation, aquaculture and irrigation of crops. Human exposure to waterborne contaminants can include:

- direct exposure through ingestion, dermal contact, inhalation of aerosols or sprays
- indirect exposure through foods contaminated by irrigation water or water used for aquaculture and seafoods contaminated by waste water discharges. Health risk associated with food contamination via a waterborne route is within the scope of addressing the risk assessment of food.

The following is a summary of documents providing more specific regulations and guidance in relation to water contamination.

16.4.1 Australian drinking water guidelines

Published by the NHMRC, the *Australian drinking water guidelines* (ADWG) are intended to provide a framework for good management of drinking-water supplies that, if implemented, will assure safety at point of use. The ADWG have been developed after consideration of the best available scientific evidence. They are designed to provide an authoritative reference on what defines safe, good-quality water, how it can be achieved and how it can be assured. They are concerned both with safety from a health point of view and with aesthetic quality.

The ADWG are not mandatory standards. However, they provide a basis for determining the quality of water to be supplied to consumers in all parts of Australia. These determinations need to consider the diverse array of regional or local factors, and take into account economic, political and cultural issues, including customer expectations, and willingness and ability to pay. The ADWG are intended for use by the Australian community and all agencies with responsibilities associated with the supply of drinking water, including catchment and water resource managers, drinking-water suppliers, water regulators and health authorities (NHMRC & NRMMC 2004).

16.4.2 Australian guidelines for water recycling

The *Australian guidelines for water recycling* are designed to provide an authoritative reference that can be used to support beneficial and sustainable recycling of waters generated from sewage, greywater and stormwater, which represent an underused resource. The guidelines are intended to be used by anyone involved in the supply, use and regulation of recycled water schemes, including government and local government agencies, regulatory agencies, health and environment agencies, operators of water and wastewater schemes, water suppliers, consultants, industry, private developers, body corporates and property managers. The guidelines describe and support a broad range of recycling options, without advocating particular choices. It is up to communities as a whole to make decisions on uses of recycled water at individual locations. The intent

BOX 4: Clean and Healthy Air for Gladstone (CHAG) project

Gladstone is a city located approximately 550 km north of Brisbane. It is a centre for substantial industrial activity and development, including aluminium smelting, mineral and gas processing, coal shipment from its port facilities, coal-fired power generation, chemicals manufacturing and projected expansion as a liquid natural gas transport hub.

Community concerns about air quality and possible impacts on public health caused the Queensland Government to expand its air quality monitoring programs in the Gladstone region and implement a more detailed EHRA, in conjunction with the further allocation of six well-equipped monitoring stations to sample an extended range of air pollutants in the airshed.

The range of air pollutants monitored in the program over a period commencing in January 2009 included the six criteria air pollutants from the ambient air quality NEPM, nine metals, nine VOCs, three carbonyl compounds, six PAHs, some strongly acidic and basic gases and vapours, and fluorides. Particulates (PM_{10}, $PM_{2.5}$ and some preliminary data collection on PM_1) were of special interest because of the potential for coal dust generation in the region. The range of air pollutants measured was informed by data from the National Pollutants Inventory on aerially emitted chemicals in the Gladstone region. It is a more extensive suite of chemicals than those proposed for monitoring in either the ambient air or air toxics NEPMs.

The planning of the monitoring program, as well as the design of the EHRA and its objectives, all included extensive consultation with industry and community stakeholders, through nominated reference groups, and via public meetings. The EHRA was supplemented with a questionnaire and hospital admission health survey, undertaken in 2008 to identify any increased incidence of specific disease patterns that could be attributed to air pollution in the Gladstone region.

Planning the EHRA included benchmarking the measured air pollution measurements (appropriately averaged over one hour for irritants, or 24-hour–12-month averages for other pollutants) against published air quality standards from Australian NEPMs and various WHO, US or other international guidelines.

An interim report covering the health risk analysis of the first six months air monitoring data was released in November 2009. Since many of the measured air pollutants produced consistently low values, often below the limit of reporting (LOR), there was an extensive discussion of how the data were censored for statistical evaluation purposes. There was also some discussion of how to manage these few pollutants where conservatism in the adopted overseas health standard put the benchmark at a level lower than the LOR. The interim analysis was undertaken at six months to evaluate whether there was a need to modify the sampling program or whether any pollutants were being detected at a level which would give rise to health concerns. Fortunately, there were none which required urgent adjustment of the program.

Reports relating to the program are available on the website at <http://www.derm.qld.gov.au/environmental_management/air/clean_and_healthy_air_for_gladstone/reports.html>.

of these guidelines is simply to provide the scientific basis for implementing those decisions in a safe and sustainable manner. National water recycling guidelines were produced in two phases:

Phase 1: Australian guidelines for water recycling: managing health and environmental risks (Natural Resource Ministerial Management Council (NRMMC), Environment Protection and Heritage Council (EPHC), Australian Health Ministers' Conference (AHMC) 2006). Phase 1 of the guidelines provides a generic 'framework for management of recycled water quality and use' that applies to all combinations of recycled water and end uses. It also provides specific guidance on the use of treated sewage and greywater for purposes other than drinking and environmental flows.

Phase 2 (Module 1): Australian guidelines for water recycling: augmentation of drinking water supplies (NRMMC–EPHC–NHMRC). This current document, the first module of Phase 2 of the guidelines, extends the guidance given in Phase 1 on the planned use of recycled water (treated sewage and stormwater) to augment drinking water supplies. The document focuses on the source of water, initial treatment processes and blending of recycled water with drinking water sources.

Phase 2 (Modules 2 and 3): – Modules 2 and 3 cover use of stormwater for uses other than drinking water augmentation and managed aquifer recharge.

These documents are accessible on the EPHC website at <http://www.ephc.gov.au/taxonomy/term/38>.

16.4.3 Guidelines for managing risks in recreational water

Published by the National Health and Medical Research Council (NHMRC), the primary aim of these guidelines is to protect the health of humans from threats posed by the recreational use of coastal, estuarine and fresh waters. Threats may include natural hazards such as surf, rip currents and aquatic organisms, and those with an artificial aspect, such as discharges of wastewater (NHMRC 2008).

16.4.4 Australian and New Zealand guidelines for fresh and marine water quality

These guidelines are published by the Australian and New Zealand Environment and Conservation Council (ANZECC) and the Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ). The purpose of the guidelines is to provide an authoritative guide for setting water quality objectives required to sustain current, or likely future, environmental values uses for natural and semi-natural water resources in Australia and New Zealand.

Part 3: Regulatory context

Chapter 17: Australian regulation of chemical hazards

Chemicals assessment in Australia is mainly the province of agencies associated with the Australian Government.[5] However, the implementation of chemicals regulatory programs, especially those aimed at management of environmental health risks, is mainly the province of state, territory and local governments.

The Environmental Health Committee (enHealth) is a Commonwealth–state coordinating body that provides policy advice in these areas. It is constituted as a subcommittee of the Australian Health Protection Committee (AHPC) (see <http://www.health.gov.au/internet/main/publishing.nsf/Content/ohp-environ-enhealth-committee.htm>).

Under its terms of reference, enHealth has responsibility for providing agreed health policy advice, implementation of the *National environmental health strategy 2007–2012*, consultation with key players, and the development and coordination of research, information and practical resources on environmental health matters at a national level. The advice development process is strongly based on collaboration and consultation.

The enHealth membership includes representatives from Commonwealth, state and territory health departments; the New Zealand Ministry of Health; the National Health and Medical Research Council; Choice; Environmental Health Australia; the Australian Local Government Association; the Commonwealth Department of the Environment, Heritage, Water and the Arts; the Commonwealth Department of Climate Change; the Public Health Association of Australia; and the Deputy Chair (who must be Australian Aboriginal or Torres Strait Islander) of the enHealth Working Group on Aboriginal and Torres Strait Islander Environmental Health.

Under the guidance of the AHPC and with reference to the *National environmental health strategy 2007–2012*, enHealth's terms of reference are to:

- provide nationally agreed environmental health policy advice, based on the best available evidence and expertise, to the Australian Health Ministers' Advisory Council (AHMAC) through the AHPC
- coordinate implementation of nationally agreed environmental health policies and approaches
- provide environmental health expertise and support for AHPC's emergency management role
- under arrangements to be agreed between AHMAC and the Environment Protection and Heritage Standing Committee (EPHSC), keep the AHPC and AHMAC informed of developments in environmental policy with significant health implications and provide expert, and where nationally agreed, health advice in environmental policy forums
- consult with consumers and other stakeholders as appropriate in developing environmental health policy advice and implementing environmental health policies
- contribute, through the Commonwealth, to international collaboration on environmental health issues
- coordinate research, share information and develop practical environmental health resources, including through expert and/or nationally agreed publications as guided by the AHPC.

17.1 OVERVIEW OF THE CHEMICAL RISK ASSESSMENT AGENCIES

Several bodies are involved in the process for undertaking chemical hazard and risk assessments. National chemicals legislation and responsible authorities are outlined in Table 23.

Essentially, if a product or chemical is not intended for agricultural or veterinary use, nor for human therapeutic or food/food additive use, then it falls to National Industrial Chemicals Notification and Assessment Scheme (NICNAS) for assessment and review. The roles and responsibilities of the agencies are set out in more detail in the website of the Australian Government National Chemicals Information Gateway at <http://apps5a.ris.environment.gov.au/pubgate/cig_public/!CIGPPUBLIC.pStart> and the NEPC website of the National Framework for Chemicals Environmental Management (NChEM) at <http://www.ephc.gov.au/taxonomy/term/75>. Although it is now somewhat out of date, the National Profile on Chemicals Management Infrastructure (Environment Australia 1998) detailed the legislative and administrative infrastructures operating in Australia at that time. The document was produced in compliance with international agreements to detail national plans for chemicals risk management.

In 2008, the Productivity Commission undertook a comprehensive review of regulation of the chemicals and plastics sector. The reports are available at <http://www.pc.gov.au/projects/study/chemicalsandplastics/docs/finalreport>. Recommendations were made about many aspects of the regulatory system, and these are currently being implemented through the Council of Australian Governments (COAG).

5 This chapter describes the responsibilities of Australian regulatory agencies as of August 2011. Government functions and departmental names tend to change over time, so that this descriptive information may only be accurate and relevant at this point of time.

Table 23: Chemicals regulation in Australia

Agency	National Industrial Chemicals Notification and Assessment Scheme (NICNAS)	Australian Pesticides and Veterinary Medicines Authority (APVMA)	Therapeutic Goods Administration (TGA)	Food Standards Australia New Zealand (FSANZ)
Portfolio	Health and Ageing	Agriculture, Fisheries and Forestry	Health and Ageing	Health and Ageing
Scope	Assessment and review, not registration based	Assessment, product registration and review	Assessment and product registration	Assessment and standard-setting
Relevant legislation	*Industrial Chemicals (Notification and Assessment) Act 1989*	*Agricultural and Veterinary Chemicals Code Act 1994 and Veterinary Chemicals Administration Act 1994*	*Therapeutic Goods Act 1989*	*Australia New Zealand Food Authority Act 1994. Food Standards Code*
About the chemicals	Industrial chemicals are varied and include dyes, solvents, adhesives, plastics, laboratory chemicals, paints and coatings, chemicals used in cleaning products, cosmetics including some sunscreens.	Agricultural products, including chemicals that destroy/repel pests or plants. Veterinary products are used to prevent, diagnose or treat disease in animals.	Therapeutic goods, including prescription medicines and medicines available over the counter (OTC). OTCs include complementary medicines (herbal, vitamin and homeopathic preparations); some sterilants and disinfectants and cosmetic-type products that make therapeutic claims.	Food additives and processing aids used in food to assist with preservation, flavouring, colouring, or modifying its functions; permitted levels of some environmental contaminants are also controlled via the Food Standards Code.

What follows is a brief synopsis of the roles and methods of operation of the two agencies listed in Table 23 that have main responsibility for chemicals likely to be released into the environment (industrial chemicals, AgVets and consumer product chemicals). Note that the agencies and responsibilities are current at 2011, but may change according to future government reorganisations.

17.2 NATIONAL INDUSTRIAL CHEMICALS NOTIFICATION AND ASSESSMENT SCHEME

NICNAS is a chemical entity-based notification and pre-market risk assessment scheme. Industrial chemicals are defined in the context of their use and by their exclusion as therapeutic goods, food or food additives, pesticides and veterinary medicines. The scope of the NICNAS assessments comprises three elements – occupational health and safety, public health and environmental impact – over the life cycle of the chemical.

NICNAS was established in July 1990 under Australian Government legislation: the *Industrial Chemicals (Notification and Assessment) Act 1989* (ICNA Act)

NICNAS's activities include:

- assessing industrial chemicals that are new to Australia for their risk to human health and the environment, before use or release into the environment
- assessing industrial chemicals that are already in use in Australia (known as 'existing chemicals') in response to health and environmental concerns
- making risk assessment and safety information on chemicals and their potential occupational health and safety, public health and environmental risks widely available to workers, the public, industry and other government agencies
- enabling the public, organisations and key stakeholders to have effective input into decision-making processes regarding the safe use of chemicals.

The risk assessment entails some or all of the following elements:

- hazard identification
- hazard assessment, incorporating the dose–response relationship
- exposure assessment
- risk characterisation, where the hazard and exposure assessments are integrated.

Based on the risk assessment findings, recommendations are made to mitigate the risks.

17.2.1 Existing chemicals

17.2.1.1 Data requirements

For existing industrial chemicals, the data requirements depend on the assessment type. A standard dataset for a full-priority existing chemical (PEC) comprises information confirming the identity of the chemical, the physico-chemical properties and use of the chemical (including import/manufacture volumes), all available toxicological and epidemiological data, detailed exposure information for workers, the public and the environment and risk management initiatives. The toxicological package includes available human, animal and *in vitro* data and ecotoxicity and biodegradability/fate data for the environmental assessment.

The dataset for a preliminary PEC may or may not include a detailed toxicological package or detailed exposure data. Risk assessment, in terms of a formal risk characterisation for specific uses, is not carried out for preliminary assessments.

The dataset for a secondary notification assessment is determined in accordance with a set of circumstances (criteria) as set out in the ICNA Act. Should these circumstances require a re-evaluation of the risks assessed in the original PEC report, then a formal risk characterisation is usually carried out.

17.2.1.2 Exposure data

Occupational and public exposure data is provided as a statutory obligation (under the ICNA Act) from applicants (for assessment). This information is supplemented from literature review, site visits, international reports (e.g. OECD SIARs), and where data is lacking from modelling. The model that has been used to date is the UK HSE EASE model, which provides estimates of airborne and dermal exposure for different occupational scenarios.

Where exposure by inhalation is the major route of exposure, and the toxicological database includes good quality inhalation data (human or animal), the common practice is to use 'external' exposure data (i.e. not to attempt to extrapolate to 'internal' dose) in the risk characterisation process (see Section 17.2.3). When 'external' exposure data is used/determined, no adjustment is made to account for reduced personal exposures resulting from the use of personal protective equipment (e.g. respiratory protection, gloves). However, where mechanical ventilation is installed, this can be factored into the EASE model, should suitable monitoring data not be available (i.e. measured when ventilation has been installed and is operational). The quality of the monitoring data should also be a factor considered in the risk characterisation and exposure standard-setting processes .

Where dermal exposure is an important route of exposure or where the toxicological database does not provide an inhalation study, internal (dose) exposure may be estimated utilising the available pharmacokinetic data and used in the risk characterisation process.

Estimates of public exposure are carried out similarly for existing chemicals and new chemicals (see 17.2.2.1).

17.2.1.3 Toxicological data

Toxicological and epidemiological/case study/clinical data is also provided as statutory obligation (under the ICNA Act) from applicants (for assessment). This data is supplemented from literature review and international reports (e.g. OECD SIARs, IPCS, IARC, ECETOC).

Currently, available toxicological and epidemiological data is evaluated in conjunction with available pharmacokinetic data, to estimate the critical no observed adverse effect levels (NOAEL) or, if not determined, the lowest observed adverse effect level (LOAEL) for both acute and chronic exposures for each relevant route of exposure (i.e. oral, dermal and inhalation). The health hazards for each endpoint are classified in accordance with the Safe Work Australia (SWA) NOHSC Approved Criteria for classifying hazardous substances.

The quality of the toxicological database should be a factor considered in the risk characterisation and exposure standard-setting processes.

17.2.2 New chemicals

For new industrial chemicals, the data requirements depend on the notification category and are stipulated under the ICNA Act. A standard dataset comprises information confirming the identity of the chemical, the physico-chemical properties and use of the chemical; detailed exposure information about how workers; the public and the environment are exposed to the chemical; and a standard toxicological package. The toxicological package includes animal and *in vitro* data for the human health assessment, and ecotoxicity and biodegradability data for the environmental assessment.

17.2.2.1 Exposure assessment

The occupational exposure assessment is conducted by establishing the use pattern of the chemical and identifying the sources of occupational exposure. Exposure is then estimated by taking into account the routes of exposure, the frequency and duration of exposure, and measured worker data (e.g. example, atmospheric and/or biological monitoring results). Information is needed for each of the scenarios where workers are potentially exposed to the chemical.

For new industrial chemicals, the occupational exposure assessment is usually qualitative, as measured data is unlikely to be available and there is insufficient information available to predict reliable quantitative estimates. Modelling (e.g. using EASE) is occasionally used.

Estimates of public exposure from consumer use (e.g. for cosmetics) are made using the expected type and frequency of use. The possibility of public exposure arising from release into the environment during transport, manufacturing or end-use is also taken into account.

17.2.2.2 Toxicological assessment

Both human and experimental animal data is assessed in accordance with international guidelines to identify the critical health effects of the chemical and to determine the dose–response relationship, with NOAELs established wherever possible. For new industrial chemicals, human data is usually not available. The health hazards of the chemical are classified in accordance with the approved criteria for classifying hazardous substances. Classification is also determined according to the Global Harmonized System for Classification and Labelling of Chemicals (GHS) (see <http://www.unece.org/trans/danger/publi/ghs/ghs_welcome_e.html>) that is expected to be adopted in future years.

For new chemicals, the toxicological database may consist of studies that have been performed with a structural analogue of the notified chemical, or with a formulation. Adequacy and applicability of the data will be taken into account when performing the assessment. Where data gaps exist, or where toxicological data has not been provided, as with some classes of polymer, the toxicological hazard may be predicted from the chemical's physical properties or the characteristics of structurally related chemicals, given that factors such as volatility, solubility and molecular weight can indicate the likely extent of absorption across biological membranes.

17.2.3 Risk characterisation

The current methodology utilised by NICNAS for both new and existing chemicals is the 'margin of exposure' (MoE) approach.

In deriving the MoE, direct comparison is made of the critical NOAEL with the measured/estimated exposures for each occupational scenario of relevance to manufacture and use in Australia. This is carried out separately for inhalation and dermal exposure (where relevant), that is, by using NOAELs derived specifically from each route of exposure.

Where exposures may be significant by both routes, the combined estimated internal dose may be used. In this case, the oral NOAEL (for the critical effect) is usually considered the more appropriate NOAEL for deriving the MoE.

The resulting MoE is then evaluated (for each route), taking into account the quality of the available database (e.g. whether derived from human data, uncertainties in the database) and nature or severity of effect (e.g. carcinogen, sensitiser). No specific values are assigned to component uncertainty factors (this is usually part of the exposure standard-setting process carried out by SWA. However, the risk characterisation process takes these uncertainties into account in evaluating the adequacy of the MoE.

Based on the magnitude of the MoE, current risk management initiatives are assessed and where found inadequate, recommendations for additional exposure reduction measures (controls) or other risk management initiatives are promulgated. Recommendations may include regulatory action by SWA or other agencies (e.g. chemicals scheduling and state and territory environmental agencies, where public health and environmental risks have been identified).

Recommendations to SWA may include: the setting of an occupational exposure standard , review of an existing exposure standard, scheduling of a substance in accordance with the model regulations for control of workplace hazardous substances and, as a last resort, phase-out of use and manufacture.

The health risk to workers is characterised by integrating the occupational exposure and toxicological assessments. For brief or short-term exposures, human data and information from acute toxicity studies in animals are taken into account to determine the risk of adverse health effects such as acute respiratory effects and skin irritation. For longer-term and repeated exposures, the health risk to workers is characterised by firstly comparing exposure estimates with NOAELs to give an MoE, and then deciding whether there is cause for concern.

Similarly, an estimate of risk to the public is characterised through the hazard of the chemical and determining whether there is any significant public exposure. The approach taken will vary with the nature of any hazard posed by the chemical and the extent of data on exposure and toxicology.

Matters taken into account when characterising the risk, include the uncertainty arising from the variability in the experimental data and inter- and intra-species variation, the nature and severity of the health effect and its relevance to humans and the reliability of the exposure information.

Where it is not possible to determine a NOAEL or LOAEL (e.g. from lack of suitable data), the risk is evaluated on the basis of qualitative or quantitative exposure relevant to the group of workers being considered. For new chemicals, a more qualitative characterisation takes place as exposures are often unknown or more difficult to predict.

17.3 AUSTRALIAN PESTICIDES AND VETERINARY MEDICINES AUTHORITY (APVMA)

The APVMA regulates agricultural and veterinary chemicals, as defined by the *Agricultural and Veterinary Chemicals Code Act (1994).* OHS risk assessments on new and existing chemicals that are under review are conducted for the APVMA by the OCS.

17.3.1 Data requirements

The Agricultural and Veterinary Chemicals Code Act makes provision for approving active constituents, registering products containing those active constituents, and reconsideration of existing chemicals that have been nominated for review. Data required for the OHS assessment of AgVet chemical products includes:

- use pattern of the product
- formulation composition of product
- physico-chemical properties of the active constituent and product
- toxicology of the active constituent and product
- exposure data.

17.3.2 Exposure data

It is a requirement under the AgVet Code Act that exposure data and adverse incident reports (when they occurred following use according to the label) must be provided to the APVMA by applicants. Exposure data may cover manufacture/formulation of AgVet products and end-use situations. Exposure data provided by applicants is supplemented from literature review, international reports (e.g. US EPA, UK MAFF), field/site visits and modelling. The model used to date is the UK Predictive Operator Exposure Model (POEM). Occasionally, exposure data from the US Pesticide Handlers Exposure Database (PHED) has been used where applicants provide subset exposure data.

The exposure assessment constitutes consideration of the use pattern of the product, identification of potential exposure scenarios and predominant routes of exposure in each case. For AgVet chemicals, it is generally accepted that skin absorption is the predominant route of exposure. In general, inhalation exposure comprises only a small proportion of total exposure, except when the product is applied in an enclosed space (e.g. fumigants). Therefore, where the toxicological database includes dermal dosing studies of appropriate quality and duration, 'external' exposure data is used in the risk characterisation process.

Pesticide exposure assessments also take into consideration the protection afforded by label-specified protective equipment. Default protection factors are utilised in the absence of specific data.

17.3.3 Toxicological data

Toxicological and epidemiological/case study data provided by applicants is evaluated by the Office of Chemical Safety (OCS). The OCS evaluation is considered in order to determine relevant endpoints and NOAEL/LOAEL(s) for use in the OHS risk assessment. The selection is based on factors including the quality of the database, the frequency of use of the product, health significance of the endpoint(s) and predominant route of exposure.

For new AgVet chemicals, the health hazards of the chemical are classified in accordance with the approved criteria for classifying hazardous substances.

17.3.4 Risk assessment

The risk assessment takes into consideration the hazard of the chemical and the potential for occupational exposure. In general, an end-use risk assessment is conducted for AgVet products. Potential exposure is determined by the use pattern of the product and current agricultural/animal husbandry practices (including existing exposure mitigation methods such as protective equipment and engineering controls).

As for industrial chemicals (NICNAS – existing chemicals), AgVet assessments utilise the 'margin of exposure' (MoE) approach. The benchmark MoE is determined on a case-by-case basis, following consideration of the quality of the database, nature and severity of the health effect and known variability in human metabolism of the chemical. In general, a tenfold factor is considered appropriate to account for inter-species extrapolation and a similar factor (10x) for intra-species variability.

Current exposure mitigation methods are evaluated quantitatively, where possible. In the absence of data or models, qualitative assessments are conducted, based on generalised information about the use pattern and 'scientific judgement'. Where current exposure assessment methods are found to result in unacceptable risk, additional exposure and risk reduction methods may be recommended.

OHS recommendations on regulatory action may include restrictions on the use of the chemical and exposure mitigation methods.

17.4 ROLE OF THE OFFICE OF CHEMICAL SAFETY (OCS) IN PUBLIC HEALTH RISK ASSESSMENT

The OCS is a professional scientific group within the Office of Health Protection in the Department of Health and Ageing that provides advice to the minister, to relevant committees of Commonwealth, state and territory government agencies, and to the public on possible risks to health associated with exposure to chemicals in the environment. These include agricultural chemicals, veterinary drugs, industrial chemicals and other chemicals that may have an impact on public health.

17.4.1 Chemical products assessment

The main function of the OCS is to assess the toxicology and public health aspects of applications for registration of new

agricultural and veterinary chemicals.

The origins of this function can be traced back to 1984 when the then Toxicology Evaluation Section (TES) was created following a Senate inquiry into hazardous chemicals, which noted the increasing use of chemicals in the environment. Increasing public concern and media attention to chemical exposure demanded greater accountability from the chemical industry and from government regulators. In addition, chemical residues in export produce (e.g. beef, wheat, dairy goods) have important implications for trade so Australia must be able to ensure standards of chemical regulation acceptable to international markets, as well as to domestic consumers. In acknowledgment of public health and trade concerns, chemicals regulation as an area of public policy has developed, recognising that the numbers of chemicals requiring assessment and the complexity of the assessment process have increased.

Recommendations on individual chemicals form an important component of the decisions by the APVMA in the registration of chemicals. OCS reports are also used to assist the Advisory Committee on Chemicals Scheduling to make recommendations to the Departmental Scheduling Delegate on poisons scheduling. The OCS provides the secretariat for the scheduling advisory committees. The recommendations of the delegate are formally incorporated into the *Standard for the uniform scheduling of medicines and poisons* (SUSMP) which forms the basis for national uniformity in drugs and poisons scheduling.

The scope of OCS assessments may be expanded to address toxicological hazards associated with a range of natural and synthetic chemicals. Examples of these include chemicals in certain consumer products, and environmental contaminants.

17.4.2 Evaluation reports

Reports are structured to allow for ready access to the main points arising from the assessment. They are written to provide sufficiently detailed information for the reader to form an independent conclusion and aim to obviate the need, during a subsequent review of the chemical (or product), to refer back to the original study data.

Reports include the following components.

- Submission summary – briefly outlines the results from all studies accompanying the application/ submission and includes a discussion of the important findings and appropriate recommendations.
- Main body of the report – contains detailed outlines of the studies conducted with the chemical, including methodology, the extent of monitoring for biological changes, all treatment-related effects and any other observations or information that may be pertinent to the assessment of the significance of the findings.

  Reports on agricultural and veterinary chemicals generally assess the types of studies set out in Chapter 9. Submitted studies usually include acute studies with the technical grade active constituent (TGAC), relevant product formulations and, where relevant, studies on the toxicity of key impurities.
- Consolidated summary – contains the integrated summaries of study results from previous submissions relating to the particular active ingredient, plus the newly evaluated data.
- Confidential business information – impurity profiles, product ingredients and information on additives in formulations, etc. are included in removable appendices at the end of the report.

17.4.3 Chemical review programs

Until the mid-1990s, there was no formal program to review 'old' pesticides and veterinary drugs in Australia. However, two programs provided an informal mechanism for reviewing a number of aspects of chemicals safety. First, a review process for individual chemicals occurred on the suspicion of human health and/or environmental concerns, or following international regulatory action.. In practice, this usually involved reviewing limited extra data related to the particular issue of concern rather than conducting a comprehensive review.

Second, the Technical Grade Active Constituent (TGAC) Scheme, introduced in 1985, while designed primarily to identify the source and ascertain the quality of technical grade materials used for formulating end-use products (EUPs) used in Australia, had significant elements of a review program in that mammalian toxicology data and environmental data were collected and reviewed. As a result of the TGAC scheme, toxicology data on more than 400 pesticides were reviewed with respect to public health considerations over a 6–7 year period. As a consequence of the above two programs, Australia was well placed to develop further more formal review arrangements.

Australia introduced a formal program to review existing agricultural and veterinary chemicals in 1994 under the title of the Existing Chemicals Review Program (ECRP), managed by the then National Registration Authority (NRA). The program carried out systematic reviews of existing agricultural and veterinary chemicals on a priority basis. This program was one of a number of initiatives arising from a 1990 Senate inquiry into aspects of the legislative, administrative and regulatory procedures for agricultural and veterinary chemicals. The ECRP stemmed largely from the fact that many registered chemicals entered the market place based on criteria now recognised as outdated by today's regulatory standards. The ECRP involved cooperative arrangements between the health portfolio (public health), environment portfolio (environmental assessment), Safe Work Australia (formerly the National Occupational Health and Safety Commission) (occupational health and safety) and the APVMA (chemistry, efficacy and agricultural issues, residues and registration). The ECRP ran in parallel with the so-called 'Special Review Program' that was in place to deal with issues relating to existing chemicals that needed to be dealt with rapidly (e.g. that a particular use practice was leading to MRL exceedances in exported food commodities). In about 2000, the ECRP and Special Review Program were combined into one Chemical Review Program.

The goal of the Chemical Review Program is to ensure agricultural and veterinary chemicals in use in Australia can be used safely and effectively. The program operates according to the principles of openness, fairness and consistency with regard to public consultation, selection of chemicals for review, and standards of assessment. All aspects of a chemical (public health, OHS, environmental, efficacy, and animal and crop safety) are considered in a review. The review program:

- works towards the goal that AgVet chemicals remain safe and effective when used according to label instructions by specifically considering toxicity and exposure patterns in relation to public health, OHS and environmental control mechanisms known and potential environmental impacts efficacy safety issues in relation to target species (animal and crop) management options to reduce identified risks
- helps maintain the protection of Australian trade and commerce in agricultural produce and livestock
- helps address community concerns and general interest in agricultural and veterinary chemicals by providing information to the public on the use of chemicals and their environmental, public health and OHS aspects
- considers public nomination of chemicals for review.

AgVet chemicals are selected for review on the basis of agreed criteria including their potential health and environment hazard(s), exposure potential, age and adequacy of the database, efficacy, international regulatory actions, and trade and other agricultural implications. The chemical selection process also incorporates a mechanism for public nominations of chemicals.

The public and occupational health aspects of reviews are assessed by staff of the OCS, which provides toxicological and chemicals policy advice as required to appropriate committees of Commonwealth, state and territory government agencies, and to the public.

The OCS also undertakes technical policy development and provides health advice on international chemicals treaty negotiations. It interacts with other government agencies on a range of environmental health issues.

An important role of OCS is to encourage and, where practical, to extend international harmonisation of chemicals regulation, including toxicological reviews and re-registration programs.

17.4.4 Assessment processes within the OCS

Toxicologists within the OCS assess mammalian toxicology and toxicokinetic data and prepare written assessment reports that carry sufficient detail of the studies and findings to allow an independent assessment of the data. As the primary emphasis is on independent assessment, limited regulatory status is given to company summaries and company-sponsored 'expert reports'. It is important that all toxicity data and the methods by which they are obtained be subjected to critical and independent scientific assessment.

Hazard/risk assessment: Given the complexity of biological data interpretation and the need for professional judgement and a flexible approach when assessing the public health risk of chemicals, it has not been the policy to establish prescriptive methodologies for hazard and risk assessment, although several guidance documents for evaluators have been drafted. In general, a qualitative approach is used to assess chemical risk. The approach taken to derive an acceptable daily intake (ADI) follows the principles outlined in *Environmental health criteria* monographs 104 and 210 prepared by the WHO/UNEP/ILO International Programme on Chemical Safety (IPCS).

While the main focus of agricultural and veterinary chemical assessments is a consideration of human exposure to pesticides through ingesting residues in food and/or drinking water, the direct dermal or inhalational exposure of the public, as users of chemicals (in the home garden/domestic setting) or as bystanders to agricultural or licensed pest control operator (PCO) use, is also taken into account.

In general, a classification system for public health aspects is not used when regulating chemicals with potential carcinogenicity. The potential human carcinogenicity of chemicals is assessed using a weight-of-evidence (WoE) approach, which takes into account epidemiological data, carcinogenic potency in animals, biological relevance and potential human exposure. Australia, in this regard, supports the general approach outlined by IPCS.

Exposure assessment: Two aspects of exposure to AgVet chemicals are evaluated. The OCS takes responsibility

for evaluating worker and bystander exposure, and uses these assessments to recommend appropriate safety directions.

Possible dietary exposure to pesticide residues is based on calculations of likely daily intakes of pesticide residues – either national theoretical maximum daily intakes (NTMDIs) or national estimated dietary intakes (NEDIs) – and these are the responsibility of the APVMA and FSANZ. The procedures used are based the *Guidelines for predicting dietary intake of pesticide residues* (1989) published by the UNEP/FAO/WHO Food Contamination Monitoring Programme in collaboration with the Codex Committee on Pesticide Residues.

Food consumption data is used in dietary risk assessment and is available from the Australian Dietary Survey (ADS) and the Market Basket Survey (MBS). The ADS estimates actual food intakes in various sub-groups of the population, taking into account such factors as age and ethnic background. The MBS measures the amount of pesticide residue in ready-to-eat food based on a typical diet for different age groups. Estimated daily food residue intake can be compared with the ADI.

Although there is no formal program to assess human exposure to pesticides in the domestic setting, pesticides for domestic use are restricted to those of low toxicity, and they have appropriate controls on availability, packaging and labelling. Additional exposure assessment may be carried out where a particular concern arises.

Risk management: Toxicological issues may raise concerns with respect to supply, availability, and use of agricultural or veterinary chemicals. The supply and availability of chemicals can be managed through APVMA's registration process; that is, approval for pesticide technical grade active constituents (TGACs) may not be granted (or be withdrawn), approval for particular uses of a pesticide or veterinary chemical may not be granted (or be withdrawn), or registrations for particular products may not be granted (or be withdrawn), in order to eliminate or reduce potential public exposure.

The use of registered AgVet products on the market can be regulated through poisons scheduling and appropriate labelling. The Delegate of the Commonwealth Department of Health and Ageing, acting on the advice of its Advisory Committee on Chemicals Scheduling (ACCS) recommends classification of drugs and poisons which are published in the *Standard for the uniform scheduling of medicines and poisons* (SUSMP); these federal recommendations are adopted by state and territory legislation. The more restrictive schedules prescribe restrictions on supply and use, as well the use of appropriate signal headings on labels.

The poisons schedule classification of a chemical and its formulated products, together with product labelling instructions (first aid and safety directions), help control the availability of products and help minimise the exposure of users.

17.5 ROLE OF THE DEPARTMENT OF SUSTAINABILITY, ENVIRONMENT, WATER, POPULATION AND COMMUNITIES (DSEWPC) IN ENVIRONMENTAL RISK ASSESSMENT

On behalf of NICNAS and the APVMA, scientific staff of DSEWPC provide assessments of potential environmental issues relating to chemicals that have been submitted for assessment by these two agencies. These assessments address potential impacts on Australian flora and fauna. Coupled with this are study requirements to enable assessment of environmental fate and persistence. Manuals explaining how these risk assessments are undertaken are available at the EPHC website.

17.5.1 Data requirements

The data requirements for environmental assessment of agricultural chemicals is set out in Section 7 of the APVMA *Agricultural (and veterinary) manual of requirements and guidelines* (Ag-MORAG and Vet-MORAG). The requirements include studies on toxicity to birds, mammals and other vertebrates, freshwater and marine organisms, non-target terrestrial invertebrates (e.g. worms, bees) and plants. Studies include short-term and longer-term exposures.

17.6 STANDARDS AUSTRALIA MODEL OF RISK MANAGEMENT

A risk management standard was first published by Standards Australia in 1999 and updated in 2004 and 2009 (AS/NZS ISO 31000:2009). This is directed towards as wide a range of risk and risk management disciplines as possible for application to a very wide range of activities or operations of any public, private or community enterprise, or group so as to establish a systematic risk management program. The standard provides a generic guide for the establishment and implementation of the risk management process involving establishing the context and the identification, analysis, evaluation, treatment, communication and ongoing monitoring of risks.

The risk management process outlined in the standard contains a model of risk assessment and uses the term 'risk analysis' to describe the process. Risk analysis is defined as a systematic use of available information to determine how often specified events may occur and the magnitude of their consequences.

The original risk management standard was followed by a further standard, *Environmental risk management: principles and process*, HB 203: 2000. This gave more extensive detail on environmental risk management using the framework established in the previous generic Australian Standard.

Both standards provide qualitative measures of consequence and likelihood. Appendixes in HB203: 2000 detail sustainability principles, links between environmental risk and environmental management systems, discussion of how the acceptability of risk may be considered, sources of information for risk identification and cost-benefit analysis.

17.7 RISK MANAGEMENT ASSOCIATED WITH GENETICALLY MODIFIED ORGANISMS

The Office of the Gene Technology Regulator (OGTR) is responsible for protecting human health and safety and the environment by identifying and managing risks posed by, or as a result of, gene technology. It assesses the environmental and human health risks associated with materials derived from or by GMOs prior to their use or release into the environment. The risk assessment processes employed are outlined in the *Risk analysis framework* (RAF), a key document for informing applicants, stakeholders, the public and other domestic and international regulatory bodies about the rationale and approach adopted by the regulator in undertaking risk analyses and arriving at risk management decisions and licence conditions. The RAF has been recently updated and released in May 2009. The outcome of the risk assessment process is a *Risk assessment and risk management plan* (RARMP) produced in response to licence applications proposing dealings involving intentional release (DIR) of GMOs into the Australian environment.

The RARMP categorises risks (including potential risks to human health) into broad categories (negligible, low, moderate, high, very high) rather than making numerical estimates of risk.

17.8 SUPPORTING PROFESSIONAL ORGANISATIONS

A number of scientific and industry professional organisations have a range of scientific expertise in toxicology, health risk assessment and environmental health among their memberships, and can provide useful support for the continuing education of professionals engaged in EHRA, within government and in the academic and commercial sectors. These organisations include:

- Australasian Society of Clinical and Experimental Pharmacologists and Toxicologists (ASCEPT)
- Royal Australian Chemical Institute (RACI)
- Society of Environmental Toxicology and Chemistry (SETAC)
- Society for Risk Analysis (SRA) – Australia and New Zealand Regional Organisation
- Australian Land and Groundwater Association (ALGA)
- Australian Contaminated Lands Consultants Association (ACLCA)
- Australasian College of Toxicology and Risk Assessment (ACTRA)

This list is not exhaustive, but a special note is made of ACTRA, which was established with support from enHealth partners in 2006 to specifically address professional development in the areas of toxicology and risk assessment relevant to EHRA.

17.8.1 The Australasian College of Toxicology and Risk Assessment

With funding support and encouragement from most of the enHealth partners, the Australasian College of Toxicology and Risk Assessment (ACTRA) was established in 2006 (see <www.actra.org.au>).

ACTRA sponsors workshops and continuing education programs in toxicology and health risk assessment, with the objective to:

- advance the study and applications of toxicology and health risk assessment as professional scientific disciplines
- cultivate (and maintain) the highest standards of professional practice and ethics of those engaged in the sciences of toxicology and health risk assessment.

ACTRA is a professional society whose members are accepted on the basis of their experience and educational achievements. Members may apply for listing on the ACTRA Register of Toxicologists and Risk Assessors. This registration process, modelled on that of the British Toxicological Society, involves a peer-reviewed assessment of their professional achievements, active service and standing in the profession.

Chapter 18: International programs and guidance on EHRA

This chapter summarises some of the EHRA guidance that has been developed in selected US, European and international programs. It is not intended to be a comprehensive review of these programs, but to provide some brief insights into their key elements and how they may have contributed to the development of an Australian approach to EHRA.

18.1 UNITED STATES PROGRAMS

The United States has been particularly active in developing guidance on toxicity assessment and EHRA methodology. Reference has been made elsewhere in this enHealth document (see Sections 1.8, 3.8, 3.9) to the NRC 2008 proposals to update US approaches to EHRA, and to the Tox21 Toxicity Testing in the 21st Century program (see Section 9.9) which seeks to supplement and/or replace traditional animal-based toxicity testing with newer *in vitro*, QSAR and *in silico* techniques.

Current approaches and challenges in US chemicals regulatory policies were recently outlined by the administrator of the US EPA in legislative hearings on the US *Toxic Substances Control Act of 1976*, see <http://www.epa.gov/aging/press/epanews/2009/2009_1202_2.htmh>.

18.1.1 Historical aspects

The development of regulatory risk assessment approaches became prominent in the United States in the 1980s but quantitative risk assessment dates to 1976 when the brief and generic EPA guidelines for cancer risk assessment were published (Hrudey 1998). Benner (2004) has reviewed the history of the use of risk assessment by the US EPA, including a summary of the US EPA standard-setting processes and a discussion on why US air and water quality standards may have developed differently from those in other parts of the world.

One important landmark was a Supreme Court decision in 1980, when an occupational safety and health administration (OSHA) standard for exposure to benzene in the workplace was struck down. The policy had been aimed at reducing exposure as far as technologically possible but did not consider whether the existing concentration posed a significant risk to health. The majority of the court concluded that under the legislation, OSHA could only regulate if benzene posed a significant risk of harm. While 'whose significant risk of harm' was not defined, the decision highlighted that some form of quantitative risk assessment was required as the basis for deciding whether the risk was large enough to warrant regulation (NRC 1994).

Following from this judgement, Congress instructed FDA to have the National Research Council (NRC) appraise federal efforts to use risk assessment in 1981.

Drawing on work done previously by the Environmental Protection Agency, the Food and Drug Administration, the Occupation Safety and Health Administration, International Agency for Research on Cancer and the National Cancer Institute, the NRC report (1983) recommended a risk assessment on specific definitions of risk without recommending specific methods for the conduct of risk assessment.

Two key recommendations of the 1983 report were:

- A clear conceptual distinction between risk assessment and risk management should be maintained. However, it was recognised it was not necessary nor advisable for a physical separation of the two activities.
- The scientific basis for risk assessment should be detailed along with default options. It was intended that the guidelines should be flexible and allow departures from the defaults if there was appropriate data to indicate that the default option was not appropriate.

The NRC committee did not recommend a specific methodology for risk assessment but noted there should be opportunities for continuing review of the science underlying the guidelines and of the associated default options (NRC 1994). The report acknowledged the critical role of science policy judgements and that these must be distinguished from scientific facts.

The Office of Science and Technology Policy brought together scientists from regulatory agencies, the National Institutes of Health and other federal agencies. This body reviewed the scientific basis of risk assessment of chemical carcinogens and adopted the framework for risk assessment proposed by the NRC. Only the Environment Protection Agency (US EPA) adopted a specific set of guidelines for carcinogen risk assessment. The US EPA carcinogenic HRA guidelines were initially issued in 1986, and updated in 2005, after a decade of consultation. The agency has subsequently gone on to issue guidelines for other adverse health effects (mutagenicity, developmental toxicity, effects of chemical mixtures, reproductive risk, exposure, etc. (see Chapter 9).

An important step in the application of these methodologies to regulatory decision making was the US EPA's adoption of risk assessment to guide decisions at major contaminated sites. It went on to apply risk assessment methodologies to decisions regarding pesticide residues in food, carcinogenic contaminants of drinking water supplies, industrial emissions of carcinogens to surface waters, and specified industrial chemicals (NRC 1994).

The linearised multi-stage model using upper-bound estimates initially underpinned US regulatory risk assessment of carcinogens. It has been labelled as 'one of the most conservative models used in QRA' (IEH 1999b). In 1996, the US EPA proposed changing from the linearised multi-stage approach to a benchmark dose approach as their default model (US EPA 1996a), and this change has been reflected in more recent guidance (US EPA 2005a; NRC 2008).

The development of US EPA risk assessment policies has also been given an introspective flavour, with the publication of an internal assessment of EPA risk assessment policies and practices by staff in the risk assessment task force (US EPA 2004a).

18.1.2 Risk Assessment Guidance for Superfund (RAGS)

An important stimulus to the development of EHRA guidance in the US was the passing of the *Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA)* that established the Superfund for contaminated site clean-up. This legislation addressed assessment and remediation of contaminated sites in the US. It established a series of guidance documents, called the RAGS series (see <http://www.epa.gov/oswer/riskassessment/ragsa>):

- RAGS Part A Vol. I (1989) – *Human health evaluation manual* contains an overview and general HRA guidance
- RAGS Part A Vol. III (2001) – policies and guiding principles on the application of probabilistic risk assessment (PRA) methods to human health and ecological risk assessment
- RAGS Part B – guidance on using EPA toxicity values and exposure information to derive risk-based preliminary remediation goals (PRGs) for a Superfund site
- RAGS Part C – guidance on the human health risk evaluations of remedial alternatives that are conducted during the feasibility study, during selection and documentation of a remedy, and during and after remedy implementation
- RAGS Part D – guidance on standardised planning, reporting and review of Superfund risk assessments
- RAGS Part E – supplemental guidance to address human health risk related to dermal exposures
- RAGS Part F – supplemental guidance to address human health risk related to inhalation exposures.

18.1.3 Cumulative and aggregate risk assessment guidance

The passage of the *Food Quality Protection Act of 1996* required US regulatory agencies to develop specific risk assessment methodologies that could address exposures to chemicals from multiple environmental sources (termed 'aggregate' risk assessment) and to consider where simultaneous or consecutive exposures to different chemicals in an environmental mixture could result in health risks additional to, or compounded by, these multiple exposures (termed 'cumulative' risk assessment). These concepts are discussed in Chapters 5 and 12.

18.1.4 Protection of children

The Food Quality Protection Act (FQPA) also included special provisions to enhance protection of children from pesticide residues in foods. It followed publication of an NRC review of the safety of pesticides in infants and children (NRC 1993). The FQPA required the US EPA to use an extra tenfold safety factor in risk assessments to take into account potential pre- and postnatal developmental toxicity, taking into consideration the completeness of the data with respect to exposure and toxicity to infants and children. It only allowed this additional safety factor to be removed where reliable data established that it was not necessary in order to provide protection for infants and children.

18.1.4 Current guidance

The US EPA has published extensive guidance on EHRA in recent years. Most of these have been specifically cited elsewhere in this enHealth document. Listing of the key guidance documents in chronological order (other than the RAGS documents listed above) shows the development of guidance in the management of health risks associated with pesticides, soil air and water contaminants, carcinogens, and reproductive and neurological toxicants:

1983 Good laboratory practice standards: toxicology testing
Pesticide programs: good laboratory practice standards
Technical assistance document for sampling and analysis of toxic organic compounds in ambient air

1988 Seven cardinal rules of risk communication

1989 First version of the *Exposure factors handbook* (updated in 2009)

1990 Compendium of methods for the determination of air pollutants in indoor air

1992 Guidelines for exposure assessment
Guidance on risk characterisation for risk managers

1994 Quality assurance handbook for air pollution measurement systems
Methods for derivation of inhalation reference concentrations (RfCs) and application of inhalation dosimetry

1995 Guidance for risk characterisation
The use of the benchmark dose approach in health risk assessment

1996 Draft revision to the guidelines for carcinogenic risk assessment

1997 Guiding principles for Monte Carlo analysis

1998 National primary drinking water regulations: interim enhanced surface water treatment

1999 Compendium of methods for the determination of inorganic compounds in ambient air

2000 Supplementary guidance for conducting health risk assessment of chemical mixtures

Benchmark dose technical guidance document

Methodology for deriving ambient water quality criteria for the protection of human health

2001 General principles for performing aggregate exposure and risk assessments

Preliminary cumulative risk assessment of the organophosphorus pesticides

2002 A review of the reference dose and reference concentration processes Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity

Draft guidance for evaluating the vapor intrusion pathway from groundwater and soils (subsurface vapor intrusion guidance)

2003 Framework for cumulative risk assessment
Considerations in risk communication: A digest of risk communication as a risk management tool Developing potency factors for pesticide mixtures: biostatistical analyses of joint dose–response

2005 Guidelines for carcinogenic risk assessment

Guidance on selecting age groups for monitoring and assessing childhood exposures to environmental contaminants

Supplemental guidance for assessing susceptibility from early-life exposure to carcinogens

2006 A framework for assessing health risks of environmental exposures to children

Approaches for the application of physiologically based pharmacokinetic (PBPK) models and supporting data in risk assessment

2007 Guidance for evaluating the oral bioavailability of metals in soils for use in human health risk assessment

Users guide for the integrated exposure uptake biokinetic model for lead in children (IEUBK)

Considerations for developing a dosimetry-based cumulative risk assessment approach for mixtures of environmental contaminants

2009 *Exposure factors handbook: 2009 update*

Recommended toxicity equivalency factors (TEFs) for human health risk assessment of dioxin and dioxin-like compounds

2010 *Development of a relative potency factor (RPF) approach for polycyclic aromatic hydrocarbon (PAH) mixtures*

A more extensive listing of US EPA risk assessment guidance documents may be found on the EPA website at <http://www.epa.gov/risk/health-risk.htm>.

18.2 UNITED KINGDOM PROGRAMS

18.2.1 Historical aspects

In 1996, the Government/Research Councils Initiative on Risk Assessment and Toxicology was established to review current practices for managing risks to health from chemicals and to promote improved risk assessment decision making. The agencies involved covered a wide variety of risks including those from food contaminants and additives, agricultural pesticides, biocides, veterinary products, occupational exposures, consumer products, air quality, water quality, land quality and human medicines. As a result of the deliberations of the initiative, a four-stage process of risk assessment was proposed consisting of:

1. Identifying the properties of chemicals that can lead to adverse (toxic) health effects (hazard identification)
2. Obtaining quantitative information about the hazard including, where possible, information on dose–response relationships (hazard characterisation)
3. Assessing exposure to the chemical (exposure assessment)
4. Comparing exposure and hazard information (risk characterisation).

The initiative described the variety of risk assessment practices used in different government departments as a step towards establishing a common framework for the procedures used, identifying the major areas of uncertainty and weakness in current risk assessment processes, and establishing where these risk assessment processes might benefit from harmonisation across departments.

There were substantial differences between departments and agencies in the degree of caution incorporated into risk assessment for factors such as:

- the size of uncertainty factors applied when there are thresholds for toxic effects
- the use of mathematical approaches for effects with or without a threshold – mathematical modelling (using probit analysis of the best available dataset) may be used as one component of the risk assessment
- the treatment of data gaps and efficiency
- the degree of protection for general population exposures compared with occupational exposures
- the degree of conservatism built into worst-case exposure estimates (IEH 1999b)
- the approaches to the assessment of genotoxic carcinogens. The UK has tended to use a qualitative weight of evidence approach to the evaluation of carcinogenic risk and has tended to avoid the use of mathematical approaches for quantitatively assessing risks from genotoxic carcinogens and they are rarely, if ever, carried out by UK regulatory agencies. The UK Department of Health's Committee on Carcinogenicity did not support the routine use of quantitative risk assessment (QRA) for chemical carcinogens (IEH 1999b).

Strategies for dealing with variability within the human population as a result of factors such as age, sex, pregnancy status, health status, lifestyle and genetic factors were important parts of the process.

Physiologically based pharmacokinetic (PBPK) modelling was considered to help to highlight and reduce the uncertainties of estimating the dose of an agent the body or parts of the body may receive after exposure.

18.2.2 Recent guidance development

Pollard et al. (2002) reviewed environmental risk management in the UK. General guidance in 2000 was summarised by the Department of the Environment, Transport and the Regions, the Environment Agency and the Institute for Environment and Health (DETR 2000). The principal sources of EHRA guidance in the UK have arisen through cooperative programs of the Committee on Toxicity (COT), Institute for Environmental Health (IEH) and the Interdepartmental Group on Health Risks from Chemicals (IGHRC). The Health and Safety Executive (HSE) has published a number of pamphlets and reports on chemical safety issues, mainly directed at workplace risk management. The HSE approach to risk assessment and decision making is summarised in HSE (1999).

To date, the IGHRC and its precursors has published 14 reports covering topics such as:

- general risk assessment methodology
- exposure assessment in EHRA
- weight of evidence
- route-to-route extrapolation
- uncertainty analysis
- PBPK modelling
- carcinogenic risk assessment – the IGHRC document discusses frameworks for carcinogen assessment while other documents on carcinogenicity and mutagenicity are published by the UK Committees on Toxicity, Carcinogenicity and Mutagenicity of chemicals in food, consumer products and the environment (COT, COC and COM).

18.3 INTERNATIONAL PROGRAMS DEALING WITH CHEMICAL SAFETY

International programs on chemical safety provide an opportunity for national government agencies, industry groups and non-governmental organisations (NGOs) to cooperate in the development of programs for the assessment and risk management of hazardous chemicals.

18.3.1 The Rio Earth Summit and the IFCS

The United Nations Conference on Environment and Development (UNCED), Rio de Janeiro, 3–14 June 1992 (the Rio Earth Summit) was an important milestone in the development of cooperative international programs on chemical safety. The conference developed a number of themes for effective global management of chemicals, and established, under its Agenda 21 Chapter 19 on 'Environmentally sound management of toxic chemicals, including prevention of illegal international traffic in toxic and dangerous products', a set of principles for developing chemicals management programs, including:

- expanding and accelerating international assessment of chemical risks
- harmonisation of classification and labelling of chemicals
- information exchange on toxic chemicals and chemical risks
- establishment of risk reduction programs
- strengthening of national capabilities and capacities for management of chemicals
- prevention of illegal international traffic in toxic and dangerous products.

The Intergovernmental Forum on Chemical Safety (IFCS) (<http://www.who.int/ifcs/en>) was established as a direct result of the Rio Earth Summit, and met six times from 1994 to 2008 to monitor the key programs in chemical safety (there were also three inter-sessional planning conference, one of which was convened in Canberra in 2006 hosted by the Australian Government). The Bahia Declaration (from IFCS Forum III) reaffirmed the commitment of governments to principles set at the 1992 Rio Summit.

The more significant outcomes of the IFCS sessions were:

- establishment of the Inter-Organization Programme for the Sound Management of Chemicals (IOMC) and later, the Strategic Approach to International Chemicals Management (SAICM)

- cooperation with chemicals management programs of the OECD and UNEP, leading to the establishment of the 2001 Stockholm Convention on Persistent Organic Pollutants (POPs), and the 1998 Rotterdam Convention on the Prior Informed Consent Procedure for Certain Hazardous Chemicals and Pesticides in International Trade.

18.3.2 International Programme on Chemical Safety (IPCS)

The IPCS is a tripartite program of the World Health Organization (WHO), International Labour Organization (ILO) and the United Nations Environment Programme (UNEP). It collaborates with other international programs (e.g. IOMC, JMPR, JECFA) in publishing monographs on chemical risk assessment. These include:

- Environmental health criteria (EHC) series
- Concise international chemical assessment documents (CICADs).

Important initiatives of the IPCS in the late 1990s were the development of programs reviewing the methodology of HRA (<http://www.who.int/ipcs/methods/en>) and the IPCS Harmonization Project (<http://www.who.int/ipcs/methods/harmonization/en/index.html>), which aimed to understand the basis for different approaches to EHRA around the world. The program objectives are to facilitate global harmonisation of approaches to risk assessment by increasing understanding and developing basic principles and guidance on specific chemical risk assessment issues. This would enable more efficient use of resources and promote greater consistency among assessments.

The project has issued several progress reports and monographs, which describe the project milestones and how the work is organised:

- Harmonization Project information brochure (2nd edition)
- *Strategic plan 2005–2009*
- Project Steering Committee meeting reports
- Stocktake of the project, including a risk assessment toolkit on how the products are used

The key project activities to date have been to address:

- combined exposures to multiple chemicals
- cancer risk assessment
- non-cancer risk assessment
- exposure assessment
- exposure assessment and risk assessment terminology
- reproductive/developmental toxicity terminology
- mutagenicity testing
- PBPK modelling
- skin sensitisation risk assessment
- chemical-specific adjustment factors
- immunotoxicity.

18.3.3 The REACH program

The Registration, Evaluation, Authorisation and Restriction of Chemical substances program (REACH) was established in 2006 as a new European Community program for regulating chemicals and their safe use.

> *The aim of REACH is to improve the protection of human health and the environment through the better and earlier identification of the intrinsic properties of chemical substances. At the same time, innovative capability and competitiveness of the EU chemicals industry should be enhanced. The benefits of the REACH system will come gradually, as more and more substances are phased into REACH.*
>
> See <http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm>.

The REACH regulation gives greater responsibility to industry to manage the risks from chemicals and to provide safety information on the substances. Manufacturers and importers are required to gather information on the properties of their chemical substances, which will allow their safe handling. The REACH information is centrally registered in a database run by the European Chemicals Agency (ECHA) in Helsinki. ECHA manages the databases necessary to operate the system, coordinates the in-depth evaluation of suspicious chemicals, and maintains a public database in which consumers and professionals can find hazard information.

There has been an extensive literature addressing REACH data requirements, their impacts on industry and regulators, and how significant data gaps may undermine the ability of industry to achieve compliance. Williams et al. (2009) have written a thoughtful summary of the historical development of REACH, and expanded on some of the onerous requirements in the legislation.

The REACH legislation has not been without its critics, nor those who argue that it should go further.

In the former category, Foth and Hayes (2008) also summarises REACH requirements. Their review notes the substantial efforts necessary to plug data gaps. However, it is also noted that the REACH legislation is likely to stimulate further research on risk assessment methodology and how toxic chemicals may be categorised. Ahlers et al. (2008) have also emphasised the challenges posed by REACH, especially on industry, which will carry the responsibility of generating and assessing the required data. Schaafsma et al. (2009) went even further and argued that REACH would fall short of its objectives if risk assessment was unable to move away from a labour-intensive and animal-consuming approach to risk assessment, towards more pragmatic assessments that group chemicals and combine assessments of hazard and exposure.

In the latter category, Santillo and Johnston (2006), writing before the legislation had been finalised, argued that a more defensible position for REACH would be to use less toxic chemicals if available, and promote the development of safer alternatives if not.

Also in the latter category, Ruden and Hansson (2010) suggest that the following six steps are needed to improve REACH:

1. Prioritisation and waiver criteria must be clarified. This means that sufficient data must be available to make initial hazard assessments on as many substances and toxicological endpoints as possible.
2. Data requirements for substances imported in quantities greater than 1 tonne/yr must be increased so that they align with current requirements for substances imported at greater than 10 tonnes/yr.
3. Testing protocols need to consider resource limitations and animal welfare issues (reduction in animal testing) but that still satisfy information requirements.
4. Substitution of high-risk chemicals should be promoted.
5. The control of substances incorporated in the articles of the legislation should be addressed.
6. Uncertainties in the process must be acknowledged in particular, a lack of data should be reported and incorporated into the risk management process.

18.3.4 Industry initiatives

Industry groups have often found it to be in their interest to fund programs that enhance the scientific basis of risk assessment. Some of these initiatives are:

- European Centre for Ecotoxicology and Toxicology of Chemicals – ECETOC (see <http://www.ecetoc.org>)
- International Life Sciences Institute – ILSI (see <http://www.ilsi.org>)
- Chemical Industry Institute of Toxicology – CIIT (see <http://www.thehamner.org/institutes/ciit>)
- Toxicology Excellence for Risk Assessment – TERA (see <http://www.tera.org>)
- Alliance for Risk Assessment – ARA (see <http://www.allianceforrisk.org>).

Some of these organisations sponsor taskforces, workshops and symposia that bring together leading scientists in the field of risk assessment. They may support publication of high-quality science in their own journals or websites, and in the general peer-reviewed scientific literature. They also provide opportunities for industry scientists to interact with regulatory scientists in the formulation of risk assessment guidance and policies.

A typical example of the collaboration between industry, government and academics is a review of the use of MoA and life-stage impacts on the human relevance of animal-based toxicity tests (Seed et al. 2005).

18.4 PROGRAMS FOR DEVELOPING ALTERNATIVES TO ANIMAL TESTING

Animal welfare has always been a controversial issue in the generation of data necessary for human health risk assessment. The rise of alliances concerned with animal welfare has prompted both government agencies and NGOs to develop approaches to toxicity testing that do not require *in vivo* dosing of test animals.

One such agency established by the European Commission to develop *in vitro* alternative toxicity tests is the European Centre for the Validation of Alternative Methods – ECVAM (see <http://ecvam.jrc.ec.europa.eu>).

The mission given to ECVAM by the EC when it was established under EC legislation in 1991 was to:

- coordinate the validation of alternative test methods at the European Union level
- act as a focal point for the exchange of information on the development of alternative test methods
- set up, maintain and manage a database on alternative procedures
- promote dialogue between legislators, industries, biomedical scientists, consumer organisations and animal welfare groups, with a view to the development, validation and international recognition of alternative test methods.

ECVAM has established working groups and taskforces covering most areas of animal testing:

- systemic toxicity – acute and chronic toxicity, neurotoxicity and immunotoxicity
- topical toxicity – skin and eye irritation
- sensitisation – skin and respiratory sensitisation
- genotoxicity and carcinogenicity
- reproductive toxicity – endocrine disruption
- toxicokinetics – biokinetics, blood–brain barrier, PBPK modelling.

The brief for ECVAM is similar to that for groups established under the US National Toxicology Program (NTP) to develop alternative toxicity testing strategies. The two groups, which work together under the NTP banner (see <http://iccvam.niehs.nih.gov>), are:

- NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)
- Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM).

The contributions of NICEATM and ICCVAM to risk assessment methodology, and the challenges still to be faced in the future, have been summarised by Birnbaum and Stokes (2010).

18.5 INTEGRATION OF HUMAN HEALTH AND ECOTOXICOLOGY

The assessment of risks to environmental flora and fauna (ecological risk assessment) is also a part of environmental health regulation, although not a primary focus of this enHealth guidance on EHRA, which is directed towards human health risk assessment; however, there have been calls for better integration of ecological and human health risk assessment.

Suter et al. (2005) have reviewed a framework developed by the IPCS that integrates risks to human health and risks to non-human organisms and ecosystems. They note that:

> ... WHO's framework recognises that stakeholders and risk managers have their own processes that are parallel to the scientific process of risk assessment and may interact with the risk assessment at various points, depending on the context. Integration of health and ecology provides consistent expressions of assessment results, incorporates the interdependence of humans and the environment, uses sentinel organisms, and improves the efficiency and quality of assessments relative to independent human health and ecological risk assessments. The advantage of the framework to toxicologists lies in the opportunity to use understanding of toxicokinetics and toxicodynamics to inform the integrated assessment of all exposed species.

Appendix 1: Guidance on the evaluation of toxicity studies

This appendix, in conjunction with the general guidance in Chapter 9, is intended to provide additional guidance to toxicologists who need to evaluate a package of toxicity tests to inform an EHRA.

A1.1 Study protocol and design

A1.1.1 Dosing regimen

The purpose of toxicity studies is to detect valid biological evidence for any toxic or oncogenic potential of the substance being investigated. Therefore, protocols should maximise the sensitivity of the test without significantly altering the accuracy and interpretability of the biological data obtained. The dose regimen has an extremely important bearing on these two critical elements.

Since the determination of dose responses for any observed effects is one of the objectives of repeat-dose studies, at least three dose levels are normally required, as well as controls. US EPA guidelines allow a limit dose of 1,000 mg/kg in chronic and sub-chronic studies. If this dose produces no observed toxic effects, and if toxicity is not expected based on data on structurally related compounds, then a full study using three dose levels might not be considered necessary. Ideally, the dose selection should maximise the detection of potential dose–response relationships and facilitate the extrapolation of these to potential hazards for other species including humans. The largest administered dose should not compromise biological interpretability of the observed responses. For example, it is generally considered that the upper dose should not:

- cause a body weight decrement from concurrent control values of greater than 10–12 per cent
- exceed 5 per cent of the total diet in a dietary study because of potential nutritional imbalances caused at higher levels
- produce severe toxic, pharmacological or physiological effects that might shorten duration of the study or otherwise compromise the study results
- alter survival in a carcinogenicity study in a significant manner due to effects other than tumour production.

The International Life Sciences Institute (ILSI) Risk Sciences Working Group on Dose Selection has published its deliberations on the selection of doses in chronic rodent bioassays (Foran & ILSI Risk Sciences Working Group on Dose Selection 1997).

Although it has been argued that responses observed at doses far in excess of levels experienced under real or potential exposure conditions legitimately fall within the classical dose–response concept, there are valid scientific concerns that such doses introduce a further layer of uncertainty into the already difficult task of evaluating animal dose responses and the assessment of their relevance to human hazard identification and risk (Paynter 1984). High doses that overwhelm normal mechanisms for metabolism, detoxification or excretion, or produce severe tissue damage (i.e. necrosis, demyelination) can make interpretation difficult or lead to inappropriate conclusions about the extent of the hazard.

With respect to selecting the low dose, it is commonly accepted that the lowest dose should not produce any evidence of toxicity (i.e. allows the establishment of a 'no observed adverse effect level' – NOAEL).

A1.1.2 Dosing route

For repeat-dose studies, the most convenient route of administration is by dietary admixture. However, depending on the possible route of exposure of the public or occupationally exposed workers to a chemical or an environmental contaminant, it may need to be investigated by the dermal or inhalation route.

For dermal exposure the material, in a suitable vehicle, is applied to the clipped skin of rats, rabbits or guinea pigs. OECD *Test guideline* TG410 recommends even application to an area representing about 10 per cent of the total body surface area. The site is generally occluded with polyethylene sheeting and gauze patches (or semi-occluded) to prevent dislodgment of material and oral ingestion by the animal, which could affect the validity or usefulness of the study. For volatile or semi-volatile materials, application and covering procedures should minimise the possibility of evaporation. Useful chapters or sections on dermal toxicity testing may be found in textbooks on toxicology (e.g. Derelanko & Hollinger 1995; Hayes 1994), and in more recent US EPA RAGS-E guidance (US EPA 2004b).

The surface area of the respiratory membrane is estimated at approximately 50–100 square metres in the normal adult compared with the estimated area of the small intestine at 250 square metres (Guyton 1991) and much more air (about 5,000 times, by volume) is inhaled each day than food or water is ingested (McClellan & Henderson 1989). Thus, exposure to airborne material is potentially greater than via dermal or oral exposure. Airborne material can be gases or vapours, liquid droplets or solutions, aerosols (solid and vapour components), or dry fibres or powders. As a consequence, to conduct inhalation toxicity studies, mechanisms needed to deliver chemicals to a test chamber (in a form that can be inhaled) are quite complex, particularly when coupled with the need to include measuring devices which can establish particle size, concentration and form of the material in the exposure chamber. Furthermore, many factors can influence the inhalation, deposition and retention of inhaled materials in the respiratory tract. Thus, the conduct of inhalation studies is considerably more complex than equivalent studies by the dietary or dermal routes.

Of critical importance, in both the conduct and assessment of such studies, in the need to establish what portion of the material delivered to the exposure chamber was in a respirable form. In addition to standard toxicology texts, some useful specific references on inhalation toxicology include McClellan and Henderson (1989), Mohr et al. (1988) and Salem (1987).

A1.2 Study findings – physiological, pharmacological or toxic?

In conducting a hazard assessment, the evaluator needs to determine whether effects seen in studies are physiological, pharmacological or toxic.

Responses produced by chemicals in humans and experimental animals may differ according to the quantity of the substance received and the duration and frequency of exposure, for example, responses to acute exposures (a single exposure or multiple exposures occurring within 24 hours or less) may be different from those produced by sub-chronic and chronic exposures. Not all observed responses within a study, irrespective of exposure duration or frequency, will represent toxicity per se. They may encompass a range of effects from physiological through pharmacological and toxic manifestations.

Although it sometimes may be difficult to make a clear distinction between these responses, an attempt to do so should be made. If an evaluator is uncertain of the type or the biological significance of a response, they should not hesitate to obtain competent advice for resolving the uncertainty. It is essential that all relevant toxicity endpoints (statistically and/or biologically significant) be identified for consideration when evaluating data for the presence or absence of non-toxic levels.

Physiological responses vary within limits that are in accord with the normal functioning of a living organism. Examples of such response are the usual respiratory and pulse rate increases associated with increased physical activity, systemic changes associated with normal pregnancy, and those associated with homeostatic mechanisms. These variable factors are not important toxicity endpoints in sub-chronic and chronic exposure studies unless their fluctuations are abnormally altered by a dose regimen. If such alterations occur at a particular dose or are part of a dose–response relationship, they should be correlated with other toxicity endpoints that may be present.

Pharmacological responses are altered physiological functions arising from interaction of a substance with a cellular receptor site. They are reversible, and are usually of limited duration following removal of the stimulus. While some of these responses may be undesirable under certain circumstances, they are distinguished from toxic (adverse) responses by generally not causing injury. An example of this type of response is the increased activity of the hepatic cytochrome P-450 containing mono-oxygenase systems (enzyme induction) caused by exposure to many pesticides, industrial chemicals and drugs (noting, however, that while not a direct adverse effect, a cytochrome P-450 inducer can, for example, alter hormonal homeostasis and effect tumour promotion, or increase an organism's susceptibility to other chemical exposures).

Toxic responses may be reversible or irreversible, but are distinguished from other types of responses by being injurious and therefore adverse and harmful to living organisms or tissues. A chemical that causes a physiological or pharmacological effect may produce a toxic response if the exposure is prolonged or if the dose is increased beyond a certain level.

The reversibility or otherwise of such responses may also depend on these two factors. The reversibility or irreversibility of a histopathological change will depend on the ability of the injured organ or tissue to repair/regenerate. For example, the liver has a relatively great ability to regenerate and many types of injury to this organ are reversible. By contrast, differentiated cells of the central nervous system are not replaced and many injuries to the CNS are irreversible.

A1.3 Assessing the quality of the data characterising the hazard

The following considerations address the acceptability of experimental studies and the documentation provided.

1. The adequacy of the experimental design and other experimental parameters, including: the appropriateness of the observational and experimental methods; frequency and duration of exposure; appropriateness of the species, strain, sex and age of the animals used; the numbers of animals used per dosage group; justification of dose, route and frequency of dosing; and the conditions under which the substance was tested.
2. There are many guidelines to the generation of scientifically valid data that concern good experimental design, laboratory practice and reporting, such as OECD and US EPA guidelines, and accepted codes of good laboratory practice (GLP) (OECD 1982; US EPA 1983). They can be useful as aids in determining report and data acceptability. However, the evaluator needs to make a judgement about how well the study, in its totality, facilitates the identification of potential adverse effects, or lack thereof, of the substance being evaluated, rather than how precisely it fits a prescribed test guideline or 'recipe'. The experience of senior evaluators can be helpful in resolving concerns about acceptability of study conduct and/or reporting.
3. The competency and completeness of the conduct and reporting of the study.
4. The effects of modifying factors that may result in major inequalities between control and test animals.

This qualitative consideration has more to do with the evaluation and interpretation of data than with acceptability of documentation. It is placed here because determining the factors that may have a major influence on toxicological data needs to be made prior to analysing the data. There are many factors influencing the responses of experimental animals to experimental treatment; some of these are discussed by Doull (1980). Some influences may be quite subtle, as exemplified by studies performed by Thompson et al. (1982), in which it was noted that the onset of acute pulmonary oedema in rats being used in immune hypersensitivity studies was sudden and seasonal. Circadian rhythms and seasonal physiological variations can subtly influence experimental results. Such factors influencing animal responses can be troublesome when their effects are confused with or misinterpreted as toxic responses to treatment. For further discussion of environmental effects on experimental parameters see Herrington and Nelbach (1942).

The acceptability of reports and other technical information is primarily a scientific judgement. Therefore, the rationale for rejecting a hazard assessment study should be succinctly stated in the evaluation document.

A1.3.1 Analysing and evaluating toxicity studies

Useful guidance documents for evaluating data and conducting assessments include the IPCS *Environmental health criteria* (EHC) monographs, namely EHC 6, 70, 104 and 141 (WHO 1978; 1987; 1990; 1992).

A1.3.2 Analysing and evaluating major study parameters

Not all observed effects of test substances are toxic effects. Rather, they may be adaptive (e.g. liver enzyme induction leading to some hepatic enlargement) or may be a manifestation of a pharmacological effect (e.g. in an animal colony suffering from various low-grade infections, an antibiotic will lower leucocyte counts in treated animals relative to controls obviously it is not appropriate to describe this as a leukopaenic effect of the chemical).

Concurrent control groups should always be used notwithstanding the value of historical control ranges in carconogenicity studies. It is generally not appropriate to rely on statistical comparisons with historical controls since the incidence of spontaneous lesions can vary significantly over time (and even between concurrent randomised control groups). Controls must be age-matched because some forms of toxicity represent no more than acceleration and/or enhancement of age-related changes. Examples of pathological changes in aged rats that may be affected by compound administration include chronic progressive glomerulonephropathy, peripheral nerve degeneration, amyloidosis and various neoplasms.

Using both non-treated and vehicle-control groups helps to assess effects due to vehicle or excipients. When a vehicle is used to deliver the doses of the agent being investigated (e.g. a lipophilic agent delivered in corn oil), the need for vehicle-treated controls is paramount. Since some parameters can be affected by animal handling (e.g. serum ALT was raised in mice that were grasped around the body compared with unhandled or tail-handled mice, see Swaim et al. 1985), control animals should be treated in the same way as test animals.

Control animals must receive as much attention during the analysis and evaluation process as do the treated ones. Any untreated animal or group may exhibit some signs of abnormality or drift from the norm for that species or strain. Because of the possibility that statistically significant differences between treated and control groups are the result of abnormal values among the controls, such differences should usually be dose-related and should delineate a trend away from the norm for that stock of animals, if they are to be indicative of a true compound-related effect.

Historical control data may be useful when evaluating the acceptability of the 'normal' data obtained from control groups (Haseman et al. 1984; Paynter 1984; Sumi et al. 1976; Tarone 1982). Any departure from the norm by the control groups should be taken into consideration, especially during the conduct of any statistical analysis. The finding of consistent departures from the norm in control groups may necessitate investigation of the source of the animals.

Ideally, historical control data should be taken from the same laboratory undertaking the study, collected using the same strain, age and sex of animals obtained from the same supplier, and only include those studies conducted within a 2–3-year span on either side of the study under review, for retrospective studies. Important information to consider includes identification of study methodology, which could have affected the results, such as pre-sampling conditions (e.g. fasting or non-fasting, assay methodology for study parameters, histopathological criteria for lesion identification, time of terminal sacrifice).

Weil and McCollister (1963) analysed toxicity endpoints, other than oncogenicity, from short- and long-term tests and concluded that only a relatively small number of endpoints were effective in delineating the lowest observed adverse effect level (LOAEL) in such tests. Body weight, liver weight, kidney weight and liver pathology delineated the LOAEL in 92 per cent of test chemicals in

sub-chronic studies and 100 per cent in chronic studies. To reach 100 per cent efficiency in sub-chronic studies, renal and testicular histopathology had to be included. Heywood (1981) surveyed the toxicological profiles of 50 compounds in rodent and non-rodent species and confirmed these observations.

A1.3.3 Mortality/survival

Reasonable efforts should be made to determine the cause or likely cause of individual deaths. The evaluation of pathological lesions or morphological changes in belatedly observed deaths is frequently complicated by post-mortem autolysis. The separation of deaths caused by factors unrelated to exposure to the test agent (e.g. acute or chronic infections, age or disease-related degenerative processes, anatomical abnormalities, negligent handling or accident) from toxicity-induced deaths is important. All data relating to moribund or dead animals during their study life, as well as the results of post-mortem examinations, should be scrutinised in an attempt to make this distinction. Note that US EPA guidelines state that the highest dose used in sub-chronic studies with non-rodents should not produce an incidence of fatalities that would prevent meaningful evaluation.

Analysis of mortality requires more than a statistical treatment of incidence at termination of a study. Survival/mortality data can be influenced by factors other than the test substance. Changes in the protocol during the course of a study can complicate the analysis, such as alterations in dosage levels can produce a confusing mortality pattern.

Any unusual mortality pattern should be explained by the test laboratory on biological or toxicological grounds. If overall mortality is high (i.e. significantly greater than expected for the particular colony and strain) for any repeat-dose study, or for a particular group within a study, a credible explanation should be provided.

An evaluation of mortality patterns within each group may indicate that mortality is clustered early or late in the course of the study, is intermittent and scattered throughout the duration of the study, or has a higher incidence in one sex than in the other. The analysis of the cause of individual deaths will aid in determining the toxicological significance of these various patterns. Early deaths within treated groups may just reflect deaths of the more susceptible animals in the test population. Alternatively, it may indicate changes in compound intake per unit body weight, in those experiments in which the quantity of test substance in the diet is kept constant. Relative to body weight, young rats ingest more food than older rats and therefore ingest relatively more of the test substance. Early deaths may therefore be the result of the higher exposure, on a mg/kg/d basis, of young animals compared with older animals.

Deaths that are clustered at a specific time period may reflect a spontaneous, epidemic disease situation of limited duration. High mortality associated with infectious agents in treated groups, in the absence of such evidence in the concurrent control group, could indicate an immunosuppressive action of the chemical being tested.

The effect of dietary intake on mortality needs to be considered. A compound administered in the diet may make the laboratory food more or less palatable, may have a pharmacological stimulant or depressant effect on appetite, or may affect the partitioning of the nutrients in the food. Likewise, decreased water consumption (e.g. in the case of an unpalatable compound administered in the water) will lead to reduced food consumption. These effects may significantly influence longevity since it has been clearly shown in animal species that long-term dietary restriction very significantly increases life span (e.g. Tucker 1979). Conversely, excessive *ad libitum* intake of highly nutritious diets can reduce life span compared with the expected average life span for an animal species/strain. To date, regulatory authorities have not come to any decision on recommending restricted diets versus *ad libitum* feeding in toxicity study guidelines. Some useful references on this topic include Keenan et al. (1998), Klinger et al. (1996), Masoro (1992) and Thurman et al. (1995).

A1.3.4 Clinical observations

Adverse clinical signs (gross observations) noted during the exposure period may correlate with toxicity endpoints or disease processes. These can be used as supportive evidence for dose–response relationships and may play a role in determining the NOEL/NOAEL. However, not all adverse clinical signs will correlate with pathological or morphological changes in organs or tissues. Some will be caused by biochemical or physiological effects, such as incoordination, muscle twitching, tremor or diarrhoea, which may indicate acetylcholinesterase inhibition without any morphological changes being evident in nervous tissue.

Many of these qualitative signs can be counted, scored for intensity and tabulated as incidences. However, statistical analysis is of limited value. The evaluator must rely on the number of individuals per group exhibiting signs of a particular type, as well as the intensity of the responses, to gain an impression of a dose–response relationship.

Clinical observations, such as palpable tumours or those that might be associated with neoplasia (e.g. haematuria, abdominal distension or impaired respiration), may be useful in defining the time a tumour was first suspected as being present. Such signs might help evaluate decreased tumour latency in long-term rodent studies. They may also aid in determining cause of death. A statement of the correlations, or the lack thereof, between clinical signs and specific toxicity endpoints should be made in the evaluation.

Useful information on gross behavioural observations in laboratory animals and abnormal behaviour patterns can be found in Bayne (1996).

The revised OECD test guidelines for 90-day oral toxicity studies in rodents and non-rodents (TGs 408 and 409) have placed additional emphasis on neurological endpoints, that is, studies should allow for identifying chemicals with the potential to cause neurotoxic effects, which may warrant further in-depth investigation. The reader is referred to the references cited in *Test guideline* 408 relating to neurotoxicity assessment, including sensory reactivity to stimuli of different types (auditory, visual, proprioceptive), grip strength, and motor activity. The OECD promulgated *Test guideline* TG 426 in 2007 to address neurotoxic effects during the developmental period in newborn animals.

A1.3.5 Body weight changes, food and water consumption

Body weight changes (gains or losses) for individual animals and groups of animals when compared with concurrent control changes during the course of a study are a criterion of some importance (Heywood 1981; Roubicek et al. 1964; Weil & McCollister 1963). Such changes are usually related to food intake, and analysis of one without an analysis of the other is of limited value. Weight decrement may not always be related to toxicity per se (Seefeld & Peterson 1984). Occasionally the incorporation of the test substance into the diet will reduce the palatability of the diet to many individuals in all treatment groups or to the majority of individuals in the higher dietary level groups. Food spillage needs to be considered when evaluating food palatability and compound intake. The same considerations apply if the compound is administered in drinking water.

This effect on body weight is often evidenced during the first two or three weeks of the study. Sometimes animals in the affected groups are able to accommodate to the diet and a gradual increase in group weight gain will occur (Nolen 1972). In sub-chronic studies, the lag in group weight gain may persist, even though the individual animal gains per gram of food consumed (food utilisation efficiency) are favourable after the accommodation, and produce a statistically significant difference between the affected group and the concurrent controls that is not related to toxicity of the test substance (McLean & McLean 1969). Sometimes the addition of the test substance will interact with one or more essential nutritional elements in the diet, thereby producing weight gain decrements or alterations of toxic responses (Casterline & Williams 1969; Conner & Newbern 1984; Rogers et al. 1974). This phenomenon may be encountered in sub-chronic studies and, when identified, can usually be overcome by acceptable means before a chronic study is initiated. Infrequently, control values for weight gain (at one or more time points) can be low, causing the other value to appear unusually high.

Diet composition, food and water consumption, and body weight gains per se can also have an important influence on many aspects of animal responses including shifts in metabolic, hormonal and homeostatic mechanisms (Kennedy 1969) as well as disease processes (Berg & Simms, 1960; Paynter 1984; Ross & Bras 1965; Tannenbaum 1940) and maturation (Innami et al. 1973). These variables and should be considered when unusual effects are observed in the absence of any indication of injury to organs and other vital systems.

Evaluating body weight changes and attendant effects is significantly aided by the graphical presentation of group mean body weights and food consumption versus compound consumption (on a mg/kg body weight basis). This allows quick identification of any unusual or sudden changes in gain or loss by any group.

A1.3.6 Haematological, clinical chemistry and urinary measurements

Haematological, clinical chemistry and urinary parameters are routinely measured in sub-chronic and chronic toxicity studies compliant with regulatory guidelines.

Normal biological variation in inter-animal values and their alteration in response to a variety of inputs means that evaluators will have to contend with much 'noise' in this area, and will frequently be presented with scattered, statistically significant effects in the absence of any evidence of clinically significant relationships to specific toxicity endpoints. To deal with 'noise' there is a need to examine whether the effect noted is within the normal range of variation (concurrent and historical controls). Some of these parameters can vary significantly with no clinical manifestations but others (e.g. serum potassium) have a very narrow normal clinical range and small differences can be important.

Frequently this data shows apparently 'random' changes in individual groups or, less commonly, non-dose-related trends in changes across several groups. If using historical control data as an aid to evaluation, only values produced by the identical methods from the same laboratory are valid in such comparisons. Literature values for normal ranges that do not specify the method by which they were obtained should be used with caution.

A good review of factors that can complicate the interpretation of findings in a toxicity study may be found in the *Handbook of toxicologic pathology* (Haschek et al. 2002).

To gain maximum information from enzyme determinations it is important to consider the most appropriate enzymes.

It is important that organ distribution and location of the enzyme in the cell is known. ALT (alanine aminotransferase) is found in greatest concentration in the liver in rats, even more so in dogs. AP (alkaline phosphatase, or ALP) is virtually absent from the liver in these two species, being mainly confined to the kidney, intestine and bone. CPK (creatine phosphokinase, or CK) is mainly located in skeletal and heart muscle, while AST (aspartate aminotransferase) is found in various concentrations in most organs. It is clear that CPK is the most appropriate enzyme to detect muscle damage, while changes in ALT would probably reflect some liver necrosis. Although AST is not organ-specific, it serves to confirm organ damage, especially for muscle and liver, if its activity changes in parallel with other enzymes. In dogs, AP is a sensitive test for biliary function, but in the rat it is of little diagnostic value since it is absent from the liver and principally derived from the intestines. For hepatocellular evaluation, ALT, AST, SDH (sorbitol dehydrogenase) and GLDH (glutamate dehydrogenase) are the most appropriate, while for hepatobiliary evaluation, AP, 5'-nucleotidase, GGT (gamma-glutamyl transferase) and total bilirubin are the most appropriate measurements. It is important to understand that many of these types of serum enzyme tests and urinalysis fail to detect minor injury or may reflect only transient or reversible changes. Therefore, evaluation and interpretation of the test results must be performed carefully and correlated with more specific, sensitive and reliable histopathological findings.

Sensitivity and specificity of the enzyme changes as diagnostic of organ pathology are greatly influenced by the species selected for testing (Tyson & Sawhney 1985). For example, in mammalian species, aspartate transaminase is not specific to any tissue and thereby elevated plasma AST activity may suggest damage to any one of many tissues. In contrast, alanine transaminase is relatively specific to the liver in the cat, dog, ferret, mouse and rat, whereas in primates, ALT is present in heart, skeletal muscle and liver. Plasma alkaline phosphatase measurement has been less useful in detecting liver cell necrosis in the dog, sheep, cow and rat but may be indicative of other types of liver damage, particularly those of a cholestatic nature in a number of species. It is evident that species differences are of great importance when specific clinical chemistries are selected for inclusion in toxicity studies.

When analysis and evaluation of clinical data indicate a dose–response relationship or a biologically important drift from concurrent control values, the effects observed should be correlated with other manifestations of toxicity. The evaluator should indicate whether or not a correlation could be established.

Standard veterinary (e.g. Bush 1991; Duncan et al. 1994; Evans 1996) and human clinical manuals (e.g. Henry 1984; Tyson & Sawhney 1985; Walach 1996) should be consulted for information about laboratory diagnostic tests and to assist in the evaluation of potential correlations between clinical chemistry, haematological, and urinary data with adverse effects.

The deliberations of a joint international committee, established to provide advice for clinical pathology testing of laboratory animal species used in regulated toxicity and safety studies, has published its recommendations, including those parameters that should be measured (Weingand et al. 1996). While these recommendations have not been formally incorporated into national or international guidelines at this stage, they are noted, as follows:

> For repeated-dose studies in rodent species, clinical pathology testing is necessary at study termination. Interim study testing may not be necessary in long-duration studies provided that it has been done in short-duration studies using dose levels not substantially lower than those used in the long-duration studies.
> For repeated-dose studies in non-rodent species, clinical pathology testing is recommended at study termination and at least once at an earlier interval. For studies of two to six weeks in duration in non-rodent species, testing is also recommended within 7 days of initiation of dosing, unless it compromises the health of the animals. If a study contains recovery groups, clinical pathology testing at study termination is recommended.
>
> The core haematology tests recommended are total leukocyte (white blood cell) count, absolute differential leukocyte count, erythrocyte (red blood cell) count, evaluation of red blood cell morphology, platelet (thrombocyte) count, haemoglobin concentration, haematocrit (or packed cell volume), mean corpuscular volume, mean corpuscular haemoglobin, and mean corpuscular haemoglobin concentration. In the absence of automated reticulocyte counting capabilities, blood smears from each animal should be prepared for reticulocyte counts. Bone marrow cytology slides should be prepared from each animal at termination. Prothrombin time and activated partial thromboplastin time (or appropriate alternatives) and platelet count are the minimum recommended laboratory tests of haemostasis. The core clinical chemistry tests recommended are glucose, urea nitrogen, creatinine, total protein, albumin, calculated globulin, calcium, sodium, potassium, total cholesterol and appropriate hepatocellular and hepatobiliary tests. For hepatocellular evaluation, measurement of a minimum of two scientifically appropriate blood tests is recommended, such as ALT, AST, SDH, GLGH or total bile acids.
>
> For hepatobiliary evaluation, measuring a minimum of two appropriate blood tests is recommended, such as AP, GGT, 5'-nucleotidase, total bilirubin or total bile acids.
>
> Urinalysis should be conducted at least once during a study. For routine urinalysis, an overnight collection (approximately 16 hours) is recommended. The core tests should include an assessment of urine appearance (colour and turbidity), volume, specific gravity or osmolality, pH, and either the quantitative or semi-quantitative determination of total protein and glucose.
>
> For carcinogenicity studies, only blood smears should be made from unscheduled sacrifices (decedents) and at study termination, to aid in the identification and differentiation of haematopoietic neoplasia.

A1.3.7 Absolute and relative organ weights

It is generally considered that histopathology is more sensitive for establishing the lowest dose producing an effect than are organ or body weight changes. Organ weights are usually reported as absolute organ weights and as relative organ weights (relative to body weight and/or brain weight). Relative organ weight comparisons are used since body weights are often affected by compound administration.

Experimentally controllable and uncontrollable factors, such as circadian rhythms, food intake, dehydration, nature of the diet, age of animals, organ workload, stress, and method of killing, have an influence on organ and body weights and the variability of such data. A review of this subject by Weil (1970) should be consulted. The most important influencing factor appears to be the method of killing and the timing of necropsy. The killing method used not only affects the appearance of the tissue, important in describing gross necropsy observations, but also, in conjunction with the timing of necropsies, may cause post-mortem shifts in organ weights (Boyd & Knight 1963).

A common problem in interpreting study findings is the misinterpretation of relative organ weight changes. For example, an increase in relative brain weight in a toxicity study in which the chemical causes a significant body weight loss or reduced body weight gain may not be indicative of a toxic effect on the brain as the brain would be spared under conditions leading to reduced body weight. Similarly, the relative weight of other organs may change because of changes in body weight rather than as a result of a specific compound effect. Tables of the relationship of relative organ weights to various levels of reduced bodyweights (produced by dietary restriction) for rats may be found in Sharer (1977). Changes in organ weight/body weight ratios as a physiological response of the organism to decreased nutrient intake and markedly reduced growth must be differentiated from organ weight changes resulting from primary toxic effects of the compound being tested.

The interpretation of organ weight changes must not be made solely on the determination of a statistically significant difference between the concurrent control value and a treatment group value. A proper evaluation will also include consideration of any correlation between organ weights (absolute and relative), histopathological and metabolic/pharmacodynamic data.

Since the evaluation of organ weight changes is such an integral part of toxicity assessment, the US Society of Toxicologic Pathology (STP) has made recommendations regarding the protocols for weighing organs and interpreting the findings in the light of the class of compound under test, the MoA for toxicity and the combined dataset for the study (Sellers et al. 2007).

A1.3.8 Post-mortem observation

Although much progress has been made in standardising nomenclature, to minimise any difficulties in this area, an experienced pathologist will describe each significant lesion type, at least once, in such detail that another competent pathologist can perceive a mental picture of the lesion and form a judgement as to its relevance to the histopathology induced by the chemical being tested.

To assist in the uniform description of pathologies, a series of articles on pathology nomenclature have been published, under the title *Standardized system of nomenclature and diagnostic criteria guides for toxicologic pathology* by the US Society of Toxicologic Pathologists (STP), in cooperation with the Armed Forces Institute of Pathology (AFIP) and the American Registry of Pathology (ARP).

Age-associated, especially geriatric, changes can have an extremely important effect on histopathology, as well as clinical chemistry, metabolic and pharmacokinetic parameters (Grice & Burek 1983; Mohr et al. 1992; 1994; 1996) and therefore, important overt, and frequently subtle, influences on observed physiological, pharmacological and toxic responses during the latter part of any long-term study. It is essential in all cases where spontaneous and/or age-associated lesions are present to differentiate between such lesions and treatment induced lesions. Grice and Burek (1983) and Benirschke et al. (1978) are very helpful in this respect, as is advice from a competent and experienced pathologist.

An overview of factors (physiological, environmental, etc.) that can complicate the interpretation of morphological findings in a toxicity study may be found in the *Handbook of toxicologic pathology* (Haschek et al. 2002).

A1.4 Analysing and evaluating study parameters in acute, developmental, reproductive and special toxicity studies

A1.4.1 Acute toxicity studies

Important endpoints in acute toxicity studies are clinical signs, gross necropsy signs and mortality incidence; each of these endpoints are discussed in detail in Sections A1.3.4, A1.3.8 and A1.3.3 respectively. Since the purpose of acute toxicity studies has moved away from establishing a strict, quantitative number for the median lethal dose to an estimate of the likely toxicity range, the emphasis is more on clinical signs and gross organ pathology than on mortality.

A1.4.2 Reproductive toxicity studies

Section A1.3 describes the major study parameters to be considered in repeat-dose toxicity studies; these endpoints may also apply to the sires and dams in developmental toxicity studies. However, the critical endpoints relate to potential toxic effects on reproductive parameters, including effects on mating behaviour (both sexes), on fertility (both sexes), the implantation of blastocysts, embryonic and foetal development and survival, parturition, lactation and postnatal survival and development. Thus, a plethora of reproductive parameters need to be assessed in one or more generations, depending on whether the study is a one-generation (OECD TG 415), two-generation (TG 416) or three-generation test. Important endpoints to assess within each generation include: time after pairing to mating; mating behaviour; percentage of females pregnant; number of pregnancies going to full term; litter size; number of live births; number of stillborns; pup viability and weight at parturition, and postnatal days 4, 7, 14 and 21 days of age the fertility index (percentage of matings resulting in pregnancy) gestation index (percentage of pregnancies resulting in live litters) viability index (percentage of pups that survive 4 or more days) and lactation index (percentage of pups alive at four days that survived to day 21 (i.e. weaning) gross necropsy and histopathology on some parents (sires and dams), with attention paid to the reproductive organs and gross necropsy on weanlings. Guidelines and procedures for assessing these endpoints are well documented (see e.g. Korach 1998).

A1.4.3 Developmental toxicity studies

The critical endpoints in developmental toxicity studies relate to developmental effects *in utero*, including death, malformations, functional deficits and developmental delays in foetuses.

Parameters assessed include: number of live litters; number of live foetuses/litter (total and by sex); sex ratio of foetuses; foetal body weights; litter weights; number and percentage of foetuses with malformations; number and percentage of litters with malformations; number and percentage of foetuses with variations; number and percentage of litters with variations; number and percentage of foetuses/litter with malformations; number and percentage of foetuses/litter with variations; and types of malformations and variations.

Guidelines and procedures for soft tissue and skeletal examination are well documented (e.g. Tyl & Marr 1997).

Additional reproductive parameters assessed include number of females pregnant; number of corpora lutea/dam; number of implants/dam; and number and percentage of pre-implantation loss/litter. While the dosing period in a standard teratology study commences after mating, conception and commencement of implantation, it is appropriate to check these parameters to see that the study has not been compromised by factors other than the compound under test. OECD TG 414 outlines the protocol for a standard developmental or 'teratology' study.

A1.4.4 Special studies

Different classes of chemicals may require special toxicology studies that are not part of the 'standard' package of studies. For example, it is common to test organophosphate (OP) pesticides for their ability to cause delayed neuropathy by conducting tests in hens (OECD TG 419), since this species is especially sensitive to inhibition of neuropathy target esterase (NTE) by OPs. Further *in vitro* and *in vivo* studies may be conducted to resolve possible mechanisms for toxic effects seen in the standard toxicology tests. Because of the wide range of types of studies that may be classified in this category, it is not possible to comment on the assessment of particular endpoints, but the evaluator should apply sound scientific judgement in reviewing these studies.

A1.4.5 Toxicokinetic and metabolism data

Toxicokinetic (absorption, distribution and elimination) and metabolic data on the handling of the substance in the test species can be very useful in evaluating and interpreting sub-chronic and chronic exposure study data, as discussed by Paynter (1984) and references cited therein.

References in this paper also discuss dose-dependent effects in the absorption process and in biotransformation interactions (Levy 1968), the potential difficulties presented by impurities, the overloading of detoxification mechanisms (Munro 1977) and other important experimental considerations (Dayton & Sanders 1983).

A number of toxicology textbooks include chapters on pharmacokinetics and toxicology assessment (e.g. Sharma & Coulombe 1996). The publication, *Science and judgement in risk assessment* (NRC 1994) has useful sections on the impact of pharmacokinetic information in risk assessment.

A1.5 Statistical tests

The objective of a toxicology study is to demonstrate responses of biological importance. Where statistical analyses are used in the judgement process, an awareness of the validity of the test and the degree of certainty (confidence) pertaining within the context of the study should be demonstrated.

There are limitations associated with using statistics in toxicology (Gad & Weil 1986):

- Statistics cannot make poor data better.
- Statistical significance may not imply biological significance.
- An effect that may have biological significance may not be statistically significant.
- The lack of statistical significance does not prove safety.

The importance and relevance of any effect observed in a study must be assessed within the limitations imposed by the study design and the species being studied.

If statistical tests have not been used, if inappropriate tests appear to have been used, or if tests not commonly employed have been used, then this should be noted and action taken such as data re-analysis.

A number of textbooks and papers on the application of statistics in experimental toxicology and the life sciences are available these include Dickens and Robinson (1996), Gad and Weil (1986), Gad and Weil (1989) and Lee (1993).

A1.6 General comments

If possible, compound-related changes in biochemical, haematological or urinalysis parameters should be linked with organ weight, gross pathology or histopathological changes.

The following points also should be noted in evaluating repeat-dose toxicity data.

Findings should be considered on the basis of both statistical significance and likely biological significance. The variability of biological data must be remembered in assessing a statistically significant result. Conversely, a finding that is not statistically significant may have biological significance when considered in the light of the likely toxicological or pharmacological action of the compound, or when combined with results from other studies. Thus, evaluators should note trends or transient changes in parameters if there is an indication that these may be related to dosing with the compound in some way. This information may be useful when comparing results across studies and in considering the overall significance or relevance of an observed effect (i.e. in one study an effect may be only a trend while in another study it may be very clearly treatment-related).

A difficult problem for evaluators is when studies have significant defects in design or reporting, such that they should be considered seriously flawed. The use of a flawed study reporting either a positive or negative outcome may provide a false sense of security, especially if there is pressure to consider it because no valid alternative study exists, and the study cannot be replaced with a valid or unflawed study. It is a matter of judgement as to how much weight should be accorded to a flawed study, and this in turn requires an element of experience in evaluating toxicity studies to make a determination of what type of flaws seriously compromise the study outcomes.

Data obtained from studies carried out many years ago should not simply be dismissed out of hand simply because they do not meet today's standards. Such studies may still provide some useful information if an experienced toxicologist is able to evaluate the study using a weight-of-evidence approach. This is a matter for scientific interpretation and judgement on a case-by-case basis.

Appendix 2: Extracts from a 2010 NHMRC discussion paper on health-based targets for microbial safety of drinking water supplies

The *Australian drinking water guidelines* (ADWG) provides the authoritative reference and benchmark on drinking water quality for Australian drinking water suppliers, health authorities and consumers. It is therefore essential that the scientific content of the guidelines is maintained and represents best practice. The ADWG was at the forefront of introducing a comprehensive risk-management-based approach to assure drinking water quality. This process, which started in the late 1990s, culminated in publication of the *Framework for management of drinking water quality* in the 2004 edition of the guidelines.

However, the ADWG has not matched international developments in setting health-based performance targets for achieving microbial safety. These targets are required to provide measurable benchmarks for effective risk management systems.

A2.1 Quantitative microbial risk assessment

Performance targets derived from quantitative definitions of microbial safety are a common theme in guidelines and standards developed by WHO, US EPA, Health Canada and New Zealand where they are used to identify appropriate barriers for ensuring safety of drinking water supplies. The *Australian guidelines for water recycling* (NRMMC, EPHC, NHMRC 2006) including the module on *Augmentation of drinking water supplies* (NRMMC, EPHC, NHMRC 2008) has adopted the WHO definition of microbial safety to underpin the setting of microbial targets for pathogenic bacteria, viruses and protozoa.

This paper canvasses the inclusion of a similar approach in the ADWG. Issues discussed are:

- a definition of microbial safety
- the principle of setting health-based performance targets for producing safe drinking water supplies
- mechanisms for setting performance targets for different types of source water
- mechanisms for setting performance targets for urban and non-urban supplies.

The benefit of such an approach is that it provides a process for identifying appropriate barriers for producing safe drinking water. The 2006 WHO GDWQ uses a combination of quantitative microbial risk assessment (QMRA) and the metric of disability-adjusted life years (DALYs) to define microbial safety.

QMRA is applied to determine the likelihood of infection and illness occurring from exposure to specific pathogens contained in water, while DALYs are used to convert the likelihood of illness into impacts or burdens of disease. The advantage of using DALYs is that the metric recognises that not all pathogens cause the same level or severity of disease. Some like *Cryptosporidium* cause mild diarrhoea in the general population while others such as *E. coli* 0157 and *Rotavirus* can result in death. In the Walkerton incident, seven people died from an outbreak involving *E. coli* 0157 and *Campylobacter* (Hrudey & Hrudey 2004). Severity of consequence is an important consideration in developing risk management systems and the greatest attention should be paid to hazards that can cause the largest harm.

DALYs represent the sum of time with an illness (i.e. loss of time in good health) and years lost through premature death. DALYs take account of each of the possible outcomes associated with a particular pathogen, for example:

- watery diarrhoea, bloody diarrhoea, haemolytic uraemic syndrome and death from infection with *E. coli* 0157
- gastroenteritis, reactive arthritis and Guillain-Barré syndrome from infection with *Campylobacter*.

Each of the outcomes is assigned a severity ranging from 0.02 to 0.12 for mild diarrhoea, and up to 1 for death.

The DALY per case is then determined by multiplying the severity by the persistence and percentage occurrence of each outcome. As shown in Box 5, the DALYs per case for illness caused by *Cryptosporidium* is 10 times less than illness caused by *Rotavirus*.

A2.2 Microbial safety and performance (treatment) targets

A key advantage of applying a quantitative metric for assigning public health impacts is that it can then be used to define safety. In the GDWQ, WHO defines microbial safety as water that does not give rise to more than 1 DALY per 1 million people per year (10^{-6} DALYs). This is equivalent to about one case of diarrhoea per 1,000 people per year. This value is considered to be the upper level of tolerable risk.

Having established the benchmark, the GDWQ shows how to use it to identify levels of treatment (performance targets) to produce safe drinking water from raw water. Setting performance targets is a four step process:

1. Determining or estimating concentrations of pathogens in the source water
2. Use of a quantitative microbial risk assessment to determine the extent of infection and illness arising from these pathogens
3. Translating the frequency of illnesses to burdens of disease and DALYs
4. Calculating required reductions in pathogen concentrations to comply with the reference level of 10^{-6} DALYs per person per year.

It is not practical, nor is there sufficient data, to develop health-based targets for all microbial pathogens. The recommended approach is to use reference organisms representing the major groups of pathogens (i.e. bacteria, viruses and protozoa). The selection criteria for reference pathogens include:

BOX 5: Disability-adjusted life years

The basic principle of the DALY is to weight each health impact in terms of its severity within the range of 0 for good health to 1 for death. Severities for outcomes of microbial infection include:

- 0.02–0.12 for mild diarrhoea
- 0.21 for reactive arthritis
- 0.23 for severe diarrhoea
- 1 for death.

The severity is then multiplied by duration of the effect and the relative frequency of occurrence in those who become ill. In the case of death, duration is regarded as the years lost in relation to normal life expectancy.

Hence, DALYs = YLL (years of life lost) + YLD (years lived with a disability/illness).

In this context, 'disability' refers to conditions that detract from good health. In the context of water-related guidelines, it generally relates to illness but in other areas it can also relate to physical or mental impairment.

Using an Australian example, infection with *Rotavirus* causes:

- mild diarrhoea (severity rating of 0.1) lasting three days in 97.5 per cent of cases
- severe diarrhoea (severity rating of 0.23) lasting seven days in 2.5 per cent of cases
- rare deaths of very young children in 0.015 per cent of cases (a death at under 1 year of age means a loss of up to 80 years of life).

$$\begin{aligned}\text{The DALY per case then} &= (0.1 \times 3/365 \times 0.975) + \\ &\quad (0.23 \times 7/365 \times 0.025) + \\ &\quad (1 \times 80 \times 0.00015) \\ &= 0.0008 + 0.0001 + 0.012 \\ &= 0.013\end{aligned}$$

Infection with *Cryptosporidium* can cause watery diarrhoea (severity weighting of 0.067) lasting for seven days with extremely rare deaths in 0.0001 per cent of cases. This equates to a DALY per case of 0.0015.

Campylobacter can cause diarrhoea of varying severity, Guillain-Barré syndrome of varying severity, reactive arthritis and occasional deaths. The calculated DALY per case is 0.0046.

Based on DALYs per case the impacts of the three pathogens is *Rotavirus* > *Campylobacter* > *Cryptosporidium*

Adapted from: NRMMC, EPHC, AHMC 2006. DALYs per case based on Havelaar and Melse 2003 and WSAA 2004.

- waterborne transmission established as a route of infection
- sufficient data available to enable a quantitative risk assessment to be performed including data on occurrence in source waters, dose responses in humans and disease burden
- a relatively high occurrence in source waters
- a relatively high survival in the environment
- a relatively high infectivity and severe disease burden
- a relatively low sensitivity to removal or inactivation by treatment processes.

There are a range of micro-organisms that could be selected as reference pathogens including *Campylobacter*, *E. coli* 0157, *Salmonella*, *Shigella*, *Rotavirus*, *Norovirus*, *Giardia* and *Cryptosporidium*. Selecting reference pathogens should take account of local conditions, including the prevalence of waterborne transmission and source water characteristics, and may vary between different countries and regions. In the GDWQ, *Campylobacter*, *Rotavirus* and *Cryptosporidium* are used as reference pathogens to illustrate the method for calculating health-based targets (WHO 2006a).

A2.3 Calculating health-based targets

Calculating health-based targets for source water containing 1 organism per litre of *Cryptosporidium*, *Campylobacter* or *Rotavirus* per litre is illustrated in Table A1. The calculations are based on the consumption of 2 litres of water per person per day. The probability of infection from single organisms is based on QMRA using published data largely derived from human dosing experiments (Haas et al.1999; Messner et al. 2001). DALYs per case were calculated as shown in Box 5. The required reductions range from 4.3 log for *Campylobacter* to 5.5 log for *Rotavirus*.

Table A1: Calculation of pathogen reductions for a source water containing 1 organism per litre of *Cryptosporidium*, Rotaviruses and *Campylobacter*[1]

Step		Equation	*Cryptosporidium*	*Campylobacter*	Rotavirus
a	Probability of infection per organism[2]		5.9×10^{-2}	1.9×10^{-2}	5.9×10^{-1}
b	Probability of infection per year (730 litres = 730 organisms)	***a* × 730**	43.1	13.9	431
c	Percentage of infection leading to illness		70%	30%t	88%
d	Probability of illness per year	***b* × *c***	30.1	4.16	379
e	DALYs per case		1.5×10^{-3}	4.6×10^{-3}	1.3×10^{-2}
f	Percentage of population susceptible to illness		100%	100%	6%
g	DALYs per person per year	***d* × *e* × *f***	4.52×10^{-2}	1.9×10^{-2}	2.96×10^{-1}
	Required reduction to achieve 10^{-6} DALY per year	***(g – 10⁻⁶) + g100***	99.998% (4.7 log)	99.995% (4.3 log)	99.9997% (5.5 log)

1 The information in this table is taken from WHO (2006a) as adapted by the *Australian guidelines for recycled water* (NRMMC, EPHC, AHMC 2006).

2 Low-dose formulae used as described in Appendix 2 of the *Australian guidelines for recycled water* (NRMMC, EPHC, AHMC 2006) as doses in safe drinking water below 10^{-5}.

Working backwards, based on consumption of 730 litres of water per year, 10^{-6} DALYs is equivalent to 2.2×10^{-5} *Cryptosporidium*, 5.2×10^{-5} *Campylobacter* and 3.4×10^{-6} *Rotavirus*, *Cryptosporidium* per litre of water. These concentrations reinforce the impracticality of direct pathogen monitoring of drinking water.

Depending on the source water quality, the log reductions required to produce safe water can then be calculated from the formula:

Log reduction = log (organisms in raw water per L ÷ 2.2×10^{-5} *Cryptosporidium*, or 5.2×10^{-5} *Campylobacter* or 3.4×10^{-6} *Rotavirus*)

The GDWQ indicates that ideally the concentration of pathogens in source water should be determined on a system-specific basis and then used to determine the log reductions that will produce safe drinking water. However, the GDWQ recognises that this is not always possible and suggests that default values could be used as an alternative. A limited range of defaults derived from international data are included in the guidelines.

A2.4 Include a definition of microbial safety

The two alternatives are:

- the definition that underpins the US EPA drinking water standards (1 infection per 10,000 people per year) or
- DALYs as applied by the WHO (10^{-6} DALYs per person per year).

The two approaches share a number of common features and, importantly, both can be used to identify quantifiable and measurable performance targets for producing safe drinking water. Typically these targets have been developed for reference pathogens such as *Cryptosporidium* or groups of pathogens such as protozoa, viruses and bacteria.

The difference between the two metrics is that application of DALYs includes an additional step. After calculating the likelihood of infection, the additional step is to consider the likelihood and severity of resulting disease. In this way, DALYs recognise that not all pathogens cause the same level or severity of disease. Infections can be asymptomatic and where symptoms occur they can range from mild diarrhoea to reactive arthritis, haemolytic uraemic syndrome and occasionally death.

A review of acceptable risk published by the American Academy of Microbiology (2006) indicated that while a goal of less than one infection per 10,000 people per year is a useful benchmark it has limitations and a more descriptive endpoint, taking into account measurements of injury, may be more advantageous. The DALY metric fulfils this requirement.

Application of DALYs requires data on disease outcomes. As discussed in Section 3.1, WHO has published a scoring system for severity of outcomes and estimates of DALYs per case for *Cryptosporidium*, *Campylobacter*, *E. coli* 0157 and *Rotavirus* (Havelaar & Melse 2003). These estimates are based on international data from developed countries. In the AGWR, an adjustment was included for diarrhoea caused by *Rotavirus* in Australia. Further refinements based on Australian data would increase the accuracy of applying DALYs to define safety of Australian drinking water supplies. A research proposal to provide this information has been developed.

The ADWG has been at the forefront of international guidelines and standards in adopting a risk management framework for assuring drinking water safety. However, the ADWG does not provide a quantitative definition for microbial safety. It also does not provide benchmarks to measure the success of risk management. In this important aspect, the ADWG lags behind comparable guidance provided by the US EPA (2006a), Health Canada (2004a and b), New Zealand (2008), WHO (2006a) and the *Australian guidelines for water recycling* (NRMMC, EPHC, AHMC 2006).

In the absence of microbial targets, many water suppliers have adopted parts of the US EPA regulations to provide performance benchmarks for assuring the safety of drinking water.

The 'do nothing' option means that the ADWG does not meet basic requirements of water suppliers or regulators. Reliance on the production of water containing 0 *E. coli* per 100 mL from unprotected water sources is not sustainable and is difficult to defend given the well-established limitations of this indicator in reflecting the presence or absence of enteric viruses and protozoa.

Internationally two alternatives have been established for defining microbial safety. The US EPA has performance targets set in drinking water standards that are based on an upper limit of one infection or illness per 10,000 people per year (US EPA 1989; 2002; 2006). WHO has adopted the metric of DALYs and defined tolerable risk as being less than 1×10^{-6} DALYs per person per year. As noted by the American Academy of Microbiology (2006), while benchmarks based on infection or illness rates are useful, a more descriptive endpoint, taking into account impacts, may be more advantageous.

The *Australian guidelines for water recycling* (NRMMC, EPHC, AHMC 2006), including the module on *Augmentation of drinking water supplies* (NRMMC, EPHC, NHMRC 2008), use DALYs and the WHO definition of tolerable risk as a basis for setting performance targets for achieving microbial safety.

Adopting a quantitative definition of microbial safety and establishing health-based performance targets has practical implications, including the need to categorise source water quality. The most direct method is to undertake system-specific monitoring based on source water characteristics. However, this is not always practically achievable. Alternative methods could be based on source water characteristics, sanitary surveys and monitoring for indicator bacteria.

The benefit of incorporating a definition of microbial safety and health-based targets is that they provide a mechanism for identifying appropriate barriers (catchment protection and treatment) to ensure safety of drinking water supplies.

Abbreviations

ACCS	Advisory Committee on Chemicals Scheduling
AChE	acetylcholinesterase
ACHHRA	Australian Centre for Human Health Risk Assessment
ACS	American Chemical Society
ACTRA	Australasian College of Toxicology and Risk Assessment
ADI	acceptable daily intake
ADS	Australian Dietary Survey
ADWG	Australian drinking water guidelines
AEF	Australian Exposure Factor
AF	attribution factors
AFIP	Armed Forces Institute of Pathology
AGWR	Australian guidelines for water recycling
AHMAC	Australian Health Ministers' Advisory Council
AHPC	Australian Health Protection Committee
AhR	aryl hydrocarbon receptor
AIHC	American Industrial Health Council
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
AUC	area under the curve
ANZECC	Australian and New Zealand Environment and Conservation Council
AP	alkaline phosphatase
APVMA	Australian Pesticides and Veterinary Medicines Authority
ARfD	acute reference dose
ARMCANZ	Agriculture and Resource Management Council of Australia and New Zealand
ARP	American Registry of Pathology
AT	averaging time
ATSDR	Agency for Toxic Substances and Disease Registry
BaP	benzo(a)pyrene
BMD	benchmark dose
BMDL	benchmark dose, lower confidence limit
BMR	benchmark risk
CHIP	Chemicals (Hazard Information and Packaging for Supply) regulation (UK)
CICAD	concise international chemical assessment documents
CIIT	Chemical Industry Institute of Toxicology
CLEA	contaminated land exposure assessment
CNS	central nervous system
COAG	Council of Australian Governments
COC	Committee on Carcinogenicity (UK)
COM	Committee on Mutagenicity (UK)
CoNTC	concentration of no toxicological concern
COPC	chemicals of potential concern
COSHH	control of substances hazardous to health
COT	Committee on Toxicity (UK)
CPK	creatine phosphokinase
CRC-CARE	Co-operative Research Centre for Contamination Assessment and Remediation of the Environment
CSAF	chemical-specific adjustment factors
CSF	cancer slope factor
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CSM	conceptual site model
CUF	combined uncertainty factor
CYP	cytochrome P450
DALY	disability-adjusted life years
DETR	Department of the Environment, Transport and the Regions
DIR	dealings involving intentional release
DSEWPC	Department of Sustainability, Environment, Water, Population and Communities
EASE	estimation and assessment of substance exposure
EC	exposure concentration
ECETOC	European Centre for Ecotoxicology and Toxicology of Chemicals
ECHA	European Chemicals Agency
ECRP	Existing Chemicals Review Program
ECVAM	European Centre for the Validation of Alternative Methods
EDC	endocrine disrupting chemicals
EHC	environmental health criteria
EHPC	Environmental Protection and Heritage Council
EHRA	environmental health risk assessment
EPA	Environment Protection Authority
EPHC	Environment Protection and Heritage Council
EPHSC	Environment Protection and Heritage Standing Committee
EUP	end-use products
FDA	Food and Drug Administration
FQPA	Food Quality Protection Act
FSANZ	Food Standards Australia New Zealand
GAF	gastrointestinal absorption factor
GD	guideline dose
GDWQ	guidelines for drinking water quality
GHS	Globally Harmonised System for classification and labelling of chemicals
GIS	geographic information systems
GLC	ground-level concentration
GLDH	glutamate dehydrogenase
GMO	genetically modified organism
GOF	goodness of fit
GV	guideline value

HEC	human-equivalent concentration
HED	human-equivalent dose
HESI	Health and Environmental Sciences Institute
HI	hazard index
HIA	health impact assessment
HIL	health investigation level
HPV	high production volume
HQ	hazard quotient
HRA	health risk assessment
HSE	health and safety executive
HSL	health screening level
IARC	International Agency for Research on Cancer
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ICNA	Industrial Chemicals Notification and Assessment (Act)
IEH	Institute for Environmental Health
IEUBK	Integrated Exposure Uptake Biokinetic Model for Lead in Children
IFCS	Intergovernmental Forum on Chemical Safety
IGAE	Inter Governmental Agreement on the Environment
IGHRC	Interdepartmental Group on Health Risks from Chemicals
IGR	ingestion rate
ILCR	increased lifetime cancer risk
ILO	International Labour Organization
ILSI	International Life Sciences Institute
IOMC	Inter-Organization Programme for the Sound Management of Chemicals
IPCS	International Programme on Chemical Safety
IRIS	Integrated Risk Information System
ITER	International Toxicity Estimates for Risk
JECFA	Joint FAO/WHO Expert Committee on Food Additives
JMPR	Joint Meeting on Pesticide Residues
LD	Longfellow DG
LOAEL	lowest observed adverse effect level
LOD	limit of determination
LOR	level of reporting
MBS	market basket survey
MF	modifying factor
ML	maximum levels
MoA	mode of action
MoE	margin of exposure
MRA	microbiological risk assessment
MRL	maximum residue level
MTD	maximum tolerated dose
NAS	National Academy of Science

NATO	North Atlantic Treaty Organisation
NCEA	National Center for Environmental Assessment
NChEM	National Framework for Chemicals Environmental Management
NEPC	National Environment Protection Council
NEPM	National Environmental Protection Measure
NGO	non-governmental organisations
NHMRC	National Health and Medical Research Council
NICEATM	NTP Interagency Center for the Evaluation of Alternative Toxicological Methods
NICNAS	National Industrial Chemicals Notification and Assessment Scheme
NIEHS	National Institute of Environmental Health Sciences
NIST	National Institute of Standards & Technology
NOAEL	no observed adverse effect level
NOEL	no observed effect level
NOHSC	National Occupational Health and Safety Commission
NRA	National Registration Authority for Agricultural and Veterinary Chemicals
NRC	National Research Council
NRMMC	Natural Resource Management Ministerial Council
NTE	neuropathy target esterase
NTMDI	national theoretical maximum daily intakes
NTP	National Toxicology Program
OCS	Office of Chemical Safety
OECD	Organisation for Economic Cooperation and Development
OGTR	Office of the Gene Technology Regulator
OHS	occupational health and safety
OSHA	Occupational Safety and Health Administration
OTC	over the counter
PAD	population-adjusted dose
PAH	polycyclic aromatic hydrocarbon
PBPK	physiologically based pharmacokinetic models
PCB	polychlorinated biphenyls
PCO	pest control operator
PEC	priority existing chemical
PEL	permissible exposure limits
PHED	Pesticide Handlers Exposure Database
POD	point of departure
POEM	Predictive Operator Exposure Model
POP	Persistent Organic Pollutant
PP	precautionary principle
PPAR	peroxisome proliferator-activated receptor
PQRA	preliminary quantitative risk assessment
PRA	probabilistic risk assessment
PRG	preliminary remediation goals

PTWI	provisional tolerable weekly intake
QA	quality assurance
QC	quality control
QRA	quantitative risk assessment
QSAR	quantitative structure–activity relationship
QMRA	quantitative microbial risk assessment
RAF	risk analysis framework
RAGS	risk assessment guidance for superfund
RARMP	risk assessment and risk management plan
REACH	registration, evaluation, and authorisation of chemicals
RfC	reference concentration
RfD	reference dose
RI	risk index
RiskIE	Risk Information Exchange
RIVM	National Institute for Public Health and the Environment (Netherlands)
RPF	relative potency factor
RSC	relative source contribution
SA	surface area
SARF	social amplification of risk framework
SF	safety factor
SGV	soil guidance values
SIDS	Standard Information Data Sets
SOT	Society of Toxicology
SRA	Society for Risk Analysis
SRM	standard reference materials
SSRA	site-specific risk assessment
STEL	short-term exposure limits
STP	Society of Toxicologic Pathology
SUSMP	Standard for the Uniform Scheduling of Medicines and Poisons
SWA	Safe Work Australia
TCDD	2,3,7,8-Tetrachlorodibenzo-p-dioxin
TDI	tolerable daily intake
TEF	toxicity equivalence factor
TEQ	toxicity equivalent
TES	Toxicology Evaluation Section
TGAC	technical grade active constituent
THM	trihalomethane
TI	tolerable intake
TLV	threshold limit values
TMI	tolerable monthly intake
TPH	total petroleum hydrocarbons
TOXNET	toxicology databases maintained by the US National Library of Medicine
TRV	toxicity reference value
TTC	threshold of toxicological concern
TWI	tolerable weekly intake
UF	uncertainty factor
UNCED	UN Conference on Environment and Development
UNEP	United Nations Environment Programme
URF	unit risk factor
VOC	volatile organic compounds
WHO	World Health Organization
WoE	weight of evidence
XRF	X-ray fluorescence
YLD	years lived with a disability
YLL	years of life lost

Glossary

Note: the terminology in this glossary has been largely based on that used in the original 2002 edition of the enHealth guidance on EHRA, plus some terms copied from the glossary of the 2010 revision of the contaminated sites NEPM, and the 2006 NHMRC guidance on ambient air quality standard-setting.

Absorbed dose	The amount of chemical that, after contact with the exchange boundary (skin, lungs, gut), actually penetrates the exchange boundary and enters the circulatory system. The amount may be the same or less than the applied dose. (See Table 6 for other types of doses used in health risk assessment.)
Accuracy	The degree to which a measurement represents the true value of the variable that is being measured (NHMRC 2000); or the degree of agreement between the average predictions of a model or the average of measurements and the true value of the quantity being predicted or measured (WHO 2003).
Acceptable daily intake (ADI)	The daily intake of a chemical that, during a lifetime, appears to be without appreciable risk on the basis of all the facts known at the time. It is expressed in milligrams per kilogram of body weight per day (mg/kg/day). For this purpose, 'without appreciable risk' is taken to mean that adverse effects will not result even after a lifetime of exposure. Furthermore, for a pesticide residue, the ADI is intended to give a guide to the maximum amount that can be taken daily in the food without appreciable risk to the consumer. Accordingly, the figure is derived as far as possible from feeding studies in animals. (See also 'tolerable daily intake' and 'reference dose'.)
Acceptable risk	This is a risk management term. The acceptability of risk depends on scientific data, social, economic and political factors, and the perceived benefits arising from exposure to an agent. (See also 'target risk'.)
Acute exposure	A contact between an agent and a target occurring over a short time, generally less than 14 days, with a single or repeated dose. (Other terms, such as 'short-term exposure' and 'single-dose' are also used.)
Adduct	A chemical moiety that is covalently bound to a large molecule such as DNA or protein.
Adverse effect	The change in the morphology, physiology, growth, development, reproduction or life span of an organism, system population or sub-population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress or an increase in susceptibility to other influences. Some adaptive changes are not generally considered to be adverse (e.g. some changes in enzyme levels).
Agent	A chemical, physical or biological substance or factor (including social factor) being assessed in the context of an environmental health risk assessment.
Aggregate/cumulative risk	Terminology derived from US legislation. The term 'aggregate risk' in this context, implies consideration of all sources of exposure to determine a total (or aggregated) exposure estimate. The term 'cumulative' risk implies that the risk associated with substances sharing a common mode of action or toxicity outcome, are aggregated across the exposure estimates for all such substances.
Air pollution	The presence of contaminants (air pollutants) in high enough concentrations in the air that could interfere with human health or welfare, or produce other harmful environmental effects.
Ambient air	An unconfined portion of the atmosphere; also open air or surrounding air .
Applied dose	Amount of an agent presented to an absorption barrier and available for absorption. The amount may be the same or more than the absorbed dose. (See Table 6 for other types of doses used in health risk assessment.)

Background level (or concentration)	The amount (or concentration) of agent in a medium (e.g. water or soil) that is not attributed to the sources under investigation in an exposure assessment. Background levels can be naturally occurring or the result of human activities.
Benchmark dose (BMD)	The dose associated with a given incidence (the benchmark risk, e.g. 1 per cent, 5 per cent or 10 per cent incidence) of effect, based on the best-fitting dose–response curve.
Benchmark risk (BMR)	A predetermined incidence of adverse response that determines the benchmark dose.
Bias	A process resulting in a tendency to produce results that differ in a systematic value from the true values. Also known as systematic error.
Bioaccessibility	The fraction of a contaminant in an exposure medium that is soluble in the relevant physiological milieu (usually the gastrointestinal tract) and available for absorption. Generically, it is the ability for a chemical to come into contact with the absorbing surfaces in an organism. It is related to solubility and dissolution, since absorption usually can only occur from a liquid or gaseous phase and not from a solid phase.
Bioavailability	A generic term defined as the fraction of a contaminant that is absorbed into the body following dermal contact, ingestion or inhalation. It is expressed as the ratio (or percentage) of the absorbed dose (systemic dose) to the administered dose. (See Table 6 for other measures of bioavailability.)
Biological monitoring	Measurement of a contaminant or metabolite in body tissue, fluid, blood or expired air.
Biomarker	Any measurement reflecting an interaction between a biological system and an environmental agent that may be chemical, physical or biological (WHO 1993). Often used to describe measurements used in biological monitoring.
Cancer or carcinogenesis	A disease of heritable, somatic mutations affecting cell growth and differentiation. That is, genetic alterations incurred in the first damaged cells are acquired in subsequent cells after cell division within the same individual. It encompasses the origin, causation and development of tumours and applies to all forms of tumours (e.g. benign and malignant).
Cancer slope factor (CSF)	The plausible upper-bound estimate of the probability of a response per unit of intake of an agent over a lifetime.
Carcinogen	Chemical, biological or physical cancer-causing agent. A distinction may be made based on the presumed mode of action (MoA) – see 'genotoxicity' and 'non-genotoxic carcinogen'.
Carcinogenicity	A property of an agent that enables it to produce tumours, whether benign or malignant.
Causality	The relating of causes to the effects they produce. Most of epidemiology concerns causality, and several types of causes can be distinguished. However, epidemiological evidence by itself is insufficient to establish causality, although it can provide powerful circumstantial evidence.
Chronic exposure	A contact between an agent and a target occurring over a continuous or repeated basis for a duration of three months or greater. (See also 'sub-chronic exposure' and 'lifetime'.)
Chronic toxicity	An adverse effect that is generally induced by prolonged exposure to a chemical. It may also include an ability to produce an adverse effect that persists over a long period of time, whether or not it occurs immediately upon exposure to a chemical or is delayed.
Chemical of potential concern (COPC)	An agent that is potentially associated with the site or exposure medium under consideration and whose data is of sufficient quality to be judged as potentially causing an adverse health effect.

Cluster	A greater than expected number of cases that occur within a group of people in a geographic area over a period of time (Queensland Health 2009).
Cluster assessment	A scientific process to determine if there is an increased number of cases of a specific disease or condition and to determine if there is a biologically plausible causal agent(s) for the disease (Queensland Health 2009).
Community	Those individuals or groups residing in a locality where a site assessment is to be conducted and who may be affected by the assessment and/or possible site contamination physically (e.g. through risks to health or the environment, loss of amenity) or non-physically such as via concern about possible contamination). The term 'wider community' may be applied to individuals and/or groups not necessarily residing in the locality of the site assessment who may have an interest in the assessment (NEPC 2010).
Conceptual site model	A description of a site including the environmental setting, geological, hydrogeological and soil characteristics, together with nature and distribution of contaminants. Potentially exposed populations and exposure pathways are identified. Presentation is usually graphical or tabular with accompanying explanatory text.
Confidence	Weight assigned by the evaluator to the quality of the information available (high, medium or low confidence) to indicate that a chemical possesses certain toxicological properties.
Confidence limit	A range of values determined by the degree of presumed random variability in a set of data, within which the value of a parameter (e.g. the mean) lies with a specified level of confidence or probability (e.g. 95 per cent). The upper and lower confidence limits refer to the values at opposite ends of the specified range.
Confounding factor	A factor that distorts the apparent effect or magnitude of the effect of a study factor or risk. Confounding factors must be controlled for in order to obtain an undistorted estimate of the effect under study.
Conservatism	A cautious or conservative approach to evaluating and managing the uncertainties inherent in a risk assessment, which reduces the probability of harm occurring.
Contaminant	Any chemical existing in the environment above background levels and representing, or potentially representing, an adverse health or environmental risk (may be synonymous with a pollutant).
Contamination	The condition of land, water or food where any chemical substance or waste has been added or detected at above background level and represents, or potentially represents, an adverse health or environmental impact (NEPC 2010).
Critical effect	The adverse effect judged to be the most important for setting an acceptable human intake or exposure. It is usually the most sensitive adverse effect, that is, that with the lowest effect level, or sometimes a more severe effect, not necessarily having the lowest effect level.
Data quality objectives (DQOs)	The establishment of the amount, nature and quality of data required to complete a specific risk assessment.
Default value	A pragmatic, fixed or standard value used in the absence of relevant data.
Deterministic	A deterministic approach uses single values or point estimates as input values in an exposure or risk estimation model. These are intended to be 'best estimates' of the value of the input variables. (see also 'probabilistic').

Disability-adjusted life years (DALYs)	For a given health condition, the sum of the years of life lost due to premature mortality in the population and the years lost due to disability for incident cases. It is a term used more commonly in quantitative microbiological risk assessment (QMRA) rather than in HRA for chemicals.
Dose	A stated quantity or concentration of a substance to which an organism, system, population or sub-population is exposed over a continuous or intermittent duration of exposure. It is generally the total amount of a chemical administered, but there may be other expressions relating to the amounts actually absorbed or taken up (see Table 6 for other types of doses used in health risk assessment). Dose is most commonly expressed as the amount of test substance per unit weight of test species (e.g. mg/kg body weight).
Dosage	A general term comprising the dose, its frequency and the duration of dosing. Dosage is properly applied to any rate or ratio involving a dose. Dosages often involve the dimension of time (e.g. mg/kg/day), but the meaning is not restricted to this relationship.
Dose–response	Relationship between the amount of chemical administered to, taken up by, or absorbed by an organism, system or (sub)population and the change developed in that organism, system or (sub) population in reaction to the agent. It is the correlative association existing between the dose administered and the response (effect) or spectrum of responses that is obtained. The concept expressed by this term is indispensable to the identification, evaluation and interpretation of most pharmacological and toxicological responses to chemicals. The basic assumptions that underlie and support the concept are: (a) the observed response is a function of the concentration at a site; (b) the concentration at a site is a function of the dose; and (c) response and dose are causally related (Eaton & Klaassen 1996). The existence of a dose–response relationship for a particular biological or toxicological response (effect) provides a defensible conclusion that the response is a result of exposure to a known substance.
Dose–response curve	Graphical representation of a dose–response relationship that is essential to any quantitative estimation of risk for a given exposure.
Endpoint	An observable or measurable biological event used as an indicator of the effect of a chemical on a biological system (cell, organism, organ etc.). It may also be expressed as a 'toxicological endpoint'.
Environmental health	Those aspects of human health determined by physical, chemical, biological and social factors in the environment. Environmental health practice covers the assessment, correction, control and prevention of environmental factors that can adversely affect health, as well as the enhancement of those aspects of the environment that can improve human health.
Environmental monitoring	The monitoring of the concentration of substances in the physical environment of air, water, soil and food.
Epidemiology	The study of the distribution and determinants of health-related states or events in specified populations, and the application of the study to the control of health problems.
Expert	An expert has: (1) training and experience in the subject area resulting in superior knowledge in the field; (2) access to relevant information; (3) an ability to process and effectively use the information; and (4) is recognised by his or her peers or those conducting the study as qualified to provide judgements about assumptions, models and model parameters at the level of detail required (NCRP 1996).
Expert/professional judgement	Opinion of an authoritative person on a particular subject.

Exposure	Concentration or amount of a particular chemical that reaches a target organism, system, population or sub-population in a specific frequency for a defined duration. Exposure is usually quantified as the concentration of the agent in the medium integrated over the time duration of contact.
Exposure assessment	The estimation (qualitative or quantitative) of the magnitude, frequency, duration, route and extent (e.g. number of organisms) of exposure to one or more contaminated media for the general population, for different sub-groups of the population, or for individuals.
Exposure concentration	The exposure mass divided by the contact volume or the exposure mass divided by the mass of contact volume, depending on the medium.
Exposure duration	The length of time over which continuous or intermittent contacts occur between a chemical and the exposed population.
Exposure event	The occurrence of continuous contact between chemical and exposed population.
Exposure frequency	The number of exposure events within an exposure duration.
Exposure route or pathway	The way a chemical enters an organism after contact (e.g. by ingestion, inhalation or dermal absorption). The pathway usually describes the course a chemical or physical agent takes from a source to an exposed organism. An exposure pathway describes a unique mechanism by which an individual or population is exposed to chemicals or physical agents at or originating from a site. Each exposure pathway includes a source or release from a source, an exposure point and an exposure route. If the exposure point differs from the source, a transport/exposure medium (e.g. air) or media (in cases of inter-media transfer) is also indicated.
Exposure scenario	A set of conditions or assumptions about sources, exposure pathways, concentration of contaminants involved and exposed population (e.g. numbers, characteristics, habits) used in the evaluation and quantification of exposure(s) in a given situation. The exposure scenario may be expressed in terms of a model, that is, a conceptual or mathematical representation of the exposure process.
Exposed population	The people who may be exposed to the contaminant. Synonymous with 'receptors'.
Extrapolation	For dose–response curves, an estimate of the response at a point outside the range of the experimental data most commonly extrapolated to low dose. Also refers to the estimation of a response in different species or by different routes than that used in the experimental study of interest.
Factor	A single factor or product of several single factors used to derive an acceptable intake. These factors account for adequacy of the study, inter-species extrapolation, inter-individual variability in humans, adequacy of the overall database, nature and extent of toxicity, public health regulatory concern and scientific uncertainty. The terms safety factor (SF), uncertainty factor (UF) and modifying factor (MF) are examples of the terminology used in different jurisdictions to imply essentially the same process.
False negative	A result that is erroneously negative leading to a determination that the factor under study is not present. In statistical inference, this is a Type 2 error.
False positive	A result that is erroneously positive leading to a determination that the factor under study is present when it is not. In statistical inference, this is a Type 1 error.
Genotoxicity	A broad term describing the ability to produce damage to the genetic material (DNA) of cells or organisms.
Genotoxic chemical	A chemical for which there is adequate evidence of the potential to interact with, and/or modify the function of genetic material.

Genotoxic carcinogen	A chemical for which there is adequate evidence that the ability to induce tumours is via a mechanism involving direct damage to DNA.
Guideline values (GVs)	Values such as concentrations in air or water that are derived after appropriate allocation of tolerable intake (TI) among the possible different media of exposure. Combined exposure from all media at the guidance values over a lifetime would be expected to be without appreciable health risk. The aim of a guidance value is to provide quantitative information from risk assessment for risk managers to enable them to make decisions concerning the protection of human health. (WHO 1994a, p. 16).
Hazard	Inherent property of a contaminant or situation having the potential to cause adverse effects when a population may be exposed to that contaminant. It is also described as the disposition of a thing, a condition or a situation to produce an adverse health or environmental effect; or an event, sequence of events or combination of circumstances that could potentially have adverse consequences (adapted from ACDP 1996). Note the definition of risk to distinguish hazard from risk.
Hazard identification	The identification of the type and nature of adverse effects that a contaminant has an inherent capacity to cause harm to an exposed population.
Hazard indices/index (HI)	The sum(s) of at least two hazard quotients.
Hazard quotient (HQ)	The ratio of the mean daily intake to the reference dose or tolerable daily intake for threshold exposure.
Health	Health is a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity (WHO 1946).
Health investigation levels (HILs)	Screening criteria based on health risk, presented in Schedule B(7) of the contaminated sites NEPM. May also be called health screening levels (HSLs) to emphasise the fact that they represent an outcome of a Tier 1-type screening level risk assessment, and may require a more refined Tier 2–3 level process to better define the risk.
Health risk assessment (HRA)	The process of estimating the potential impact of a chemical, biological, physical or social agent on a specified human population system under a specific set of conditions and for a certain time frame. May also be described as a process intended to calculate or estimate the risk to a given target organism, system or (sub)population, including the identification of attendant uncertainties following exposure to a particular contaminant, taking into account the inherent characteristics of the agent of concern as well as the characteristics of the specific target system.
Heuristics	A psychological term used to describe the process whereby people frame their perceptions of risk, based on 'rules of thumb' and other emotional (affective) factors by which we make judgements about everyday occurrences.
Hormesis	Demonstrated beneficial effects of an agent at low (but not homeopathic) doses but with toxicity occurring at higher doses. Also used to describe 'hockey-stick' or other J-shaped non-monotonic dose–response relationships where biological effects may appear to become greater as the dose decreases.
Immunotoxicity	The ability to produce an adverse effect on the functioning of organs and cells involved in immune competence (IEH 1999b).
In vitro/in silico	Describes tests undertaken in test tubes, culture dishes or other systems where a non-living organism is exposed to a test agent. *In silico* techniques refer to modern genomic methodologies where genes or DNA arrays on microchips are the responsive agents.

Integrated Risk Information System (IRIS)	The database of the US EPA that provides the agency's adopted hazard and dose–response assessment for chemical and radiological agents. Used as guidance and to provide consistency in the agency's regulatory decisions designed to reduce risk related to environmental exposures.
LD_{50}	The quantity of a chemical compound that, when applied directly to test organisms via inhalation, oral or dermal exposure, is estimated to be fatal to 50 per cent of those organisms under the stated conditions of the test.
Lowest observed effect level (LOEL)	The lowest concentration or amount of a substance found by experiment or observation that causes alterations of morphology, functional capacity, growth, development or life span of target organisms. WHO (1990) defines it as the lowest dose of a substance that causes changes distinguishable from those observed in normal (control) animals.
Lowest observed adverse effect level (LOAEL)	The lowest concentration or amount of a substance found by experiment or observation that causes adverse alterations of morphology, functional capacity, growth, development or life span of target organisms.
Level of detection (LOD)	The minimum concentration or mass of analyte that can be detected at a known confidence level.
Level (limit) of reporting (LOR)	The value calculated from the instrumentation detection limits and with appropriate scale-up factors applied. The scale-up factors are affected by the procedures, methods and the size of the sample.
Lifestyle factors	Behaviours or habits that are a matter of individual choice and that may impinge in the outcomes of a risk assessment. Examples include smoking, poor diet and alcohol intake.
Lifetime	A figure used in exposure assessment and risk characterisation representing the average life span of an organism. Seventy years has been conventionally used for humans, but newer demographic data suggests that human life spans are expanding.
Metabolite	A substance that is the product of biochemical alteration of the parent compound in an organism.
Mode of action (MoA)	A description of observable key events or processes from interaction of an agent with a cell or tissue through operational and anatomical changes to the disease state (EPA 2005).
Model	A mathematical representation of a biological system intended to mimic the behaviour of the real system, allowing description about empirical data and predictions about untested states of the system.
Mutagenicity	The ability to produce a permanent, heritable change in the amount or structure of genetic material of cells or organisms (IEH 1999b) (see also 'genotoxicity').
National Environment Protection Measure (NEPM)	National guidance on assessment and management of environmental pollution, established under the National Environment Protection Act. NEPMs are broad framework-setting statutory instruments defined in the NEPC Act. They outline agreed national objectives for protecting or managing particular aspects of the environment. Establishment, maintenance and review of NEPMs is the responsibility of the Environment Protection and Heritage Council (EPHC), which incorporates the National Environment Protection Council (NEPC), a statutory body under the NEPC Acts of the Commonwealth, states and the territories. The EPHC addresses broad national policy issues relating to environmental protection, particularly in regard to air, water and waste matters.
Neurotoxicity	The ability to produce an adverse effect in the central or peripheral nervous system (IEH 1999b).

No observed adverse effect level (NOAEL) (may also be cited as 'no observable adverse effect level')	The highest concentration or amount of a substance, found by experiment or observation, that causes no observable alterations of morphology, functional capacity, growth, development or life span of target organisms. The NOAEL is the next dose below the LOAEL in the series of doses tested in a study, where no toxic (i.e. adverse) effects are observed. It may also be worded in more detail thus: The NOAEL is defined as the highest exposure at which there is no statistically or biologically significant increase in the frequency of an adverse effect when compared with a control group (National Academy of Sciences, National Research Council 1994). The definition of NOEL is equivalent, but with the removal of the term, 'adverse'. Often, the difficult issue in the use of the terms NOEL or NOAEL is in deciding whether a compound-related effect noted in a particular study is necessarily an 'adverse' effect. Alterations of morphology, functional capacity, growth, development or life span of the target organism may be detected, which are judged not to be adverse.
No observed effect level (NOEL)	The 'highest dose of a substance administered to a group of experimental animals at which there is an absence of observable effects on morphology, functional capacity, growth, development or life span that are observed or measured at higher dose levels used in the study. Thus, dosing animals at the NOEL should not produce any biologically significant differences between the group of chemically exposed animals and an unexposed control group of animals maintained under identical conditions. The NOEL is expressed in milligrams of chemical per kilogram of body weight per day (mg/kg bw/day) or, in a feeding study, in ppm in food (converted to mg/kg bw of compound intake by measured or estimated food intake over the period of the study). The NOEL has been simply defined as the highest dose of a substance that causes no changes distinguishable from those observed in normal (control) animals (WHO 1990).
Non-genotoxic carcinogen	An agent that induces tumours via a mechanism that does not involve direct damage to genetic material (DNA); sometimes referred to as 'epigenetic'.
Physiologically based pharmacokinetic (PBPK) model	Modelling the dose or degree of exposure to a chemical at a target tissue, cell or receptor by integration of pharmacokinetic data with anatomical, physiological and biochemical data (IEH 1999b).
Particulate matter (PM_{10}, $PM_{2.5}$)	The fraction of particles passing an inlet with a 50 per cent cut-off efficiency at an aerodynamic diameter of 10 µm (PM_{10}) or 2.5 µm ($PM_{2.5}$). May also be referred to as ultrafine particulate matter.
Pica	A behaviour exhibited occasionally by young children characterised by the deliberate ingestion of non-nutritive substances, such as soil.
Point of departure (POD)	A point on a dose–response curve that is defined by the available data and close to the range of observed data points, from which extrapolation techniques (e.g. linearised extrapolation and/or application of safety/uncertainty factors) are used to estimate a toxicity reference value.
Probabilistic	A probabilistic approach uses frequency distributions of parameters from which input data is randomly selected for repeated calculations to generate a frequency distribution of the output (exposure or risk). (See also 'deterministic')
Provisional tolerable weekly intake (PTWI)	The tolerable intake of a chemical expressed as a weekly amount. The term was established by WHO (1972) for several heavy metals which 'are able to accumulate within the body at a rate and to an extent determined by the level of intake and by the chemical form of the heavy metal present in food' (WHO 1989).
Public health	The science and art of preventing disease, prolonging life and promoting health through the organised efforts of society.

REACH program	The Registration, Evaluation, Authorisation and Restriction of Chemical substances program (REACH), established in 2006 as a new European Community program for regulating chemicals and their safe use.
Read across	An extrapolation technique that may be applied when information on the toxicological properties of a substance is missing or incomplete. It relies on extrapolating from the toxicological profile of a known, and related, substance to the substance under consideration.
Reproductive toxicity	The ability to produce an adverse effect on any aspect of reproductive capacity, function or outcome. It includes effects on the embryo, foetus, neonate and prepubertal organism and on adult reproductive and neuroendocrine systems (IEH 1999b).
Reference dose (RfD)	An estimate (with uncertainty factors spanning perhaps an order of magnitude) of the daily exposure (mg/kg/day) to the general human population (including sensitive sub-groups) that is likely to be without an appreciable risk of deleterious effects during a lifetime of exposure. It is derived from the NOAEL or the LOAEL by application of uncertainty factors that reflect various types of data used to estimate RfD and an additional modifying factor, which is based on professional judgement of the entire database of the chemical (IRIS 1996). The RfD is equivalent in meaning to tolerable daily intake (TDI) and acceptable daily intake (ADI). Usually doses less than the RfD are not likely to be associated with adverse health risks, and are therefore less likely to be of regulatory concern. As the frequency and/or magnitude of the exposures exceeding the RfD increase, the probability of adverse effects in a human population increases. However, all doses below the RfD are not assumed to be 'acceptable' (or risk-free) and nor are all doses that exceed the RfD necessarily 'unacceptable' (i.e. likely to result in adverse effects) (US EPA). The term acute reference dose (ARfD) is used to designate a level of exposure (using the same types of uncertainty and other qualifiers) that is likely to be without an appreciable risk or deleterious effect after a single dose or short period of exposure.
Risk	The probability that, in a certain time frame, an adverse outcome will occur in a person, group of people, plants, animals and/or the ecology of a specified area that is exposed to a particular dose or concentration of a hazardous agent, that is, it depends on both the intrinsic toxicity of the agent and the level of exposure. Risk differs from hazard primarily because risk considers probability.
Risk characterisation	The qualitative and, wherever possible, quantitative determination, including attendant uncertainties, of the probability of occurrence of known and potential adverse effects of an agent in a given organism, system or (sub)population under defined exposure conditions.
Risk communication	An interactive two-way process involving the exchange among individuals, groups and institutions of information and expert opinion about the nature, severity and acceptability of risks and the decisions taken to combat them. It usually involves an interactive exchange of information about health and environmental risks among risk assessors, managers, news media, interested groups and the general public (see also 'stakeholders').
Risk management	The process of evaluating alternative actions, selecting options and implementing them in response to risk assessments. The decision making will incorporate scientific, technological, social, economic and political information. The process requires value judgements (e.g. on tolerability and reasonableness of costs).
Safety	Practical certainty that adverse effects will not result from exposure to an agent under defined circumstances. It is the reciprocal of risk. Safety does not demand zero risk and would be a meaningless term if it did.
Safety factor (SF)	See 'factor'. Composite (reductive) factor by which an observed or estimated no observed adverse effect level (NOAEL) is divided to arrive at a criterion or standard that is considered safe or without appreciable risk.

Sensitive groups	Refers to populations with both susceptibility and vulnerability factors (see 'susceptibility' and 'vulnerability').
Sensitivity analysis	The process of changing one variable while leaving the others constant and determining the effect on the output. The procedure commonly involves fixing each uncertain quantity, one at a time, at its credible lower bound and then its upper bound (holding all other at their medians), and then computing the outcomes for each combination of values (USEPA 1992). It can be used to test the effects of both uncertainty and variability in input values.
Skin irritancy	A local inflammatory reaction affecting the skin.
Stakeholder	One who has an interest in a project or who may be affected by it.
Stochastic	A random probabilistic phenomenon.
Structure–activity relationship (SAR)	The relationship between the biological activity of a chemical or series of chemicals and their molecular structure. The relationships can be described qualitatively and quantitatively.
Sub-chronic exposure	A contact between an agent and a target of intermediate duration between acute and chronic. Different bodies vary on their definitions of the duration of 'sub-chronic' exposure, since it varies with species. US EPA uses up to 10 per cent of an organism's lifetime; however, between three and six months is often used when discussing sub-chronic exposure to people (see also 'chronic exposure').
Susceptibility	Refers to intrinsic biological factors that can increase the health risk of an individual at a given exposure level; examples of susceptibility factors include: genetic factors, late-age and early-life, and prior or existing disease.
Target risk	The risk level assessed by extrapolation of a dose–response relationship, suggesting an exposure level where the risk could be considered to be 'negligible, tolerable or acceptable' to the risk manager. (See also 'acceptable risk').
Teratogenicity	The ability to produce a structural malformation or defect in an embryo or foetus (IEH 1999b).
Threshold	The lowest dose or exposure level that will produce a toxic effect, and below which no toxicity is observed (IEH 1999b). A non-threshold dose–response relationship implies that the response incidence is only zero at zero exposure, and that a finite level of risk may be determined (using extrapolation methodology) at any exposure level above zero. Linear extrapolation typically refers to extrapolation to the zero exposure or zero effect origin of a dose–response curve.
Tolerable intake (TI)	An estimate of the intake of a substance that over a lifetime is without appreciable health risk (WHO 1994a). Examples are the ADI, TDI and reference dose.
Tolerable daily intake (TDI)	An estimate of the daily intake of a substance that can occur over a lifetime without appreciable health risk. It may have different units depending on the route of administration (WHO 1994a). The term 'acceptable daily intake' is used for chemicals such as pesticides (herbicides, insecticides and antifungals) that are deliberately used on food crops or food-producing animals and for which some level of residues may be expected to occur in food. The term 'tolerable daily intake' is used when the chemical is a potential food or environmental contaminant. While exposure should not occur, a TDI is an established health limit below which lifetime exposure should not have any adverse health effects. (See also 'acceptable daily intake' and 'reference dose'.)
Tolerable weekly (monthly) intake (TWI/TMI)	The tolerable intake (TI) expressed as a weekly or monthly amount.
Toxicity	Inherent property of a chemical to cause an adverse biological effect.

Toxicity equivalence (TEQ)	A method of expressing the combined (assumed additive) toxicity of a group of like chemicals that share a common mode of action. The TEQ is based on summing exposure estimates for individual components of a mixture multiplied by an estimate of their toxic potency (toxicity equivalence factor – TEF) relative to a reference substance. An alternative US terminology for the TEF is relative potency factor (RPF).
Toxicity reference value (TRV)	Measures of tolerable intake or acceptable risk, such as reference doses and cancer slope factors.
Tumour	A mass of abnormal, disorganised cells arising from pre-existing tissue that is characterised by excessive and uncoordinated cell proliferation or growth and by abnormal differentiation (specialisation). There are two types of tumours: benign and malignant. Benign tumours morphologically resemble their tissue of origin, grow slowly (may also stop growing) and form encapsulated masses; they do not infiltrate other tissues, they do not metastasise and are rarely fatal unless they cause physical disruption of a critical body function (e.g. a brain tumour). Malignant tumours (also called carcinomas) resemble their parent tissue less closely and are composed of increasingly abnormal cells genetically, morphologically and functionally. Most grow rapidly, spread progressively through adjacent tissues and metastasise to distant tissues.
Tumour initiation	The first step in carcinogenesis whereby a small number of cells (or one cell) are irreversibly changed due to genetic damage.
Tumour progression	The stage in carcinogenesis when tumours acquire the features of malignant growth.
Tumour promotion	The process by which initiated cells undergo clonal expansion (reproduction of a genetically damaged cell) to form overt tumours.
Uncertainty	Lack or incompleteness of information or knowledge about toxicological profile of a substance or the correct value to be input in to a risk assessment, such as a specific exposure measure or estimate.
Uncertainty factor	See 'factor': A numerical factor applied to the no observed adverse effect level (NOAEL) to derive an exposure level considered to be without appreciable risk to health (the NOAEL is divided by the uncertainty factor). The magnitude of the uncertainty factor depends on the nature of the toxicity observed, the quality of the toxicological data available, and whether the effects were observed in humans or animals (IEH 1999b).
Unit risk factor (URF)	An expression of the incremental risk associated with increase in exposure by a single unit of exposure measure. It may also be expressed as the plausible upper-bound estimate of the probability of a response from a chemical over a lifetime. It is derived from the slope of the linearised dose–response relationship and usually expressed in units of concentration for a specified medium (e.g. incremental risk per $\mu g/m^3$ in air).
Variability	True differences in attributes or values due to diversity or heterogeneity. This may include measurable factors that differ (e.g. height is variable across populations). The major types of variability are temporal, spatial and inter-individual. They may be discrete (e.g. albinism) or continuous (e.g. body weight). It may be readily identifiable (e.g. presence of albinism) or difficult to identify (e.g. ability to detoxify a particular chemical metabolite).
Vulnerability	Refers to human populations at higher risk due to environmental factors; examples of vulnerability factors include poverty, malnutrition, poor sanitation, climate change and stress associated with mental health diseases.
Weight of evidence (WoE)	Considerations in assessing the interpretation of published information about toxicity, quality of testing methods, size and power of study design, consistency of results across studies, and biological plausibility of exposure–response relationships and statistical associations.

References

Abelson PR (1994). Risk assessments of low-level exposures (editorial). *Science*, 265:1507.

Adami H-O, Berry CL, Breckenridge CB, Smith LL, Swenberg JA, Trichopoulos D, Weiss NS and Pastoor TP (2011). Toxicology and epidemiology: Improving the science with a framework for combining toxicological and epidemiological evidence to establish causal inference. *Toxicol Sci* 122:223–234.

Advisory Committee on Dangerous Pathogens (ACDP) (1996). *Microbiological risk assessment: An interim report*. London: HMSO.

Agency for Toxic Substances and Disease Registry (ATSDR) (1992) Agency for Toxic Substances and Disease Registry. Public Health Assessment Guidance Manual Division of Health Studies, ATSDR. Washington. Available at <http://www.atsdr.cdc.gov/HAC/phamanual/toc.html>.

Agency for Toxic Substances and Disease Registry (ATSDR) (1996). Guidance for ATSDR Health Studies. Division of Health Studies. Washington, DC: ATSDR. <http://www.atsdr.cdc.gov/HS/gd1.html>.

Agency for Toxic Substances and Disease Registry (ATSDR) (2004). *Guidance manual for the assessment of joint toxic action of chemical mixtures*. US Department of Health and Human Services Public Health Service, Division of Toxicology.

Agency for Toxic Substances and Disease Registry (ATSDR) (2006) *Evaluating vapor intrusion pathways at hazardous waste sites*. <http://www.atsdr.cdc.gov/document/evaluating_vapor_intrusion.pdf>.

Ahlers J, Stock F and Werschkun B (2008). Integrated testing and intelligent assessment: New challenges under REACH. *Envir Sci Pollut Res* 15:565–572.

AIHC Task Force (1994). *Exposure factors sourcebook*. Environmental Health Risk Assessment Sub-committee, Washington, DC: American Industrial Health Council.

Aitio A, Jarvisalo J, Riihimaki V and Hernberg S (1988). *Biologic monitoring*. In C. Zenz (Ed.), *Occupational medicine: Principles and practical applications*. Chicago: Year Book Medical Publishers.

Albertini R, Bird, M, Doerrer, N, Needham L, Robinson S, Sheldon L and Zenick H (2006). The use of biomonitoring data in exposure and human health risk assessments. *Environ Health Perspect* 114:1755–1762.

Alsop WR, Ruffle B, Frohmberg EJ and Anderson PD (1993). Development of risk-based clean-up standards using Monte Carlo analysis. In *Developing clean-up standards for contaminated soil, sediment, and groundwater: How clean is clean?* Washington, DC: ATSDR/US EPA.

Altenburger R, Walter H and Grote M (2004) What contributes to the combined effects of a complex mixture? *Environ Sci Technol* 38:6353–6362.

American Academy of Microbiology (2006). *Clean water: What is acceptable microbial risk?* Washington, DC: American Academy of Microbiology.

American Academy of Pediatrics (1999). Testimony from the Committee on Genetics. Presentation to the Public Advisory Committee of the Food and Drug Administration. Dermatologic and Ophthalmic Drugs Advisory Committee. Presented 4 September 1997 by James W. Hanson, MD. <http://www.aap.org/visit/thal01.htm>.

American Conference of Governmental Industrial Hygienists (ACGIH) (1990–1991), *Threshold limit values for chemical substances and physical agents and biological exposure indices*. Cincinnati: American Conference of Governmental Industrial Hygienists.

American Industrial Health Council (AIHC) (1989). *Presentation of risk assessments of carcinogens: Report of the Ad Hoc Study Group on Risk Assessment*. Washington, DC: American Industrial Health Council.

Ames BN (1987). Six common errors relating to environmental pollution. *Regulat Toxicol Pharmacol* 7:379–383.

Ames BN, Durston WE, Yamasaki E and Lee FD (1973). Carcinogens are mutagens: A simple test system combining liver homogenates for activation and bacteria for detection. *Proceedings of the National Academy of Sciences of the USA* 70:2281–2285.

Andersen ME and Dennison JE (2004). Mechanistic approaches for mixture risk assessments: Present capabilities with simple mixtures and future directions. *Environ Toxicol Pharmacol* 16:1–11.

Andersen ME and Krewski D (2009). Toxicity testing in the 21st century: Bringing the vision to life. *Toxicol Sci* 107:324–330.

Angerer J, Bird M, Burke TA, Doerrer NG, Needham L, Robinson SH, Sheldon L and Zenick H (2006). Strategic biomonitoring initiatives: Moving the science forward. *Toxicol Sci* 93:3–10.

Appel MJ, Bouman HGM, Pieters MN and W. Slob W (2001). Evaluation of the applicability of the benchmark approach to existing toxicological data framework: Chemical compounds in the working place. Bilthoven: National Institute of Public Health and the Environment. RIVM report 601930 001, TNO report V99.1098:1–103.

Arnold SF and Price PS (2007). Modeling mixtures resulting from concurrent exposures to multiple sources. *Toxicol Appl Pharmacol* 223:121–124.

Australian Institute of Health and Welfare (AIHW) (2000). *The seventh biennial health report of the Australian Institute of Health and Welfare*. Canberra: AIHW.

Axelrod D, Burns K, Davis, D, and von Larabeke N (2004). Hormesis: An inappropriate extrapolation from the specific to the universal. *Int J Occupat Environ Health* 10:335–339.

Bader, EL, Hrudey,SE and Froese, KL (2004). Urinary excretion half life of trichloroacetic acid as a biomarker of exposure to chlorinated drinking water disinfection by-products. *Occup Environ Med* 61: 715–716.

Baker D and Nieuwenhuijsen MJ (2008). *Environmental epidemiology: Study methods and application*. New York: Oxford University Press.

Barlow SM (2005) Threshold of toxicological concern: A tool for assessing substances of unknown toxicity present at low levels in the diet. Belgium: ILSI Europe Concise Monograph Series.

Barlow SM, Kozianowski G, Wurtzen G and Schlatter J (2001). Threshold of toxicological concern for chemical substances present in the diet. *Food Chem Toxicol* 39:893–905.

Barnett V and Lewis T (1994). *Outliers in statistical data*. 3rd edn. New York: John Wiley & Sons.

Bayne K (1996). Revised guide for the care and use of laboratory animals. *Physiologist* 39:199, 208–211.

Bellar TA, Lichtenberg JJ and Kroner RC (1974). The occurrence of organohalides in chlorinated drinking waters. *Journal of the American Water Works Association* 66:703–706.

Belpomme D, Irigaray P, Hardell L, Clapp R, Montagnier L, Epstein S and Sasco AJ (2007). The multitude and diversity of environmental carcinogens. *Environ Res* 105:414–429.

Benfenati E, Benigni R, DeMarini DM, Helma C, Kirkland D, Martin TM, Mazzatorta P, Ouedraogo-Arras G, Richard AM, Schilter B, Schoonen WGEJ, Snyder RD and Yang C (2009). Predictive models for carcinogenicity and mutagenicity: Frameworks, state-of-the-art, and perspectives. *J Environ Sci Hlth Part C* 27:57–90.

Benirschke K, Garner FM. and Jones TC (Eds) (1978). *Pathology of laboratory animals*. Vols I and II. New York: Springer Verlag.

Benner TC (2004). Brief survey of EPA standard-setting and health assessment. *Environ Sci Technol*. 34:3457–3464.

Berg BN and Simms HS (1960). Nutrition and longevity in the rat: II. Longevity and onset of disease with different levels of food intake *J Nutr* 71:225–263.

Biello D (2006). Mixing it up: Harmless levels of chemicals prove toxic together. *Scientific American* 294:18.

Birnbaum LS and Stokes WS (2010). Safety testing: Moving towards alternative methods. *Environ Health Perspect* 118:A12–13.

Birnbaum LS, Harris MW, Stocking LM, Clark AM and Morrissey RE (1989). Retinoic acid and 2,3,7,8-tetrachlorobenzo-p-dioxin selectively enhances teratogenesis in c57bl/6n mice. *Toxicol Appl Pharmacol* 98:487–500.

Biosecurity Australia (2006). *Final import risk analysis report for apples from New Zealand. Part B*. Canberra: Biosecurity Australia.

Bjarnason (2004). *Toxicology of chemical mixtures: A review of mixtures assessment for the Canadian Defence R & D as part of a NATO project*. NATO HFM-057/RTG-009.

Blackburn K, Stickney JA, Carlson-Lynch HL, McGinnis PM, Chappell L and Felter SP (2005). Application of the threshold of toxicological concern approach to ingredients in personal and household care products. *Regulat Tox Pharmacol* 43:249–259.

Blando JD and Cohn P (2004). Exposure and health risk from swimming in outdoor pools contaminated by trichloroethylene. *Human and Ecological Risk Assessment*. 10:717–731.

Boekelheide K (2007). Toxicological highlights: Mixed messages. *Toxicol Sci* 99:1–2.

Boffetta P, McLaughlin JK, La Vecchia C, Tarone RE, Lipworth L and Blot WJ (2008). False-positive results in cancer epidemiology: A plea for epistemological modesty. *J Nat Cancer Inst* 100:988–995.

Bogen KT, Cullen AC, Frey HC and Price PS (2009). Probabilistic exposure analysis for chemical risk characterization. *Tox Sci* 109:4–17

Bonita R, Beaglehole R and Kjellstrom T (2007). *Basic epidemiology*. 2nd edn. Geneva: World Health Organization.

Boobis AR (2007). The future of mixtures toxicology. *Toxicology* 240:142.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D and Farland W (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit Rev Toxicol* 36:781–92.

Borak J and Sirianni G (2005). Hormesis: Implications for cancer risk assessment. *Dose Response*, 3:443–451.

Borgert CJ, Quill TF, McCarty LS and Mason AM (2004). Can mode of action predict mixture toxicity for risk assessment? *Tox Appl Pharmacol* 201:85–96.

Box GEP (1979). Robustness in the strategy of scientific model building. In RL Launer and GN Wilkinson GN (Eds). *Robustness in statistics*. New York: Academic Press.

Boyd EM and Knight LM (1963). Postmortem shifts in the weight and water levels of body organs. *Toxicol Appl Pharmacol* 5:119–128.

Breslow NE and Day NE (1987). *Statistical methods in cancer research. Vol. II – The design and analysis of cohort studies*. International Agency for Research on Cancer, Scientific Publication 82, Lyon.

Bucher JR and Portier C (2004). Human carcinogenic risk evaluation. Part V: The National Toxicology Program vision for assessing the human carcinogenic hazard of chemicals. *Toxicol Sci* 82:363–366.

Bull RJ, Brown JM, Meierhenry EA, Jorgenson TA, Robinson M and Stober JA (1986). Enhancement of the hepatoxicity of chloroform in B6C3F1 mice by corn oil: Implications for chloroform carcinogenesis. *Environ Health Perspectives* 69:49–58.

Budzt-Jorgensen E, Keiding N and Grandjean P (2004). Effects of exposure imprecision on estimation of the benchmark dose. *Risk Analysis* 24:1689–1696.

Burmaster DE and Anderson PD (1994). Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. *Risk Analysis* 14:477–481.

Busby JS Alcock RE and Macgillivray BH (2009). Interrupting the social amplification of risk process: A case study in collective emissions reduction. *Environ Science & Policy*, 12:297–308.

Bush BM (1991). *Interpretation of laboratory results for small animal clinicians*. London: Blackwell Scientific Publications.

Butterworth BE (2006). A classification framework and practical guidance for establishing a mode of action for chemical carcinogens. *Regulat Toxicol Pharmacol* 45:119–143.

Butterworth BE, Aylwood LL and Hays SM (2007). A mechanism-based cancer risk assessment for 1,4-dichlorobenzene. *Regulat Toxicol Pharmacol* 49:138–148.

Calabrese EJ (1985). *Toxic susceptibility: Male/female differences*. New York: John Wiley & Sons.

Calabrese EJ (2004). Hormesis: Basic, generizable, central to toxicology and a method to improve the risk-assessment process. *Int J Occup Environ Health* 10:466–467.

Calabrese EJ (2005). Toxicological awakenings: The rebirth of hormesis as a central pillar of toxicology. *Toxicol Appl Pharmacol* 204:1–8.

Calabrese EJ (2008). Hormesis: Why it is important to toxicology and toxicologists. *Environ Toxicol Chem* 27:1451–1474.

Calabrese EJ and Baldwin LA (1994). Improved method of selection of the NOAEL. *Regulat Toxicol Pharmacol* 19:48–50.

Calabrese EJ and Gilbert (1993). Lack of total independence of uncertainty factors (UFs): Implication for size of the total uncertainty factor. *Regulat Toxicol Pharmacol* 17:44–51.

Calabrese EJ and Stanek EJ (1995). Resolving intertracer inconsistencies in soil ingestion estimation. *Environ Health Perspect*.103:454–457.

Calabrese, EJ and Stanek EJ (1991). A guide to interpreting soil ingestion studies: II. Qualitative and quantitative evidence of soil ingestion. *Regulat Toxicol and Pharmacol* 13:278–292.

Calamari D (1999). Assessment of environmental distribution: The use of fugacity models for this purpose. In UNEP/IPCS, *Chemical risk assessment*. Training Module No. 3. WHO/PCS/99.2. Geneva: WHO.

Calamari D (Ed.) (1993). *Chemical exposure prediction*. Chelsea: Lewis Publishers.

Calder MC (2008). *Contemporary risk assessment in safeguarding children*. Lyme Regis, UK: Russell House Publishing.

Callahan MA and Sexton K (2007). If 'cumulative risk assessment' is the answer, what is the question? *Environ Health Perspect* 115:799–806.

Canadian Council of Ministers of the Environment (CCME) (2010). Canadian soil quality guidelines: Carcinogenic and other polycyclic aromatic hydrocarbons (environmental and human health effects). Scientific Criteria Document (revised) PN 1445.

Carlson L and Gelber K (2001). *Ideas for community consultation: A discussion on principles and procedures for making consultation work*. NSW Department for Urban Affairs and Planning.

Casterline JL Jr and Williams CH (1969). Effect of pesticide administration upon esterase activities in serum and tissue of rats fed variable casein diets. *Toxicol Appl Pharmacol* 14:266–275.

Cedergreen N, Christensen AM, Kamper A, Kudsk P, Mathiassen SK and Streibig JC and Sorensen H (2008). A review of independent action compared to concentration addition as reference models for mixtures of compounds with different molecular target sites. *Env Toxicol Chem* 27:1621–1632.

Centre for Research into the Built Environment (CRBE) (1995). *Guideline values for polynuclear aromatic hydrocarbon contamination in soils expressed as benzo(a)pyrene toxic equivalents*. Draft prepared for the Department of the Environment, EPG 1/6/18. Centre for Research into the Built Environment, Nottingham Trent University.

Chen J, Chen YJ, Teuschler L, Rice G, Kamernik K, Protzel A, and Kodel RL (2003). Cumulative risk assessment for quantitative response data. *Environmetrics* 14:339–353.

Chess C and Hance, BJ (1994) *Communicating with the public: 10 questions environmental managers should ask*. Center for Environmental Communication, State University of New Jersey. Rutgers, New Jersey.

Cleveland WS (1994). The elements of graphing data. Summit, NJ: Hobart Press.

Code of Federal Regulations (CFR) (2001). *Threshold of regulation for substances used in food-contact articles*. 21 CFR 170.39, United States Government Printing Office.

Codex (1999). Principles and guidelines for the conduct of microbiological risk assessment, Cac/GI-30.

Codex Committee on Food Hygiene (CCFH) (2003). Proposed draft principles and guidelines for the conduct of microbiological risk management. Codex Alimentarius Commission. Joint FAO/WHO Food Standards Program. CCFH, 35th session, Orlando, 27 Jan – 1 Feb 2003.

Cohen SM, Klaunig J, Meek ME, Hill RN, Pastoor T, Lehman-McKeeman L, Bucher J, Longfellow DG, Seed J, Dellarco V, Fenner-Crisp P, and Patton D (2004). Evaluating the human relevance of chemically induced animal tumours. *Toxicol Sci* 78:181–186.

Conner MW and Newbern PM (1984). Drug-nutrient interactions and their implications for safety evaluations. *Fund Appl Toxicol* 4:5341–5356.

Connolly RB and Lutz WK (2004). Nonmonotonic dose–response relationships: Mechanistic basis, kinetic modelling and implications for risk assessment. *Toxicol Sci*. 77:151–157.

Co-operative Research Centre for Contamination Assessment and Remediation of the Environment Pty Ltd (CRC-CARE) (2009). *HSLs for petroleum hydrocarbons in soil and groundwater*. Technical development guideline. CRC-CARE, Technical Report series No. 10.

Covello VT and Merkhofer MW (1993). *Risk assessment methods: Approaches for assessing health and environmental risks*. New York: Plenham Press.

Cramer GM, Ford RA and Hall RL (1978). Estimation of toxic hazard: A decision tree approach. *Food Cosmet Toxicol* 16:255–276.

Creswell, JW (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks, CA: SAGE publications.

Creton S, Dewhurst IC, Earl LK, Gehen SC, Guest RL, Hotchkiss JA, Indans I, Woolhiser MR and Billington R (2010). Acute toxicity testing of chemicals: Opportunities to avoid redundant testing and use alternative approaches. *Crit Rev Toxicol* 40:50–83.

Crofton KM, Craft ES, Hedge JM, Gennings C, Simmons JE, Carchman RA, Carter WH Jr and DeVito MJ (2005). Thyroid-hormone-disrupting chemicals: Evidence for dose-dependent additivity or synergism. *Environ Health Perspect* 113:1549–1554.

Cromar N, Cameron S & Fallowfield H (2004) *Environmental health in Australia and New Zealand*. Oxford University Press.

Crome P (1999). The elderly. In Institute for Environment and Health (IEH). *Risk assessment strategies in relation to population sub groups*. Leicester: Institute for Environment and Health.

Crump K (1984). A new method for determining allowable daily intakes. *Fund Appl Toxicol* 4:854–891.

Crump K (1985). Mechanisms leading to dose–response models. In P Ricci (Ed.), *Principles of health risk assessment*. Englewood Cliffs, NJ: Prentice Hall.

Cullen AC and Frey HC (1999). *Probabilistic techniques in exposure assessment: A handbook for dealing with uncertainty and variability in models and inputs*. New York: Plenum Press.

Davis G, Patterson B and Trefry M. (2009b). *Biodegradation of petroleum hydrocarbon vapours*. Technical Report No. 12. Adelaide: CRC CARE.

Davis G, Trefry M and Patterson B (2009a). *Petroleum vapour model comparison*. Technical Report No. 9. Adelaide: CRC CARE.

Dayton PG and Sanders JE (1983) Dose-dependent pharmacokinetics: Emphasis on Phase I metabolism. *Drug Metab Rev* 14:347–405.

de Bruijn JHM, Jager DT, Kalf DF, Mensink BJWG, Montforts MHMM, Sijm DTHM, Smit CS, van Vlaardingen PLA, Verbruggen EMJ, van Wezel AP (2001). *Guidance on deriving environmental risk limits*. RIVM report 601501 012. National Institute of Public Health and the Environment (RIVM), Bilthoven.

De Rosa CT, El-Masri HA, Pohl H, Cibulas W and Mumtaz MM (2004). Implications of chemical mixtures in public health practice. *J Toxicol Environ Hlth, Part B*, 7:339–350.

De Rosa CT, Hicks HE, Ashizawa AE, Pohl HR and Mumtaz MM (2006). A regional approach to assess the impact of living in a chemical world. *Ann NY Acad Sci* 1076: 829–38.

Department of Health (DOH) (1991). *Guidelines for the evaluation of chemicals for carcinogenicity. Committee On Carcinogenicity of Chemicals in Food, Consumer Products and the Environment*. Department of Health, Report RHSS 42, London: HMSO.

Department of Health (DOH) (1998). *Communication about risks to public health: Pointers to good practice*. London: HMSO.

Department of the Environment, Transport and the Regions (DETR), the Environment Agency, Institute for Environment and Health (2000). *Guidelines for environmental risk assessment and management*. <http://www.defra.gov.uk/environment/quality/risk/eramguide>.

Derelanko MJ and Hollinger MA (Eds) (1995). *CRC handbook of toxicology*. Boca Raton, Florida: CRC Press.

Di Marco PN and Buckett KJ (1996). Derivation of health investigation levels for beryllium and beryllium compounds. In A Langley, B Markey and H Hill (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission, pp. 1–20.

Di Marco PN, Priestly BG and Buckett KJ (1998). Carcinogen risk assessment: Can we harmonise? *Toxicol Letters* 102–103:241–246.

Dickens A and Robinson J (1996). Statistical approaches. In GO Evans (Ed.), *Animal clinical chemistry: A primer for toxicologists*. London: Taylor and Francis, pp. 45–48.

Doerrer N (2007). Integration of human biomonitoring exposure data into risk assessment: HESI initiatives and perspectives. *Int J Hyg Environ Health* 210:247–251.

Dolan DG, Naumann BD, Sargent EV, Maier A and Doursen M (2005). Application of the threshold of toxicological concern concept to pharmaceutical manufacturing operations. *Regulat Toxicol Pharmacol* 43:1–9.

Doull J (1980). Factors influencing toxicology. In *Casarett and Doull's Toxicology*, 2nd edn, pp. 70–83, Macmillan Pub. Co., New York. (Note: This reference is to an old edition of this well-known textbook but the latest edition (at page 22) still refers to this particular chapter in the 2nd edition for detailed discussion on factors influencing animal responses to toxic substances).

Drew R & Frangos J (2007). The Concentration of No Toxicological Concern (CoNTC): A risk assessment screening tool for air toxics. *J Toxicol Env Hlth*. Part A 70:1584–1593.

Duncan JR, Prasse KW and Mahaffey EA (1994). *Veterinary laboratory medicine: Clinical pathology*. 3rd edn. Ames, Iowa: Iowa State Press.

Dybing E, Soderlund EJ (1999). Situations with enhanced chemical risks due to toxicokinetic and toxicodynamic factors. *Regul Toxicol Pharmacol*. 30:S27–30.

EEC (2003). EC (European Commission Health and Consumer Protection Directorate). General Directorate D – Food Safety: Production and Distribution Chain, D3 – Chemical and Physical Risks: Surveillance Food Contact Materials Practical Guide, updated 15 April 2003.

Elespuru RK, Agarwal R, Atrakchi AH, Bigger AH, Heflich RH, Jagannath DR, Levt DD, Moore MM, Ouyang Y, Robinson TW, Sotomayer RE, Cimino MC and Dearfield KL (2009). Current and future application of genetic toxicity assays: The role and value of in vitro mammalian assays. *Toxicol Sci* 109:172–179.

Elinder C-G (1985). Normal values for cadmium in human tissues, blood and urine in different countries. In L Friberg et al. (Eds), *Cadmium and health: A toxicological and epidemiological appraisal*, Exposure, dose and metabolism. Boca Raton, Florida: CRC Press Inc. Vol. 1. 81–102.

Elliott P, Cuzick, English D, Stern R (1992). *Geographical and environmental epidemiology: Methods for small-area studies*. Oxford: World Health Organization.

Ellison SLR, Gregory S and Hardcastle W (1998). Quantifying uncertainty in qualititative analysis. *The Analyst* 123:1155–1161.

enHealth (2001). *Health impact assessment guidelines*. enHealth, Canberra: Commonwealth of Australia.

enHealth (2003). *Guidelines for economic evaluation of environmental health planning and assessment*. Vol. 1 – The Guidelines Vol. 2 – Case studies. Canberra: Department of Health and Ageing and enHealth Council.

enHealth (2006) *Responding to environmental health incidents: Community involvement handbook*. Melbourne: NPHP.

Environment Australia (1998). National profile of chemicals management infrastructure in Australia. Environment Australia, available at <http://www.environment.gov.au/settlements/chemicals/publications/pubs/profile.pdf>.

Environment Protection Authority (EPA) NSW (1997). *Contaminated sites: Guidelines for consultants reporting on contaminated sites*. Sydney: NSW Environment Protection Authority.

Evans GO (1996). *Animal clinical chemistry: A primer for toxicologists*. London: Taylor and Francis.

Falk-Filipsson A, Hanberg A, Victorin K, Warholm M and Wallén M (2007). Assessment factors: Applications in health risk assessment of chemicals. *Environ Res* 104:108–127.

Fao V and Allesio L (1983). Biological monitoring. In L Parmeggiani (Ed.), *Encyclopaedia of occupational health and safety*. Geneva: International Labour Organization. pp. 271–274.

Faustman EM and Omenn GS (1996). Risk assessment. In Klaassen CD (Ed.), *Casarett and Doull's toxicology: The basic science of poisons*. 5th edn. New York: McGraw Hill, pp. 5–88.

Fawell J (2000). Risk assessment case study: Chloroform and related substances. *Food Chem Toxicol* 38:S91–S95.

Federal Focus (1996). Principles for evaluating epidemiologic data in regulatory risk assessment: Recommendations for implementing the 'London principles' and for risk assessment guidance. Developed by an Expert Panel at London Conference October, 1995. Washington: Federal Focus.

Ferguson CC (1994). *The Contaminated Land Exposure Assessment Model (CLEA): Technical basis and algorithms*. Draft report submitted to Department of the Environment. Nottingham: The Nottingham Trent University Centre for Research into the Built Environment.

Feron VJ, Cassee FR, Groten JP, van Vliet PW and van Zorge JA (2002). International issues on human health effects of exposure to chemical mixtures. *Environ Health Perspect* 110 (Suppl. 6):893–899.

Feron VJ, van Vliet W and Notten WRF (2004). Exposure to combinations of substances: A system for assessing health risks. *Environ Toxicol Pharmacol* 18:215–222.

Fiksel JR and Scow KM (1983). Human exposure and health risks assessments using outputs of environmental fate models. In RL Swann and Eschenroedera (Eds), *Fate of chemicals in the environment compartmental and multi-media models*. ACS Symposium series 225. Washington: American Chemical Society.

Finley B et al. (1994). Recommended distributions for exposure factors frequently used in health risk assessment. *Risk Analysis* 14:533–53.

Finucane, ML (2004). *The psychology of risk judgments and decisions*. In N. Cromar, S. Cameron, and H. Fallowfield (Eds), *Environmental health in Australia and New Zealand* (pp. 142–155). New York: Oxford University Press.

Fitzgerald DJ (1993). Carcinogenic soil contaminants: An Australian approach. In AJ Langley, M van Alphen (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Fitzgerald J (1998). The benchmark dose approach and health-based investigations level for polycyclic aromatic hydrocarbons (PAHs). In A Langley, P Imray, W Lock and H Hill (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Fjeld RA, Eisenberg NA and Compton KL (2007). *Quantitative environmental risk assessment for human health*. Hoboken, NJ: John Wiley & Sons.

Foran JA and The International Life Sciences Institute (ILSI) (1997) Risk Sciences Working Group on Dose Selection: Principles for the selection of doses in chronic rodent bioassays. *Env Hlth Persp* 105:18–20.

Foth H and Hayes AW (2008). Concept of REACH and impact on evaluation of chemicals. *Human Exptl Toxicol* 27:5–21.

Fowle JR (1989). Summary and perspectives: Panel discussion on toxicological exposure assessment: State of the art. *J Am Coll Toxicol* 8:865–870.

Fowle JR and Sexton K (1992). EPA priorities for biological research in environmental health. *Environ Health Perspect*. 98:235–241.

Fowles JR, Alexeeff GV and Dodge D (1999). The use of benchmark dose methodology with acute inhalation lethality data. *Regul Toxicol Pharmacol* 29:262–278.

Frawley JP (1967). Scientific evidence and common sense as a basis for food-packaging regulations. *Food Cosmet Toxicol* 5:293–308.

Friberg LT (1985). The rationale of biological monitoring of chemicals: With special reference to metals. *American Industrial Hygiene Association Journal* 46:633–641.

Froese, KL, Sinclair, M and Hrudey, SE (2002). Trichloroacetic acid as a biomarker of exposure to disinfection by-products in drinking water: A human exposure trial in Adelaide, Australia. *Environ Health Perspect* 110: 679–687.

Fryer ME, Collins CD, Colvile RN, Ferrier H and Nieuwenhuijsen MJ (2004) Evaluation of currently used exposure models to define a human exposure model for use in chemical risk assessment in the UK *IGHRC Report 28*.

Fryer M, Collins CD, Ferrier H, Colville RN and Nieuwenhuijsen MJ (2006). Human exposure modelling for chemical risk assessment: a review of current approaches and research and policy implications. *Environ Sci Policy* 26:261–274.

Gad SC and Weil CS (1986). *Statistics and experimental design for toxicologists*. New Jersey: Telford Press.

Gad SC and Weil CS (1989). Statistics for toxicologists. In AW Hayes (Ed.). *Principles and methods of toxicology*. 3rd edn. New York: Raven Press, pp. 221–274.

Gaylor DW and Gold LS (1998). Regulatory cancer risk assessment based on a quick estimate of a benchmark dose derived from the maximum tolerated dose. *Regul Toxicol Pharmacol* 28:222–225.

Gaylor GW, Kodell RL, Chen JJ and Krewski D (1999). A unified approach to risk assessment for cancer and noncancer endpoints based on benchmark doses and uncertainty/safety factors. *Regul Toxicol Pharmacol* 29:151–157.

Gilbert RO (1987). *Statistical methods for environmental pollution monitoring*. New York: Van Nostrand Reinhold.

Glik DC (2007). Risk communication for public health emergencies. *Ann Rev Public Hlth* 28:33–59.

Golden RJ, Holm SE, Robinson D, Julkunen PH and Reese EA (1997). Chloroform mode of action: Implications for cancer risk assessment. *Regul Toxicol Pharmacol* 26:142–155.

Goldfrank LR, Flomenbaum NE, Lewin NA, Weisman RS and Howland MA (Eds). (1990).

Goldfrank's toxicologic emergencies. East Norwalk, Connecticut: Prentice-Hall.

Goldstein BD (2010). Risk assessment of environmental chemicals: If it ain't broke ... *Risk Analysis* published online 8 July 2010 DOI:10.1111/j.1539-6924.2010.01447x

Good J (1993). Towards standardised sampling and analysis for contaminated sites. In AJ Langley, M van Alphen (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission. pp. 43–49.

Government/Research Councils Initiative on Risk assessment and Toxicology (1999). *Risk assessment strategies in relation to population sub-groups*. Report of a workshop organised by the Risk assessment and Toxicology Steering Committee. Leicester, UK: Institute for Environment and Health.

Gray MN, Aylward LL and Keenan RE (2006). Relative cancer potencies of selected dioxin-like compounds on a body-burden basis: Comparison to current toxic equivalency factors (TEFs). *J Toxicol Environ Health Part A* 69:907–17.

Greenhalgh T (1997). Papers that summarise other papers: Systematic reviews and meta-analysis. *British Medical Journal*. 315:672–675.

Gregg MB (Ed.) (1996). *Field epidemiology*. New York: Oxford University Press.

Grice HC and Burek JD (1983). Age associated (geriatric) pathology: Its impact on long term toxicity studies. In *Current Issues in Toxicology*, New York: Springer Verlag.

Groten JP, Feron VJ and Sühnel J (2001). Toxicology of simple and complex mixtures. *Trends Pharmacol Sci* 22:316–322.

Guyton AC (1991) *Textbook of medical physiology*, 8th edn. Philadelphia: WB Saunders Company.

Guyton KZ, Kyle AD, Aubrecht J, Cogliano VJ, Eastmond DA, Jackson M, Keshava N, Sandy MS, Sonanwane B, Zhang L, Waters MD and Smith MT (2009). Improving prediction of chemical carcinogenicity by considering multiple mechanisms and applying toxicogenomic approaches. *Mutation Res*. 681:230–240.

Haas CN, Rose JB and Gerba CP (1999). *Quantitative microbial risk assessment*. New York: Wiley.

Haber LT, Maier A, Gentry PR, Clewell HJ, Dourson ML (2002). Genetic polymorphisms in assessing interindividual variability in delivered dose. *Regul Toxicol Pharmacol*. 35:177–197.

Haddad S, Béliveau M, Tardif R and Krishnan K (2001). A PBPK modeling-based approach to account for interactions in the health risk assessment of chemical mixtures. *Toxicol Sci* 63:125–131.

Haimes YY, Barry T, Lambert JH (Eds) (1994). When and how you can specify a probability distribution when you don't know much? Proceedings of a Workshop organised by the US EPA and University of Virginia,1993, Charlottesville, Virginia. In *Risk Analysis*, 14:661–706.

Hamm AK, Carter WH and Gennings C (2005). Analysis of an interaction threshold in a mixture of drugs and/or chemicals. *Statist Med*. 24:2493–2507.

Haschek WM, Rousseaux CG and Wallig MA (Eds) (2002). *Handbook of toxicologic pathology*. Vol 1. Academic Press, San Diego, CA.

Haseman JK, Huff J and Boorman GA (1984). Use of historical control data in carcinogenicity studies in rodents. *Toxicol Pathol* 12:126–135.

Hattis D and Silver K (1994). Human inter-individual variability. A major source of uncertainty in assessing risks for non cancer health effects. *Risk Analysis*, 14:421–432.

Hattis D, Goble, R and Chu M (2005). Age-related differences in susceptibility to carcinogenesis. II. Approaches for application and uncertainty analyses for individual genetically acting carcinogens. *Environ Health Perspect*, 113:509–516.

Hattis D, Goble, R, Russ A, Chu M and Ericson (2004). Age-related differences in susceptibility to carcinogenesis: A quantitative analysis of empirical animal bioassay data. *Environ Health Perspect* 112:1152–1158.

Havelaar AH and Melse JM (2003). *Quantifying public health risk in the WHO Guidelines for drinking-water quality: A burden of disease approach*. RIVM report 734301022, Rijksinstituut voor Volksgezondheid en Milieu, Netherlands.

Havelaar AH, de Wit MA, van Konigsveld R and van Kempen E (2000). *Campylobacter* spp. *Epidemiol Infection* 125:505–522.

Haws LC, Su SH, Harris M, DeVito MJ, Walker NJ, Farland WH, Finley B and Birnbaum LS (2006). Development of a refined database of mammalian relative potency estimates for dioxin-like compounds. *Toxicol Sci*. 89:4–30.

Hayes AW (Ed.) (1994). *Principles and methods of toxicology*. 3rd edn. New York: Raven Press.

Hays SM, Aylward LL, LaKind JS, Bartels MJ, Barton HA, Boogaard PJ, Brunk C, DiZio S, Dourson M, Goldstein DA,

Lipscomb J, Kilpatrick ME, Krewski D, Krishnan K, Nordberg M, Okino M, Tan Y-M, Viau C and Yager JW (2008). Guidelines for the development of biomonitoring equivalents: Report from the Biomonitoring Equivalents Workshop. *Regul Toxicol Pharmacol* 51:S4–S15.

Health and Safety Executive (HSE) (1999). *Reducing risks, protecting people: HSE's decision-making process*. <http://www.hse.gov.uk/risk/theory/r2p2.pdf>.

Health and Safety Executive (HSE) (2010). Improving health and safety: An analysis of HSE's risk communication in the 21st century. (RR785). UK: London: Prepared by King's College for the Health and Safety Executive.

Health Canada (1994). *Human health risk assessment for priority substances: Canadian Environmental Protection Act*. Canada Communications Group – Publishing.

Health Canada (2004a) *Guidelines for Canadian drinking water quality: Supporting documentation – protozoa: Giardia and Cryptosporidium*. Water Quality and Health Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Health Canada (2004b) *Guidelines for Canadian Drinking Water Quality: Supporting Documentation – Enteric viruses*. Water Quality and Health Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario.

Health Council of the Netherlands (2003). Benchmark dose method: Health-based recommended exposure limits in new perspective. The Hague: Health Council of the Netherlands. Publication No. 2003/06E.

Heath L, Pollard SJT, Hrudey SE and Smith G (2010). *Engaging the community: A handbook for professionals managing contaminated land*, CRC for Contamination Assessment and Remediation of the Environment, Adelaide, Australia.

Hellard ME, Sinclair MI, Forbes AB and Fairley CK (2001). A randomized, blinded, controlled trial investigating the gastrointestinal health effects of drinking water quality. *Environ Health Perspect* 109: 773–778.

Helsel DR (1990). Less than obvious: Statistical treatment of data below the detection limit. *Environ Sci Technol* 24:1767–1774.

Hemminki K (1992). Significance of DNA and protein adducts. In H Vainio, DB McGregor and AJ McMichael (Eds) *Mechanisms of carcinogenesis in risk identification*. International Agency for Research of Cancer, Lyon. pp. 525–534.

Henry JB (Ed) (1984). *Todd-Sanford-Davidsohn clinical diagnosis and management by laboratory methods*. 17th edn. Philadelphia: Saunders Co.

Hernandez LG, van Steeg H, Luitjen M and van Benthem J (2009) Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *Mutation Research* 682:94–109.

Herrington LP and Nelbach JH (1942). Relation of gland weights to growth and aging processes in rats exposed to certain environmental conditions. *Endocrinol* 30:375–386.

Herzberg RC and Teuschler LK (2002). Evaluating quantitative formulas for dose-response assessment of chemical mixtures. *Environ Health Perspect* 110 (Suppl. 6):965–970.

Heywood R (1981). Target organ toxicity. *Toxicol Lett* 8:349–358.

Heyworth JS (1991). Sampling and statistical analysis for assessing contaminated land sites. In O El Saadi, AJ Langley. (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Higgins JPT, Green S (2011) (Eds). *Cochrane handbook for systematic reviews of interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from <www.cochrane-handbook.org>.

Hill AB (1965). The environment and disease: Association or causation? *Proc Royal Soc Med*. 58, 295–300.

Hill RN, Clemens TL, Liu DK, Vesell ES and Johnson WD (1975). Genetic control of chloroform toxicity in mice. *Science* 190:159–161.

Hines RN, Sargent D, Autrup H, Birnbaum LS, Brent RL, Doerrer NG, Cohen Hubal EA, Juberg DR, Lauranet C, Luebke R, Olejniczak K, Portier CJ and Slikker W (2010). Approaches for assessing risks to sensitive populations: Lessons learned from evaluating risks in the pediatric population. *Tox Sci* 113:4–26.

Hrudey SE (1998). Quantitative cancer risk assessment – pitfalls and progress. In RE Hester and RM Harrison. *Issues in environmental science and technology*, No. 9. Cambridge, UK: Royal Society of Chemistry Risk Assessment and Risk Management.

Hrudey SE (2009). Chlorination disinfection by-products, public health risk tradeoffs and me. *Water Research*. 43:2057–2092.

Hrudey SE, Bathija A, Bull RJ, Dodds L, Jenkins G, King W, Murphy P, Narotsky M, Reif J, Savitz D and Thomas B (2003). *Executive summary. International expert workshop for Health Canada to identify critical endpoints for assessment of health risks related to Trihalomethanes (THMs) in drinking water*. March 2003.

Hrudey SE and Hrudey EJ (2004). *Safe drinking water: Lessons learned from recent outbreaks in affluent nations*. London IWA.

Hrudey SE and Krewski D (1995). Is there a safe level of exposure to a carcinogen? *Envir Sci Technol* 29:370A-375A.

Hrudey S. and Leiss W (2003). Risk management and precaution: Insights on the cautious use of evidence. *Environ Health Perspect* 111: 1577–1581.

Hrudey SE, Chen W and Rousseaux CG (1996). *Bioavailability in environmental risk assessment*. Boca Raton, Florida: Lewis Publishers/CRC Press.

Hubal EAC (2009). Forum series – Part VI: Biologically relevant exposure science for 21st century toxicity testing. *Toxicol Sci*. 111:226–232.

Hunt VR (1982). *Work and the health of women*. Boca Raton, Florida: CRC Press.

IGHRC (2004) Guidelines for good exposure assessment practice for human health effects of chemicals. (cr10) MRC Institute for Environment and Health, University of Leicester.

IGHRC (2009) Chemical mixtures: A framework for assessing risk to human health (cr14). Institute of Environment and Health, UK: Cranfield University.

Innami S, Yang MG, Mickelsen O and Hafs HD (1973). The influence of high fat diets on estrous cycles, sperm production and fertility of rats. *Proc Soc Exp Biol Med* 143:63–68.

Institute for Environment and Health (IEH) (1999a). Developing new approaches to assessing risk to human health from chemicals. Risk and Toxicology Steering Committee. Leicester: IEH

Institute for Environment and Health (IEH) (1999b). Risk assessment approaches used by UK Government for evaluating human health effects of chemicals. Leicester: IEH.

Institute for Environment and Health (IEH) (1999c). Risk assessment strategies in relation to population sub groups. Leicester: IEH.

Inter-Departmental Liaison Group on Risk Assessment (ILGRA) (2002). *The precautionary principle: policy and application*. <http://www.hse.gov.uk/aboutus/meetings/committees/ilgra/pppa.htm>.

International Agency for Research on Cancer (IARC) (1978). Some N-Nitroso compounds. Vol. 17. *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*. Lyon: IARC.

International Agency for Research on Cancer (IARC) (1991). Chlorinated drinking-water; chlorination by-products; some other halogenated compounds; cobalt and cobalt compounds. Vol. 52. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. Lyon: IARC.

International Agency for Research on Cancer (IARC) (1999). *The use of short- and medium-term tests for carcinogens and data on genetic effects in carcinogenic hazard evaluation*. No. 146. IARC Scientific Publications. Lyon: IARC.

International Agency for Research on Cancer (IARC) (2004). Some drinking-water disinfectants and contaminants, including arsenic. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*. Lyon: IARC.

International Agency for Research on Cancer (IARC) (2006). Formaldehyde, 2-butoxyethanol and 1-*tert*-butoxypropanol-2-ol. Vol. 88. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Lyon: IARC.

International Life Sciences Institute (ILSI) (1996). Risk Science Institute Pathogen Risk Assessment Working Group. A conceptual framework to assess the risks of human disease following exposure to pathogens. *Risk Analysis*. 16:841–848. Blackwell Scientific.

International Life Sciences Institute (ILSI) (2000). Risk Science Institute. Revised Framework for Microbial Risk Assessment. Summary Report of an ILSI Risk Science Institute workshop.

International Programme on Chemical Safety (IPCS) (2000). Disinfectants and disinfectant by-products. *Environmental Health Criteria 216*. Geneva: WHO.

International Programme on Chemical Safety (IPCS) (2005). *Chemical specific adjustment factors for interspecies differences and human variation: Guidance document for use of data in dose/concentration-response assessment*. Harmonization Project Document No. 2. Geneva: WHO.

International Programme on Chemical Safety (IPCS) (2008). *Uncertainty and data quality in exposure assessment. Part 1: Guidance document on characterizing and communicating uncertainty in exposure assessment. Part 2: Hallmarks of data quality in chemical exposure assessment*. Harmonization Project Document No. 6. Geneva: WHO.

International Programme on Chemical Safety (IPCS) (2009a). *Assessment of combined exposures to multiple chemicals*. Report of a WHO/IPCS International Workshop on aggregate/cumulative risk assessment. Harmonization Project Document No. 7. Geneva: WHO.

International Programme on Chemical Safety (IPCS) (2009b). *Risk assessment of combined exposures to multiple chemicals: A WHO/IPCS framework*. Harmonization program draft document for public comment and peer review. Geneva: WHO.

International Programme on Chemical Safety (IPCS) (2009c). Principles for modelling dose–response for the risk assessment of chemicals. *Environmental Health Criteria 239*. Geneva: WHO.

Jardine C, Hrudey S, Shortreed J, Craig L, Krewski D, Fugal C and McColl S (2005). Risk management frameworks for human health and environmental risks. *J Toxicol Envir Health Part B* 6:569–718.

Johannsen (1990). Risk assessment of carcinogenic and noncarcinogenic chemicals. *Crit Rev Toxicol* 20:341–367.

Johnson PC and Ettinger RA (1991). Heuristic model for predicting the intrusive rate of contaminant vapours into buildings. *Envir Sci Technol* 25:1445–1452.

Joint FAO/WHO Expert Committee on Food Additives (JECFA) (2002). Evaluation of certain food additives and contaminants. WHO Technical Report Series. No. 909.

Jorgenson TA, Meierhenry JR, Rushbook CJ, Bull RJ and Robinson M (1985). Carcinogenicity of chloroform in drinking water to male Osborne-Mendel rats and female B6C3F1 mice. *Fund Appl Toxicol* 5:760–769.

Kaiser J (2003). Hormesis: Sipping from a poisoned chalice. *Science* 302:376–39.

Keenan KP, Laroque P and Dixit R (1998). Need for dietary control by caloric restriction in rodent toxicology and carcinogenicity studies. *J Toxicol Environ Health B Crit Rev* 1:135–148.

Keith LH (1990). Environmental sampling: A summary. *Environ Sci Technol*, 24:610–617.

Keith LH (Ed.) (1988). *Principles of environmental sampling*. Washington, DC: American Chemical Society.

Kelly KE (1991). *The myth of 10^{-6} as a definition of 'acceptable risk'*. Presented at the 84th Annual Meeting and Exhibition of the Air and Waste Management Association, Vancouver, BC, 16–21 June.

Kennedy GC (1969). Interactions between feeding behavior and hormones during growth. *Ann NY Acad Sci* 157:1049–1061.

Kennedy Jr GI and Valentine R (1994). Inhalation toxicology. In AW Hayes (Ed.), *Principles and methods of toxicology*, 3rd edn. New York: Raven Press. pp. 805–838.

Kim YJ and Platt U (2008). Advanced environmental monitoring. Netherlands: Springer. Available online at Springerlink <http://www.springerlink.com>.

Klaassen CD (Ed) (1996). *Casarett and Doull's toxicology: The basic science of poisons*, 5th edn. New York: McGraw-Hill. p. 4.

Klinger MM, MacCarter GD and Boozer CN (1996). Body weight and composition in the Sprague-Dawley rat: Comparison of three outbred sources. *Lab Animal Sci* 46:67–69.

Knafla A, Phillipps KA, Brecher RW, Petrovic S and Richardson M (2006). Development of a dermal cancer slope factor for benzo[*a*]pyrene. *Regul Pharmacol Toxicol* 45:159–168.

Korach KS (Ed) (1998). *Reproductive and developmental toxicology*. New York: Marcel Dekker.

Kortenkamp A, Backhaus T and Faust M (2009). *State of the art report on mixture toxicity*. Report to the Commission of the European Union (Directorate General for the Environment).

Kortenkamp A, Faust M, Scholze M and Backhaus T (2007). Low-level exposure to multiple chemicals: reason for human health concerns? *Environ Health Perspect* 115:Suppl 1 106–114.

Krewski D, Gaylor D, Soms, A and Szyskowicz, M (1993). An overview of the report: Correlation between carcinogenic potency and the maximum tolerated dose: Implications for risk assessment. *Risk Anal* 13: 383–398.

Kroes R, Galli C, Munro I, Schilter B, Tran L-A, Walker R and Wurtzen G (2000). Threshold of toxicological concern for chemical substances in the diet: A practical tool for assessing the need for toxicity testing. *Food Chem Toxicol* 38:255–312.

Kroes R, Renwick AG, Cheeseman M, Kleiner J, Mangelsdorf I, Piersma AB. Schilter B, Schlatter J, van Schothorst F, Vos JG and G Wurtzen G (2004). Structure-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. *Food Chem Toxicol* 42:65–83.

Lambert JC and Lipscomb JC (2007). Mode of action as a determining factor in additivity models for chemical mixture risk assessment. *Regul Toxicol Pharmacol* 49:183–194.

Langley AJ (1991a). Exposure pathways. In O El Saadi, AJ Langley (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Langley AJ (1991b). Protecting the public: Biological monitoring. In O El Saadi, AJ Langley (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Langley AJ (1993a). Refining exposure assessment. In AJ Langley, M Van Alphen (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Langley AJ (1993b). The representation of data and the appraisal of contaminated land. In AJ Langley, M Van Alphen (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Langley A (2003) What does it mean when the risk assessment says 4.73×10^{-5}? *NSW Public Hlth Bull* 14:166–167.

Langley AJ, Sabordo L (1996). Exposure factors in risk assessment. In A Langley, B Markey and H Hill (Eds), *The health risk assessment and management of contaminated sites*. Adelaide: South Australian Health Commission.

Langley AJ, Taylor A and Dal Grande E (1998). 1996 Australian exposure factors. In A Langley, P Imray, W Lock and H Hill (Eds), *The health risk assessment and management of contaminated sites*, Adelaide. South Australian Health Commission.

Larson JL, Wolf DC and Butterworth BE (1994). Induced cytotoxicity and cell proliferation in the hepatotoxicity of chloroform in female B6C3F1 mice: Comparison of administration by gavage in corn oil vs. ad libitum in drinking water. *Fund Appl Toxicol* 22:90–102.

Larson JL, Wolf DC, Morgan MK and Butterworth BE (1995). Toxicity and cell proliferation in the liver, kidneys and nasal passages of female F-344 rats, induced by chloroform administered by gavage. *Food Chem Toxicol* 33:443–456.

Last JM (1988). *A dictionary of epidemiology*. 2nd edn. Oxford: Oxford University Press.

Last JM (2010). Connecting the dots: The power of words and the diversity of epidemiological information. *J Epidemiol Community Hlth* 64:106–108.

Lauwerys R (1984). Basic concepts of human exposures. In A Berlin, M Draper, K Hemminki and H Vainio (Eds), *Monitoring exposure to carcinogenic and mutagenic agents*. IARC Scientific Publications No. 59. Lyon: IARC. pp. 31–36.

Lead Safe (1997). *A guide for health care professionals*. Lead Education Program. NSW: NSW Lead Reference Centre.

Lee PN (1993). Statistics. In D Anderson and DM Conning (Eds), *Experimental toxicology: The basic issues*. Cambridge, UK: Royal Society of Chemistry.

Levy B and Wegman DH (1988). *Occupational health: Recognising and preventing work-related disease*. 2nd edn. Little Brown, Boston.

Levy G (1968). Dose dependent effects in pharmacokinetics. In *Important fundamental principles in drug Evaluation* pp. 141–172.

Lewis SC, Lynch JR and Nikiforov AI (1990). A new approach to deriving community exposure guidelines for 'no observed adverse effect levels'. *Regul Toxicol Pharmacol* 3:17–34.

Lilienfield AM and Lilienfield DE (1980). *Foundations of epidemiology*. 2nd edn. Oxford University Press.

Lock WH (1996). *Composite sampling*. National Environmental Health Monographs Soil Series No. 3. Adelaide: South Australian Health Commission.

Lord DW (1987). Appropriate site investigations. In T Cairney (Ed), *Reclaiming contaminated land*. Glasgow: Blackie.

Lovell DP and Thomas G (1996). Quantitative risk assessment and the limitations of the linearised multistage model. *Human Exptl Toxicol* 15:87–104.

Lu FC and Sielken Jr RL (1991). Assessment of safety/risk of chemicals: Inception and evolution of the ADI and dose–response modelling procedures. *Toxicology Letters* 59:5–40.

Lucas RM and McMichael AJ (2005). Association or causation: Evaluating links between 'environment and disease'. Bulletin WHO 83:792–795.

McAuley C and Hrudey S (2006). Towards meaningful stakeholder comprehension of sour gas facility risk assessments. *J Environ Eng Sci* 5:1–11.

McCarty LS and Borgert CJ (2006a). Review of the toxicity of chemical mixtures containing at least one organochlorine. *Regul Toxicol Pharmacol* 45:104–118.

McCarty LS and Borgert CJ (2006b). Review of the toxicity of chemical mixtures: theory, policy, and regulatory practice. *Regul Toxicol Pharmacol* 45:119–143.

MacDonald J, French JE, Gerson RJ, Goodman J, Inoue T, Jacobs A, Kasper P, Keller D, Lavin A, Long G, McCullough B, Sistare FD, Storer R and van der Laan JW (2004). The utility of genetically modified mouse assays for identifying human carcinogens: A basic understanding and path forward. *Toxicol Sci* 77:188–194.

Mackay D (1991). *Multimedia environmental models: the fugacity approach*. Michigan: Lewis Publishers.

McClellan RO and Henderson RF (Eds) (1989). *Concepts in inhalation toxicology*. Hemisphere Publishing Corp, New York.

McKone TE (1993). CalTOX. A multimedia total-exposure model for hazardous-wastes sites. Part 1: Executive Summary prepared for the State of California. Department of Toxic Substances Control UCRL-CR-111456 Pt 1. Livermore, CA: Lawrence Livermore National Laboratory.

McKone TE (1994). Uncertainty in variability in human exposures to soil contaminants through home grown food: A Monte Carlo assessment. *Risk Analysis*, 14:449–463.

McLean AEM and McLean EK (1969) Diet and toxicity. *Br Med Bull* 25:278–28.

McMichael AJ (1991). Setting environmental exposure standards: current concepts and controversies. *Int J Envir Hlth Res* 1:2–13.

Manahan SE (1993). *Fundamentals of Environmental Chemistry*. Boca Raton, Florida: Lewis Publishers.

Mason AM, Borgert CJ, Bus JS, Mumtaz MM, Simmons JE and Sipes IG (2007). Improving the scientific foundation for mixtures joint toxicity and risk assessment: Contributions from the SOT mixtures project. *Toxicol Appl Pharmacol* 223:99–103.

Masoro EJ (1992). Aging and proliferative homeostasis: Modulation by food restriction in rodents. *Lab Animal Sci* 42:132–137.

Meek ME (Bette), Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehmann-McKeeman LD, Longfellow DG, Pastoor T, Seed J and Patton DE (2003). A framework for human relevance analysis of information on carcinogenic modes of action. *Crit Rev Toxicol*. 33:591–653.

Mena KD, Gerba CP, Haas CN and Rose JB (2003). Risk assessment of waterborne coxsachievirus. *J Amer Water Works Assoc*. 95:122–131.

Menzie CA, MacDonell MM and Mumtaz M (2007). A phased approach for assessing combined effects from multiple stressors. *Environ Health Perspect* 115:807–816.

Messner MJ, Chappell CL and Okhuysen PC (2001). Risk assessment for *Cryptosporidium*: a hierarchical Bayesian analysis of human response data. *Water Research* 35:3934–3940.

Mills CJ, Bull RJ, Cantor KP, Reif J, Hrudey SE and Huston P (1998). Health risks of drinking water chlorination by-products: Report of an expert working group. *Chronic diseases in Canada* 19:91–102.

Moerbeek M, Piersma AH and Wout Slob W (2004). A comparison of three methods for calculating confidence intervals for the benchmark dose. *Risk Analysis* 24:31–40.

Mohr U, Dungworth D, Kimmerle G, Lewkowski J, McClellan R and Stöber W (Eds) (1988). *Inhalation toxicology: The design and interpretation of inhalation studies and their use in risk assessment*. ILSI Monograph. New York: Springer-Verlag.

Mohr U, Dungworth DL and Capen CC (Eds) (1992). *Pathobiology of the Aging Rat. Vol. 1*. Washington, DC: ILSI Press.

Mohr U, Dungworth DL and Capen CC (Eds) (1994) *Pathobiology of the Aging Rat. Vol. 2*. Washington, DC: ILSI Press.

Mohr U, Dungworth DL, Ward J, Capen CC, Carlton W and Sundberg J (Eds) (1996). *Pathobiology of the Aging Mouse. Vols 1 and 2*. Washington, DC: ILSI Press.

Monosson E (2005). Chemical mixtures: considering the evolution of toxicology and chemical assessment. *Env Health Perspect* 113:383–390.

Moolgavkar SH, Moller H and Woodward A (1999). Principles of the epidemiological approach to QEP. In SH Moolgavkar, L Krewski, L Zeise et al. (Eds), *Quantitative estimation and prediction of human cancer risks*. Lyon: International Agency for Research on Cancer, Scientific Publication 131.

Morisseau C, Merzlikin O, Lin A, He G, Feng W, Padilla I, Denison MS, Pessah IN and Hammock BD (2009). Toxicology in the fast lane: Application of high-throughput bioassays to detect modulation of key enzymes and receptors. *Environ Health Perspect* 117:1867–1872.

Mumtaz MM, Ruiz P and De Rosa, CT (2007). Toxicity assessment of unintentional exposure to multiple chemicals. *Toxicol Appl Pharmacol* 223:104–113.

Mundt KA, Tritschler JP, Dell LD (1998). Validity of epidemiological data in risk assessment applications. *Human Ecological Risk Assessment*. 4:675–683.

Munro IC (1977). Considerations in chronic toxicity testing: The chemical, the dose, the design. *J Environ Path Toxicol* 1:183–197.

Munro IC (1990). Safety assessment procedures for indirect food additives: an overview. *Regul Toxicol Pharmacol* 12:2–12.

Munro IC, Ford RA, Kennepohl E and Sprenger JG (1996). Correlation of structural class with no-observed-effect-levels: A proposal for establishing a threshold of concern. *Food Chem Toxicol* 34:829–867.

Munro IC, Kennepohl E and Kroes R (1999). A procedure for the safety evaluation of flavouring substances. *Food Chem Toxicol* 37:207–232.

Murphy BL (1998). Dealing with the uncertainty in risk assessment. *Human and Ecological Risk Assessment*. 4 (3):685–699.

Mushak P (2007). Hormesis and its place in non-monotonic dose–response relationships: Some scientific reality checks. *Environ Health Perspect*. 115:500–506.

Nachamkin I, Engberg J, Gutacker M, Meinersman RJ, Li CY, Arzate P, Teeple E, Fussin V, Ho TW, Asbury AK, Griffin JW, McKhann GM and Pifaretti JC (2001). *Campylobacter jejuni* HS:19 associated with Guillan-Barre syndrome and gastroenteritis. *J Infectious Dis* 184:221–226.

Nadkarni RA (1991). The quest for quality in the laboratory. *Analytical Chemistry* 63(13): 675A.

National Academy of Science (NAS) (1991). *Human exposure assessment for airborne pollutants*. Washington, DC: National Academy Press.

National Cancer Institute (NCI) (1976). Carcinogenesis Bioassay of Chloroform. Bethesda, MD: National Cancer Institute.

National Environment Protection Council (NEPC) (2009). An Australian approach to setting air quality standards. Technical Support Document <http://www.ephc.gov.au/airquality/aaq_nepm>.

National Environment Protection Council (NEPC) (2010). *National Environment Protection Measure*. Schedule B4. Guideline on Health Risk Assessment Methodology.

National Environment Protection Council (NEPC) (2011). Methodology for setting air quality standards in Australia. Part A. <http://www.ephc.gov.au/node/4>.

National Health and Medical Research Council (NHMRC) (1999a). *Toxicity assessment for carcinogenic soil contaminants*. Canberra: Commonwealth of Australia.

National Health and Medical Research Council (NHMRC) (1999b). *A guide to the development, implementation and evaluation of clinical practice guidelines*. Canberra: Australian Government Publishing Service.

National Health and Medical Research Council (NHMRC) (2000). *How to review the evidence: Systematic identification and review of the scientific literature*. Canberra: Commonwealth of Australia.

National Health and Medical Research Council (NHMRC) (2006). *Ambient air quality standards setting: An approach to health-based hazard assessment*. Canberra: Commonwealth of Australia.

National Health and Medical Research Council (NHMRC) (2008). *Guidelines for managing risk in recreational water*. Canberra: Commonwealth of Australia.

National Health and Medical Research Council and Natural Resource Management Ministerial Council (NHMRC, NRMMC) (2004). *Australian drinking water guidelines*. Canberra: NHMRC and NRMMC.

National Research Council (NRC) (1983). *Risk assessment in the federal government: managing the process*. Washington, DC: National Academy Press.

National Research Council (NRC) (1992). *Biological markers in immunotoxicology*. Washington, DC: National Academy Press.

National Research Council (NRC) (1993). *Pesticides in the diets of infants and children*. Washington, DC: National Academy Press.

National Research Council (NRC) (1994). *Science and judgement in risk assessment*. Washington, DC: National Academy Press.

National Research Council (NRC) (2003). *Bioavailability of contaminants in soils and sediments: processes, tools, and applications*. Washington, DC: The National Academies Press.

National Research Council (NRC) (2006). *Toxicity testing for assessment of environmental agents*: interim report. Washington, DC: National Academy Press.

National Research Council (NRC) (2007). *Toxicity testing in the 21st century: A vision and a strategy*. Washington, DC: National Academy Press.

National Research Council (NRC) (2008). *Science and decisions: Advancing risk assessment*. Washington, DC: National Academy Press.

National Research Council (NRC) (2010). *Review of the Environmental Protection Agency's draft IRIS assessment of tetrachlorethylene*. <http://www.nap.edu/catalog/12863.html>.

National Toxicology Program (NTP) (2006). NTP toxicology and carcinogenesis studies of a mixture of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) (CAS No. 1746-01-6), 2,3,4,7,8-Pentachlorodibenzofuran (PeCDF) (CAS No. 57117-31-4), and 3,3',4,4',5-Pentachlorobiphenyl (*PCB 126) (CAS No. 57465-28-8) in Female Harlan Sprague-Dawley Rats (Gavage Studies). Natl Toxicol Program Tech Rep Ser.* September 526:1–180.

Natural Resource Management Ministerial Council, Environment Protection and Heritage Council, Australian Health Minister's Conference (NRMMC, EPHC, AHMC) (2006) *Australian guidelines for water recycling: managing health and environmental risks. Phase 1*. Canberra: Australian Government.

Natural Resource Management Ministerial Council, Environment Protection and Heritage Council, National Health and Medical Research Council (NRMMC, EPHC, NHMRC) (2008). *Australian guidelines for water recycling: managing health and environmental risks: Augmentation of drinking water supplies*. Canberra: Australian Government.

Natural Resource Management Ministerial Council, Environment Protection and Heritage Council, National Health and Medical Research Council (NRMMC, EPHC, NHMRC) (2009a) *Australian guidelines for water recycling: managing health and environmental risks: Stormwater harvesting and reuse*. Canberra: Australian Government.

Natural Resource Management Ministerial Council, Environment Protection and Heritage Council, National Health and Medical Research Council (NRMMC, EPHC, NHMRC) (2009b) *Australian guidelines for water recycling: managing health and environmental risks: Managed aquifer recharge*. Canberra: Australian Government.

New Zealand Ministry for the Environment (2003). *Good practice guide for assessing and managing odour in New Zealand*. Air Quality Report 36. Wellington NZ: Ministry for the Environment.

Nolen G (1972). Effects of various restricted dietary regimens on the growth, health and longevity of albino rats. *J Nutr* 102:1477–1494.

New South Wales Department of Environment and Conservation (NSW DEC) (2006). *Assessment and management of odour from stationary sources in NSW*. NSW: Department of Environment and Conservation.

Nurminen M (1995). Linkage failures in ecological studies. *World Health Statistical Quarterly*. 48:78–84.

Office of Chemical Safety (OCS) (2004). Human health risk assessment of dioxins in Australia. National dioxins program. Technical Report No. 12. Australian Government Department of Health and Ageing, Canberra: Australian Government Department of the Environment and Heritage.

Office of the Gene Technology Regulator (OGTR) (2009) *Risk analysis framework*. Canberra: Department of Health and Ageing.

Organisation for Economic Cooperation and Development (OECD) (1982). *Good laboratory practice in the testing of chemicals*. Paris: OECD, p. 62.

Organisation for Economic Cooperation and Development (OECD) (2009). *OECD guidelines for the testing of chemicals*. Section 4: Health effects. <http://www.oecd.org/document/40/0,3343,en_2649_34377_37051368_1_1_1_1,00.html>.

Paustenbach DJ (1989). A survey of health risk assessment. In DJ Paustenbach (Ed.), *The risk assessment of environmental and human health hazards: A textbook of case studies*. New York: John Wiley & Sons.

Paustenbach DJ (1995). The practice of health risk assessment in the United States (1975–1995): How the US and other countries can benefit from that experience. *Human and Ecological Risk Assessment*, 1:29–79.

Paustenbach DJ (2002). *Human and ecological risk assessment: Theory and practice*. New York: John Wiley & Sons.

Paynter OE (1984). Oncogenic potential: Guidance for analysis and evaluation of long-term rodent studies. Evaluation Procedure #1000.1 Office of Pesticide and Toxic Substances. Washington, DC: EPA.

Perkins JL (1997). *Modern Industrial Hygiene. Vol. 1. Recognition and evaluation of chemical agents*. New York: Van Nostrand Reinhold.

Pershagen G (1999). Research priorities in environmental health. *Brit Med J* 318:1636–1637.

Peters RG, Covello VT and McCallum DB (1997). The determinants of trust and credibility in environmental risk communication: An empirical study. *Risk Analysis* 17:43–54.

Pigeon N, Kasperson RE and Slovic P (2003). *The social amplification of risk*. Cambridge: Cambridge University Press.

Pohl HR and Abadin HG (2008). Chemical mixtures: Evaluation of risk for child-specific exposures in a multi-stressor environment. *Toxicol Appl Pharmacol*. 233:116–125.

Pohl HR, Mumtaz MM, Scinicariello F and Hansen H (2009). Binary weight-of-evidence evaluations of chemical interactions – 15 years of experience. *Regul Pharmacol Toxicol* 54:264–271.

Polgar S and Thomas SA (1991). *Introduction to research in the health sciences*. 2nd edn. Melbourne: Churchill Livingstone.

Pollard SJ, Yearsley R, Reynard N, Meadowcroft IC, Duarte-Davidson R and Duerden SL (2002). Current directions in the practice of environmental risk assessment in the United Kingdom. *Environ Sci Technol* 36:530–538.

Pontius FW (2000). Chloroform: Science, Policy and Politics. *Journal American Water Works Association* 92:12–18.

Presidential/Congressional Commission on Risk Assessment and Risk Management (P/CCRARM) (1997). *Risk assessment and risk management in regulatory decision making*. P/CCRARM.

Purchase IFH and Auton TR (1995). Thresholds in chemical carcinogenesis. *Regul Toxicol Pharmacol* 22:199–205.

Queensland Government (2001a). *A guide to effective community engagement*. Community Engagement Unit Strategic and Executive Services. <http://www.emergency.qld.gov.au/publications/pdf/guide_to_eff_comm_engagement.pdf>.

Queensland Government (2001b). *Charter for Community Engagement*. Department of Emergency Services. <http://www.emergency.qld.gov.au/publications/pdf/charter_for_ce.pdf>.

Queensland Health (2009). *Queensland Health guidelines: Assessment of clusters of non-communicable disease*. Available at <www.qld.gov.au>.

Reckelhoff-Dangel C and Petersen D (2007). *Risk communication in action: The risk communication workbook*. US EPA Office of Research & Development, EPA/625/R-05/003.

Rees N (1999). Current dietary exposure assessment methodology. In Institute for Environment and Health (IEH), *Risk assessment strategies in relation to population sub groups*. Leicester: Institute for Environment and Health.

Renner R (2004). Redrawing the dose–response curve. *Environ Sci Technol*. 38:90A-95A.

Renwick AG (1999) Human variability and risk assessment/safety assurances. In Institute for Environment and Health (IEH), *Risk assessment strategies in relation to population sub groups*. Leicester: Institute for Environment and Health.

Renwick AG (2004). Toxicology databases and the concept of thresholds of toxicological concern as used by the JECFA for the safety evaluation of flavouring agents. *Toxicology Letters* 149:223–234.

Renwick AG (2005). Structure-based thresholds of toxicological concern: guidance for application to substances present at low levels in the diet. *Toxicol Appl Pharmacol* 207 (Supplement 1): 585–591.

Renwick AG and Walker R (1993). An analysis of the risk of exceeding acceptable or tolerable daily intake. *Regul Toxicol Pharmacol* 18:463–480.

Rice G, MacDonell M, Hertzberg RC, Teuschler L, Picel K, Butler J, Chang Y-S and Hartmann H (2008). An approach for assessing human exposures to chemical mixtures in the environment. *Toxicol Appl Pharmacol* 233:126–136.

Rider CV and LeBlanc GA (2005). An integrated addition and interaction model for assessing toxicity of chemical mixtures. *Toxicol Sci* 87:520–528.

Riegelman RK (1981). *Studying a study and testing a test*. Boston: Little, Brown and Company.

RIVM (2009). Relative oral bioavailability of lead from Dutch made grounds. RIVM report 711701086/2009. Rijksinstitut voor Volksgezondheid en Milieu (National Institute for Public Health and the Environment), Netherlands.

Roberts RG (1992). Overview of similarities and differences between children and adults: Implications for risk assessment. In PS Guzelian, CJ Henry, SS Olin (Eds), *Similarities and differences between children and adult: implications for risk assessment*. Washington, DC: International Life Sciences Institute.

Robson MG and Toscano WA Eds (2007). *Risk assessment for public health*. New York: John Wiley & Sons.

Rodricks, JV (2007). 2006 American College of Toxicology Distinguished Service Award Lecture: Has quantitative risk assessment been of benefit to the science of toxicology. *Int. J Toxicol*. 26: 3–12.

Rogers AE, Sanchez O, Feinsod FM and Newberne PM (1974). Dietary enhancement of nitrosamine carcinogenesis. Cancer Res 34:96–99.

Rook JJ (1974). Formation of haloforms during chlorination of natural waters. *Water Treatment and Examination*. 23:234–243.

Roseman JM (1998). What can be done to ensure the usefulness of epidemiological data for risk assessment purposes? *Human and Ecological Risk Assessment*. 4:737–746.

Ross MH and Bras G (1965). Tumour incidence patterns and nutrition in the rat. *J Nutr* 87:245–260.

Rothman KJ and Greenland S (1997). Causation and causal inference. In Detels et al. (Eds) *Oxford Textbook of Public Health*. 3rd edn. New York: Oxford Medical Publications.

Roubicek CB, Pahnish OF and Taylor RL (1964). Growth of rats at two temperatures. *Growth* 28:157–164.

Ruden C and Hansson SO (2010). Registration, evaluation, and authorization of chemicals (REACH) is but the first step: How far will it take us? Six further steps to improve the European chemicals legislation. *Environ Health Perspect* 118:6–10.

Rudwaleit M, Richter S, Braun J and Sieper J (2001). Low incidence of reactive arthritis in children following a salmonella outbreak. *Ann. Rheumat. Dis*. 60:1055–1057.

Ruffle B, Burmaster DE, Anderson PD and Gordon HD (1994). Lognormal distributions for fish consumption by the general US population. *Risk Analysis*, 14:395–404.

Rushton L (2000). Reporting of occupational and environmental research: use and misuse of statistical and epidemiological methods. *Occupational and Environmental Medicine*, 57:1–9.

Ryan PB, Burke TA, Cohen Hubal EA, Cura JJ and McKone TE (2007). Using biomarkers to inform cumulative risk assessment. *Environ Health Perspect* 115:833–840.

Ryker SJ and Small MJ (2008). Combining occurrence and toxicity information to identify priorities for drinking-water mixture research. *Risk Analysis* 28:653–666.

Salem H (Ed) (1987). *Inhalation toxicology: Research methods, applications and evaluation*. New York and Basel: Marcel Dekker.

Samet JM, Schnatter R and Gibb H (1998). Invited commentary: Epidemiology and risk assessment. *American Journal of Epidemiology*.

Sand S, Falk-Filipsson A and Victorin K (2002). Evaluation of the benchmark dose method for dichotomous data: Model dependence and model selection. *Regul Toxicol Pharmacol* 36:184–197.

Sand S, Victorin K and Filipsson AF (2008). The current state of knowledge on the use of the benchmark dose concept in risk assessment. *J Appl. Toxicol*. 28:405–421.

Sand S, von Rosen D, Victorin K and Falk-Filipsson A (2006). Identification of a critical dose level for risk assessment: developments in benchmark dose analysis of continuous endpoints. *Toxicol Sci* 90:241–251.

Sand SJ, von Rosen D and Falk-Filipsson A (2003). Benchmark calculations in risk assessment using continuous dose–response information: The influence of variance and the determination of a cut-off value. *Risk Analysis* 23:1059–1068.

Sanderson DM and Earnshaw CG (1991). Computer prediction of possible toxicity from chemical structure; the DEREK system. *Human Experimental Toxicol* 10:261–273.

Sandman, P (1993). *Responding to community outrage: strategies for effective risk communication*, American Industrial Hygiene Association, Fairfax, Va.

Santillo D and Johnston P (2006). Effect thresholds and 'adequate control' of risks: The fatal flaws in the EU council's position on authorisation within REACH. *Environ Sci Pollut Res Int* 13:425–31.

Saunders LD, Chen W and Hrudey SE (1997). Studies of hazardous waste sites and human health: evaluation of internal validity. *Environmental Review*. 5:167–180.

Schaafsma G, Kroese ED, Tielemans ELJP, van Sandt JJM and van Leeuwen CJ (2009). REACH, non-testing approaches and the urgent need for a change in mind set. *Regul Toxicol Pharmacol* 53:70–80.

Scherer G. 2005. Biomonitoring of inhaled complex mixtures: Ambient air, diesel exhaust and cigarette smoke. *Experimental and Toxicologic Pathology* 57:75–110.

Schreider SJ, Barrow C, Birchfield N, Dearfield K, Devlin D, Henry S, Kramer M, Schappelle S, Soloman K, Weed DL and Embry MR (2010). Enhancing the credibility of decisions based on scientific conclusions: Transparency is imperative. *Tox Sci* 116: 5–7.

Seed J, Carney EW, Corley RA, Crofton KM, DeSesso JM, Foster PM, Kavlock R, Kimmel G, Klaunig J, Meek ME, Preston RJ, Slikker W, Tabacova S and Williams GM (2005). Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. *Crit. Rev Toxicol*. 35:663–672.

Seefeld MD and Peterson RE (1984). Digestible energy and efficiency of feed utilisation in rats treated with 2,3,7,8,-tetrachlorodibenzo-p-dioxin. *Tox. Appl. Pharmacol*. 74:214–222.

Seigneur C, Constantinou E and Permutt T (1992). Uncertainty analysis of health risk estimates. Doc. Number. 2460-009-510. Paper prepared for the Electric Power Research Institute, Palo Alto.

Sellers RS, Morton D, Bindhu M, Roome N, Johnson JK Yano BL, Perry R and Schafer K (2007) Society of Toxicologic Pathology Position Paper: Organ Weight Recommendations for Toxicology Studies. *Toxicologic Pathology*, 35:751–755.

Sexton K and Hattis D (2006). Assessing cumulative health risks from exposure to environmental mixtures: Three fundamental questions. *Environ Health Perspect* 115:825–832.

Sharer K (1977). The effect of chronic underfeeding on organ weights in rats: How to interpret organ weight changes in cases of marked growth retardation in toxicity tests? *Toxicology* 7:45–56.

Sharma RP and Coulombe RA Jr (1996). Pharmacokinetics and risk assessment. In AM Fan and LW Chang) (Eds), *Toxicology and risk assessment*. New York: Marcel Dekker Inc, pp. 81–99.

Silvaggio T and Mattison DR (1994). Setting occupational health standards: Toxicokinetic differences among and between men and women. *J Occup Med*, 36: 849–854.

Silverstein MA (1990), Medical screening, surveillance, and the prevention of occupational disease (Editorial). *J Occup Med* 32:1032–1036.

Simon TW (1999.) Two-dimensional Monte Carlo simulation and beyond: A comparison of several probabilistic risk assessment methods applied to a Superfund site. *Human and Ecological Risk Assessment* 5:823–843.

Simpkins JW, Swenberg JS, Weiss N, Brusick D, Eldridge JC, Stevens JT, Handa RJ, Hovey RC, Plant TM, Pastoor TP and Breckenridge CB (2011). Atrazine and breast cancer: A framework assessment of the toxicological and epidemiological evidence. *Toxicol Sci* 123:441–459.

Singh A, Singh A and Nocerino JM (2007). Technical Guide. EPA/600/R-07/041 (NTIS PB2007-107919)

Slob W (2002). Dose–response modeling of continuous endpoints. *Toxicol Sci* 66:298–312.

Slob W, Mirjam Moerbeek M, Eija Rauniomaa E and Aldert H. Piersma AH (2005). A statistical evaluation of toxicity study designs for the estimation of the benchmark dose in continuous endpoints. *Toxicol Sci* 84:167–185.

Slovic P (2000). *The perception of risk*. London: Earthscan Publications.

Smialowicz RJ, DeVito MJ, Williams WC and Birnbaum LS (2008). Relative potency based on hepatic enzyme induction predicts immunosuppressive effects of a mixture of PCDDS/PCDFS and PCBS. *Toxicol Appl Pharmacol* 227:477–484.

Smith LL, Brent, Robert L, Cohen SM, Doerrer NG, Goodman JI, Greim H, Holsapple MP and Lightfoot RM (2008). Predicting future human and environmental health challenges: The health and environmental sciences institute's scientific mapping exercise. *Critical Rev Toxicol* 38:817–845.

Smith RL (1994). Monte Carlo simulation for human exposure assessment at a Superfund site. *Risk Analysis*, 14(4):433–439.

Snodgrass WR (1992). Physiological and biochemical differences between children and adults as determinance of toxic response to an environmental pollutant. In PS Guzelian, CJ Henry, SS Olin (Eds), *Similarities and differences between children and adults: implications for risk assessment*. Washington: International Life Sciences Institute.

Standards Australia (2009). Australia/New Zealand Standard. *Risk management. Principles and guidelines* AS/NZS ISO 31000:2009.

Starr G, Langley A and Taylor A (2000). *Environmental health risk perception in Australia: A research report to the Commonwealth Department of Health and Aged Care.* Adelaide: Centre for Population Studies in Epidemiology, Department of Human Services.

Stebbing M (2010). Exploring risk: Current issues in health risk perception in Australia. PhD thesis, Monash University.

Stebbing M, Katz E, Priestly B and Abramson M (2008). Cancer cluster in the news: Risk perception, risk communication and the media. *Australian Epidemiologist* 15:19–34.

Suk WA and Olden K (2005). Multidisciplinary research: Strategies for assessing chemical mixtures to reduce risk of exposure and disease. *Human Ecological Risk Assess.* 11:141–151.

Sumi N, Stavrou D, Frohberg H and Jochmann G (1976). The incidence of spontaneous tumours of the central nervous system of Wistar rats. *Arch Toxicol* 35:1–13.

Suter G, Vermeire T, Munns W and Sekizawa J (2005). Framework for the integration of health and ecological risk assessment. *Toxicol Appl. Pharmacol.* 207: Suppl 1 611–616.

Swaim LD, Taylor HW and Jersey GC (1985). The effect of handling techniques on serum ALT activity in mice. *J Appl Toxicol* 5:160–162.

Swenberg JA, Fryar-Tita E, Jeong Y-C, Boysen G, Starr T, Walker VE and Albertini RJ (2008). Biomarkers in toxicology and risk assessment: Informing critical dose–response relationships. *Chem Res. Toxicol.* 21:253–265.

Symons J M (2001). The early history of disinfection by-products: A personal chronicle (Part I). *Environmental Engineer* 37:20–26.

Tal A (1997). Assessing the environmental movement's attitudes towards risk assessment. *Environmental Science and Technology*, 31 (10):470A–476A.

Tannenbaum A (1940). Relationship of body weight to cancer incidence. *Arch Pathol* 30:509–517.

Tarone RE (1982). The use of historical control information in testing for a trend in proportions. *Biometrics* 38:215–220.

Taylor R and Langley A. (1998). Exposure scenarios and exposure settings. National Environmental Health Forum Monographs Soil Series No. 2. 2nd edn. Adelaide: South Australian Health Commission.

Ten Berge WF, Zwart A, Appelman LM. 1986. Concentration-time mortality response relationship of irritant and systemically acting vapours and gases. *J Hazrd Mater* 13:301–309.

Teunis P, Takumi K and Shinigawa K (2004). *Escherischia coli* O157:H7 from outbreak data. *Risk Analysis* 24:401–407.

Teuschler L, Klaunig J, Carney E, Chambers J et al. (2002). Support of science-based decisions concerning the evaluation of the toxicology of mixtures: A new beginning. *Regul Toxicol Pharmacol* 36:34–39.

Teuschler LK (2007). Deciding which chemical mixtures risk assessment methods work best for what mixtures. *Toxicol Appl Pharmacol* 223:139–147.

Thayer KA, Melnick R, Huff J, Burns K and Davis D (2006). Hormesis: A new religion? *Environ Health Perspect.* 114:A632–633.

Thomas SP and Hrudey SE (1997). *Risk of death in Canada: What we know and how we know it.* Edmonton: University of Alberta Press.

Thompson A and Taylor BN (2008). Guide for the use of the International System of Units (SI). National Institute of Standards & Technology (NIST) Special Publication 811.

Thompson GE, Scheel LD and Lowe D (1982). Seasonal alteration in response to stress or physiological change. *Drug and Chem. Toxicol* 5:189–199.

Thornton KW and Paulsen SG. (1998). Can anything significant come out of monitoring? *Human and Ecological Risk Assessment*, 4(4):797–805.

Thurman JD, Moeller RB and Turturro A (1995). Proliferative lesions of the testis in ad-libitum-fed and food-restricted Fischer F-344 and FBNF1 rats. *Lab Animal Sci* 45:635–640.

Thybaud V, Aardema M, Clements J, Dearfirled K, Galloway S, Hayashi M, Jacobsen-Kram D, Kirkland D, MacGregor JT, Marzin D, Ohyama W, Schuler M, Suzuki H and Zeiger E (2007). Strategy for genotoxicity testing: Hazard identification and risk assessment in relation to *in vitro* testing. *Mutation Res.* 627:41–58.

Travis KZ, Pate I, Zoe K and Welsh ZK (2005). The role of the benchmark dose in a regulatory context. *Regul Toxicol Pharmacol* 43:280–291.

Trivedi B (2007). Toxic cocktail. *New Scientist* Sept 44–47.

Trosko JE and Upham BL (2005). The emperor wears no clothes in the field of carcinogen risk assessment: Ignored concepts in cancer risk assessment. *Mutagenesis* 20:81–92.

Tucker MJ (1979). The effect of long-term food restriction on tumours in rodents. *Int J Cancer,* 23:803–807.

Tufte ER (1983). *The visual display of quantitative information.* Cheshire, Connecticut: Graphics Press.

Tufte ER (1990). *Envisioning information.* Cheshire, Connecticut: Graphics Press.

Tuomisto J, Pekkanen J, Kirivanta H, Tukiainen E, Vartiainen T, Viluksela M and Tuomisto JT (2005). Dioxin cancer risk – example of hormesis? *Dose Response* 3:332–341.

Turczynowicz L and Robinson NI (2007). Exposure assessment modeling for volatiles – towards an Australian indoor vapour intrusion model. *J Toxicol Env Hlth Part A* 70:1619–1634.

Tyl RW and Marr MC (1997) In RD Hood (Ed.), *Handbook of developmental toxicology*. Boca Raton, Florida: CRC Press.

Tyson CA and Sawhney DS (1985). *Organ function tests in toxicology evaluation.* Park Ridge, NJ: Noyes Publications.

United States Environmental Protection Agency (US EPA) (1983) Environmental Protection Agency. *Good laboratory practice standards: Toxicology testing.* Fed Reg 48: page 53922, and: Pesticide Programs: Good Laboratory Practice Standards. Final Rule. Fed Reg 46 (No. 230): pp. 53946–53969, Tuesday 11/29/83.

United States Environmental Protection Agency (US EPA) (1984). *Risk assessment and management: Framework for decision making.* EPA 600/9-85-002.

United States Environmental Protection Agency (US EPA) (1988). *7 Cardinal rules of risk communication.* OPA-87-020. Washington: US Environmental Protection Agency.

United States Environmental Protection Agency (US EPA) (1989). *Risk assessment guidance for Superfund. Vol 1. Human health evaluation manual, (Part A).* Interim final. EPA/540/1-89/002. Washington: US Environmental Protection Agency, Office of Emergency and Remedial Response.

United States Environmental Protection Agency (US EPA) (1992). *Guidelines for exposure assessment* notice. Part VI. Federal Register 29 May 1992 57(104):22888–22938.

United States Environmental Protection Agency (US EPA) (1993). *Provisional guidance for quantitative risk assessment of polycyclic aromatic hydrocarbons.* EPA/600/R-93/089. Environment Criteria and Assessment Office, US Environmental Protection Agency, Cincinatti, Ohio.

United States Environmental Protection Agency (US EPA) (1994) *Air emissions models for waste and wastewater* U.S. EPA contract no. 68d10118 November 1994 prepared for U.S. Environmental Protection Agency Office of Air Quality Planning and Standards. EPA-453/R-94-080A.

United States Environmental Protection Agency (US EPA) (1995a) *Guidance for risk characterisation.* US Environmental Protection Agency Science Policy Council. Washington.

United States Environmental Protection Agency (US EPA) (1995b). *The use of the benchmark dose approach in health risk assessment.* Risk Assessment Forum, EPA/630/R-94-007, United States Environmental Protection Agency, Washington DC.

United States Environmental Protection Agency (US EPA) (1996a). *Draft Revision to the guidelines for carcinogenic risk assessment.* Office of Health and Environmental Assessment, Office of Research and Development, United States Environmental Protection Agency, Washington DC.

United States Environmental Protection Agency (US EPA) (1996b). *Guidance for data quality assessment. Practical methods for data analysis.* EPA QA/G-9. EPA/600/R-96/064.

United States Environmental Protection Agency (US EPA) (1997a). *Exposure factor handbook.* National Center for Environmental Assessment. Office of Research & Development.

United States Environmental Protection Agency (US EPA) (1997b). *Guiding principles for Monte Carlo analysis*, Risk Assessment Forum, Washington, DC: US EPA.

United States Environmental Protection Agency (US EPA) (1998a). National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts Notice of Data Availability: Proposed Rule, *Federal Register*. 63:15673–15692.

United States Environmental Protection Agency (US EPA) (1998b). 40 CFR Parts 9, 141, and 142 – Disinfectants and Disinfection Byproducts; Final Rule. National Primary Drinking Water Regulations, *Federal Register*. 63:69389–69476.

United States Environmental Protection Agency (US EPA) (2000a) *Supplementary guidance for conducting health risk assessment of chemical mixtures.* Risk Assessment Forum. pp. 1–209.

United States Environmental Protection Agency (US EPA) (2000b) *Benchmark dose technical guidance document.* EPA/630/R-00/001. Risk Assessment Forum.

United States Environmental Protection Agency (US EPA) (2000c) *Methodology for deriving ambient water quality criteria for the protection of human health. Technical support document: Vol. 1 risk assessment*. Washington, DC, Office of Water, Office of Science and Technology, US Environmental Protection Agency EPA-822-B-00-005.

United States Environmental Protection Agency (US EPA) (2001a). *General principles for performing aggregate exposure and risk assessments*. Office of Pesticide Programs, Office of Prevention, Pesticides and Toxic Substances, Washington, DC: US EPA.

United States Environmental Protection Agency (US EPA) (2001b). *Preliminary cumulative risk assessment of the organophosphorus pesticides*. Office of Pesticide Programs, Office of Prevention, Pesticides, and Toxic Substances, Washington, DC: US EPA.

United States Environmental Protection Agency (US EPA) (2002a). *Guidance on cumulative risk assessment of pesticide chemicals that have a common mechanism of toxicity*. Office of Pesticide Programs, Office of Prevention, Pesticides and Toxic Substances, Washington, DC: US EPA.

United States Environmental Protection Agency (US EPA) (2002b). *OSWER draft guidance for evaluating the vapor intrusion pathway from groundwater and soils (subsurface vapor intrusion guidance)*. EPA530-D-02-004.

United States Environmental Protection Agency (US EPA) (2003a). *Framework for cumulative risk assessment*. Risk assessment forum. US Environmental Protection Agency, Washington, DC. EPA/630/P-02/001F.

United States Environmental Protection Agency (US EPA) (2003b). *Considerations in risk communication: A digest of risk communication as a risk management tool*. United States Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Cincinnati.

United States Environmental Protection Agency (US EPA) (2003c). *Recommendations of the Technical Review Workgroup for Lead: An approach to assessing risks associated with adult exposure to lead in soil*. Technical Working Group for Lead. EPA-540-R-03-001.

United States Environmental Protection Agency (US EPA) (2004a). *An examination of EPA risk assessment principles and practices*. Office of the Science Advisor EPA/100/B-04/001.

United States Environmental Protection Agency (US EPA) (2004b). *Risk assessment guidance for Superfund Vol 1: Human health evaluation manual (Part E Supplemental guidance for dermal risk assessment)*. EPA/540/R/99/005.

United States Environmental Protection Agency (US EPA) (2005a). *Guidelines for carcinogenic risk assessment*. EPA/630/P-03/001B.

United States Environmental Protection Agency (US EPA) (2005b). *Guidance on selecting age groups for monitoring and assessing childhood exposures to environmental contaminants*. EPA/630/P-03/00F.

United States Environmental Protection Agency (US EPA) (2005c). *Supplemental guidance for assessing susceptibility from early-life exposure to carcinogens*. EPA/630/R-03/003F.

United States Environmental Protection Agency (US EPA) (2006a) *A framework for assessing health risks of environmental exposures to children*. EPA/600/R-05/093A.

United States Environmental Protection Agency (US EPA) (2006b). *Approaches for the application of physiologically based pharmacokinetic (PBPK) models and supporting data in risk assessment*. EPA/600/R-05/043F.

United States Environmental Protection Agency (US EPA) (2007a). *Guidance for evaluating the oral bioavailability of metals in soils for use in human health risk assessment*. OSWER 9285.7-80

United States Environmental Protection Agency (US EPA) (2007b). *Users guide for the integrated exposure uptake biokinetic model for lead in children (IEUBK)*. EPA 9285.7-42.

United States Environmental Protection Agency (US EPA) (2007c). *Considerations for developing a dosimetry-based cumulative risk assessment approach for mixtures of environmental contaminants*. National Center for Environmental Assessment. Office of Research and Development. US Environmental Protection Agency. Cincinatti. EPA/600/R-07/064.

United States Environmental Protection Agency (US EPA) (2009a). *Risk assessment guidance for Superfund Vol 1: Human health evaluation manual (Part F Supplemental guidance for inhalation risk assessment)*. EPA-540-R-070-002.

United States Environmental Protection Agency (US EPA) (2009b). *Exposure factors handbook: 2009 update*. EPA/600/R-09/052A.

United States Environmental Protection Agency (US EPA) (2009c). *Recommended toxicity equivalency factors (TEFs) for human health risk assessment of dioxin and dioxin-like compounds*. External review draft, released September 2009.

United States Environmental Protection Agency (US EPA) (2010). *Development of a relative potency factor (RPF) approach for polycyclic aromatic hydrocarbon (PAH) mixtures*. In support of summary information on the Integrated Risk Information System (IRIS). EPA/635/R-08/012A.

United States Food and Drug Administration (FDA) (1995). Food additives threshold of regulation for substances used in food-contact articles: final rule. *Federal Register* 60(136):36582–36594.

Vahter M (Ed.) (1982). *Assessment of human exposure to lead and cadmium through biological monitoring*, National Swedish Institute of Environmental Medicine and Karolinska Institute, Stockholm.

Vainio H and Wilbourn J (1992). Identification of carcinogens within the IARC Monograph program. *Scand. J Work Environ. Health*, 18 (Suppl 1):64–73.

Van den Berg M, Birnbaum LS, Denison M, De Vito M, Farland W, Feeley M, Fiedler H et al. (2006). The 2005 World Health Organization re-evaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. *Regul Toxicol Pharmacol*. 45:9–23.

Vose D (2008) *Risk analysis. A quantitative guide*. 3rd edn. John Wiley & Sons.

Walach JB (1996). Interpretation of diagnostic tests. 6th edn. Little Brown, Boston.

Wallace LA and Ott WR (1982). Personal monitors: A state-of-the-art survey. *J Air Pollut Control Assoc*. 32, 601.

Walker NJ, Crockett PW, Nyska A, Brix AE, Jokinen MP, Sells DM, Hailey JR, Easterling M, Haseman JK, Yin M, Wyde ME, Bucher JR and Portier CJ (2005). Dose-additive carcinogenicity of a defined mixture of 'dioxin-like compounds'. *Environ Health Perspect* 113:43–48.

Ward J (2007). The two-year rodent carcinogenesis bioassay: Will it survive? *J Toxicol Pathol*. 20:13–19.

Water Services Association of Australia (WSAA) (2004). *Health risk assessment of firefighting from recycled water mains*. Occasional Paper No. 11, WSAA.

Waters M and Jackson M (2008). Databases applicable to quantitative hazard/risk assessment: Towards a predictive systems toxicology. *Toxicol Appl Pharmacol* 233:34–44.

Weed DL. (2005). Weight of evidence: A review of concept and methods. *Risk Anal* 25:1545–1557.

Weil CS (1970). Significance of organ-weight changes in food safety evaluation. In *Metabolic aspects of food safety*. pp. 419–454, New York: Academic Press.

Weil SC and McCollister DD (1963). Relationship between short- and long-term feeding studies in designing an effective toxicity test. *Agr Food Chem* 11:486–491.

Weingand K, Brown G, Hall R, Davies D, Gossett K, Neptun D, Waner T, Matsuzawa T, Salemink P, Froelke W, Provost JP, Dal-Negro G, Batchelor J, Nomura M, Groetsch H, Boink A, Kimball J, Woodman D, York M, Fabianson-Johnson E, Lupart M and Melloni E (1996). *Harmonization of animal clinical pathology testing in toxicity and safety studies*. Joint Scientific Committee for International Harmonization of Clinical Pathology Testing.

Western Australia Department of Environment and Conservation (2006). *Contaminated sites management series: community consultation guideline*. Land and Water Quality Branch.

Western Australia Department of Health. (2006) *Public health consultation: A guide for developers*, <http://www.public.health.wa.gov.au/cproot/1503/2/Public_Health_Consultation_Guide.pdf>.

Wexler P (2008) Online toxicology resources in support of risk assessment from the U.S. National Library of Medicine. *Toxicol Appl Pharmacol* 233:63.

White RH, Cote I, Zeise L, Fox M, Dominici F, Burke TA and White PD (2009). State-of-the-science workshop report: Issues and approaches in low-dose response extrapolation for environmental health risk assessment. *Env Health Perspect* 117:283–287.

Williams ES, Panko J and Paustenbach DJ (2009). The European Union's REACH regulation: A review of its history and requirements. *Crit. Rev. Toxicol.* 39:553–575.

Wilson RM, Sigal EA, Bacigalupo CM, Willes RF and Munro IC (2000). Derivation of risk management criteria for chemicals of unknown toxic potency at contaminated sites. *Human and Ecological Risk Assessment* 6:131–139.

Woodall GM (2008). An exposure-response database for detailed toxicity data. *Toxicol Appl Pharmacol* 233:14–16.

Woodall GM and Goldberg RB (2008). Summary of the workshop on the power of aggregated toxicity data. *Toxicol Appl Pharmacol* 233:71–75.

World Health Organization (1999a). *Chemical risk assessment*. Training module No. 3. WHO/PCS/99.2. Geneva: WHO.

World Health Organization (1999b). *Principles for the assessment of risks to human health of chemicals*. Environmental Health Criteria 210. IPCS/WHO. Geneva: WHO.

World Health Organization (WHO) (1978). *Principles and methods for evaluating the toxicity of chemicals*. Environmental Health Criteria 6. International Programme on Chemical Safety (IPCS)/ WHO. Geneva: WHO.

World Health Organization (WHO) (1987). *Principles for the safety assessment of food additives and contaminants in food*. Environmental Health Criteria 70, IPCS/ WHO. Geneva: WHO.

World Health Organization (WHO) (1990). *Principles for the toxicological assessment of pesticide residues in food*. Environmental Health Criteria 104, IPCS/ WHO. Geneva: WHO.

World Health Organization (WHO) (1991). *Investigating environmental disease outbreaks: A training manual*. Environmental and Occupational Epidemiology series. WHO/PEP/91.35. Geneva: WHO.

World Health Organization (WHO) (1992). *Quality management for chemical safety testing*. Environmental Health Criteria, 141, IPCS/WHO. Geneva: WHO.

World Health Organization (WHO) (1993a). *Guidelines for drinking water quality*. Vol. 1 recommendations. 2nd edn. WHO. Geneva: WHO.

World Health Organization (WHO) (1993b). *Biomarkers and risk assessment: Concepts and principles*. Environmental Health Criteria 155. IPCS/ WHO. Geneva: WHO.

World Health Organization (WHO) (1994a). *Assessing human health risks of chemicals: Derivation of guidance values for health-based exposure limits*. Environmental Health Criteria 170. IPCS/ WHO. Geneva: WHO.

World Health Organization (WHO) (1996). *WHO/FAO/IASEA Report on trace elements in human nutrition and human health*. Geneva: WHO.

World Health Organization (WHO) (1999a). *Chemical risk assessment*. Training Module No. 3 WHO/PCS/99.2. Geneva: WHO.

World Health Organization (WHO) (2000a). *Global burden of disease.* Geneva: WHO.

World Health Organization (WHO) (2000b). *Air quality guidelines for Eur*ope. 2nd edn. WHO regional publications. European series. No. 91.

World Health Organization (WHO) (2001). *Biomarkers in risk assessment: Validity and validation*. Environmental Health Criteria 222. IPCS/WHO. Geneva: WHO.

World Health Organization (WHO) (2006a) (World Health Organization). *Guidelines for drinking-water quality*, 3rd edn, incorporating first addendum, Geneva: WHO. <http://www.who.int/ water_sanitation_health/dwq/gdwq3rev/en/ index.html>.

World Health Organization (WHO) (2006b). *Elemental speciation in human health risk assessment*. Environmental Health Criteria 234. IPCS/WHO. Geneva: WHO.

World Health Organization (WHO) (2006c). *Dermal absorption*. Environmental Health Criteria 235. IPCS/ WHO. Geneva: WHO.

World Health Organization (WHO) (2006d). *Principles for evaluating health risk in children associated with exposure to chemicals*. Environmental Health Criteria 237. IPCS/WHO. Geneva: WHO.

World Health Organization (WHO) (2009). *Risk characterization of microbiological hazards in food*. Guidelines. Microbiological risk assessment series 17. WHO Food and Agriculture Organization of the United Nations. Geneva: WHO.

World Health Organization (WHO) (2010). *WHO guidelines for indoor air quality: Selected pollutants*. WHO Regional Office for Europe.

Wullenweber A, Kroner O, Kohrman M, Maier A, Dourson M, Rak A, Wexler P and Tomljanovic C (2008). Resources for global risk assessment: The International Toxicity Estimates for Risk Assessment (ITER) and Risk Information Exchange (RiskIE) databases. *Toxicol Appl. Pharmacol.* 233:45–53.

Yang RSH and Dennison JE (2007). Initial analyses of the relationship between 'thresholds' of toxicity for individual chemicals and 'interaction thresholds' for chemical mixtures. *Toxicol Appl. Pharmacol.* 223:133–138.

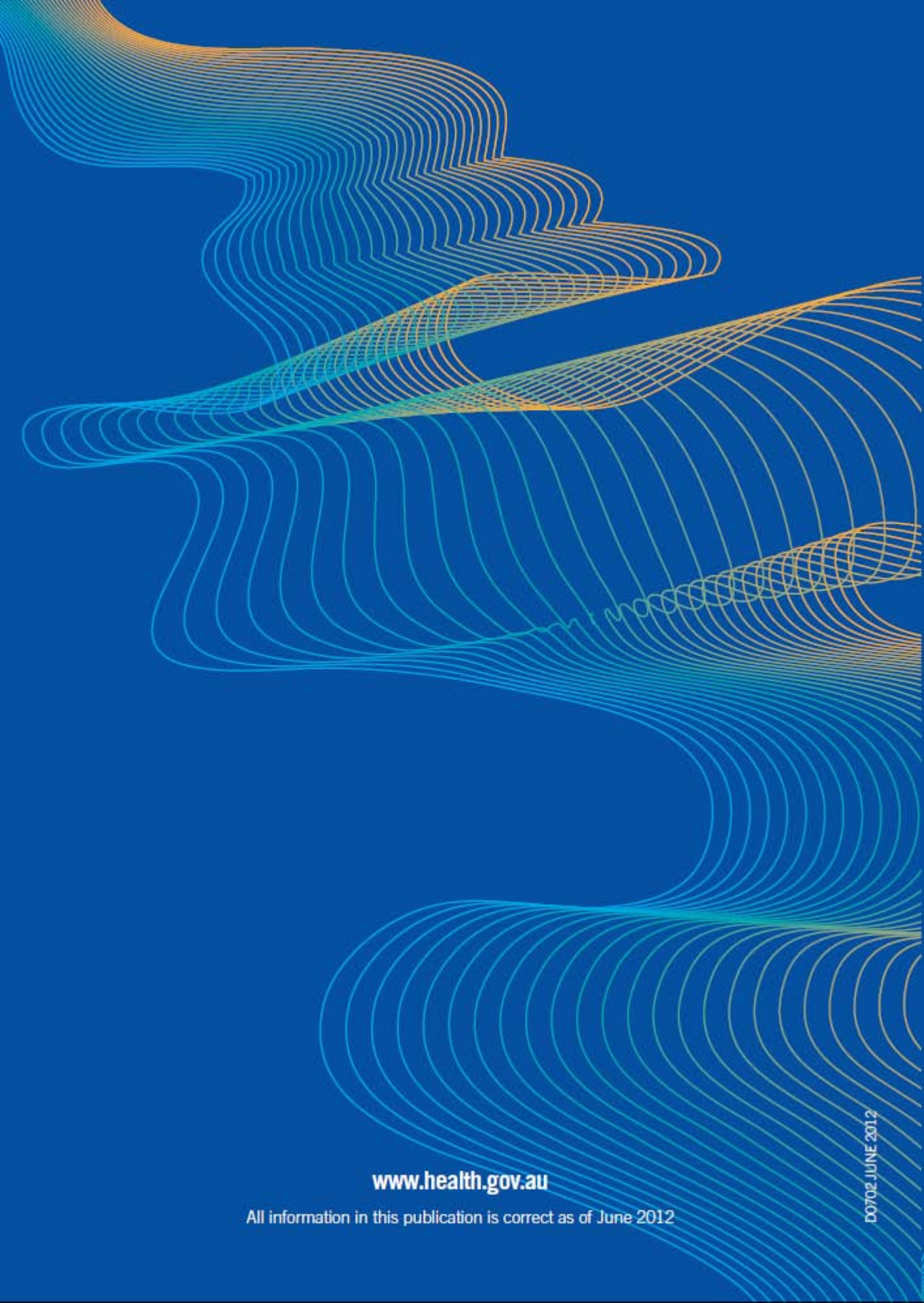
Yassi A, Kjellstrom T, de Kok T and Guidotti T (2001). *Basic environmental health*. World Organization, United Nations Environment Programme, United Nations Education Science and Cultural Organization, Council of Rectors of European Universities.

Zeliger HI (2003). Toxic effects of chemical mixtures. *Arch. Environ. Health* 58:23–29.

Zhang W, Gabos S, Schopflocher D, Li XF, Gati WP and Hrudey SE (2009). Validation of urinary trichloroacetic acid as a biomarker of exposure to drinking water disinfection by-products. *J Water Health* 7: 359–371.

Environmental health risk assessment: – Guidelines for assessing human health risks from environmental hazards

ISBN: 978-1-74241-766-0
Online ISBN: 978-1-74241-767-7
Publications approval number: D0702



www.health.gov.au

All information in this publication is correct as of June 2012

D0702 JUNE 2012